# UNIVERSITÉ DE STRASBOURG

*École doctorale des Sciences de la Vie et de la Santé*
**IGBMC - CNRS UMR 7104 - Inserm U 964**

# THÈSE présentée par :

## Nikolaos Vosnakis

soutenue le : **16 décembre 2014**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

## Investigating the role of human HAT (histone acetyltransferase) containing complexes, ATAC and SAGA, in living cells

**THÈSE dirigée par :**
  **M. Laszlo TORA**    Directeur de recherche, université de Strasbourg

**RAPPORTEURS :**
  **M. Edouard BERTRAND**    Chargé de Recherches, Institut de Génétique Moléculaire de Montpellier
  **M. Marc TIMMERS**    Professeur, University Medical center d'Utrecht

**AUTRES MEMBRES DU JURY :**
  **Mme Evi SOUTOGLOU**    Chargé de Recherches, Université de Strasbourg

# *Acknowledgements*

A special thanks goes to all the good friends from Strasbourg that tolerate(d) me and made these four years special. It would be a long list to mention. Thank you for the fun, memorable moments and the multidimensional support.

Finally, I would like thank my family for always being there for me.

# ABSTRACT

SAGA and ATAC are multisubunit Histone Acetyl-Transferase (HAT) containing complexes that facilitate RNA polymerase II (Pol II) transcription by altering the chromatin state via post-translational modifications of histones. ATAC and SAGA share a set of subunits among which GCN5, that exhibits different catalytic properties in the context of each complex. Many studies have described the roles of the complexes in gene-specific transcription regulation and provided information about their genome wide distribution and function. However, little is known about the dynamic properties of the complexes in living cells. In addition, structural and functional studies suggest a modular organisation of the complexes. Nevertheless, there is little information on the regulation of their assembly and how the common subunits are differentially incorporated in the two complexes. The first goal of my Ph.D. project was to investigate the intranuclear dynamics of human ATAC and SAGA complexes in living cells and study how it can be related to transcription regulation. I used an experimental approach based on live-cell imaging methods such as FRAP (Fluorescence Recovery After Photobleaching) and FCS (Fluorescence Correlation Spectroscopy). I exploited these two techniques, to characterise the mobility of the two complexes and compare their dynamic properties with other important actors of Pol II transcription. I showed that all tested ATAC and SAGA subunits are highly mobile in the nucleus of living cells and they only exhibit very transient interactions with chromatin. Moreover, I investigated the existence of a possible connection between active transcription and the dynamics of the two complexes. Thus, the data collected from this set of experiments show, for the first time in living cells, that the very dynamic behaviour of the ATAC and SAGA is a key property of both complexes which explains certain aspects of their function as chromatin modifiers. During my Ph.D. studies, I also became interested in the mechanisms that regulate the dynamic properties of SAGA and ATAC subunits related to the assembly of the complexes. This question was investigated by a combination of imaging experiments (in fixed and living cells) and quantitative proteomics. FLIP (Fluorescence Loss In Photobleaching), showed that exogenously overexpressed ATAC- and SAGA-specific HAT-module subunits (ADA2a and ADA2b respectively) differ significantly in their subcellular localisation dynamics. In addition, I showed that the relative abundance of GCN5 and ADA2a affects the subcellular distribution of the latter protein. With the application of quantitative proteomic analysis, based on MudPIT (Multidimensional Protein Identification Technology), the findings of the imaging experiments were expanded on endogenous proteins and provided strong evidence that the cytoplasmic and nuclear assembly pathways of SAGA and ATAC complexes are different. Altogether, our data contributed to our understanding of the way ATAC and SAGA exhibit their functions and revealed novel protein-protein interaction related to the assembly of the complexes.

# Table of contents

# Table of Figures

## Table of Tables

# 1 Introduction

The availability of the right type of proteins at the right amounts is a fundamental prerequisite of homeostasis for all living organisms. In eukaryotes, and particularly in metazoans, protein abundance mainly depends on the tight regulation of mRNA production. Although some genes are constitutively expressed, the expression level of others has to be adjusted depending on the differentiation stage of a cell and the conditions of the environment. The amount of mRNA and the corresponding protein produced from a gene can be also changed by regulatory pathways related to the processing, the stability or the transport of the transcript in the cytoplasm. In addition, there are cellular mechanisms that regulate translation efficiency and protein stability. However, the most common strategy for gene expression regulation involves changes of the transcription rate of a gene.

Transcription of most protein encoding genes in eukaryotes is performed by RNA polymerase II (Pol II). Pol II function must be highly regulated so that it transcribes genes in a very specific spatiotemporal manner dictated by cell cycle, metabolic or tissue specific needs of a cell. Moreover, the molecular pathways regulating the function of Pol II must be sensitive enough to respond rapidly in changes and signals of the extra or intra-cellular environment. Such type of regulation requires the interaction between sequence-specific DNA-binding factors, general transcription factors (GTFs), co-activators, the epigenetic state of target sequences and Pol II.

In the beginning of this chapter, some of the main steps of the above interactions will be summarised to provide the general functional insights on the role of factors that are particularly relevant to this study. Second, a more focused selection of information regarding the role of the SAGA co-activator complex in transcription regulation will be described. In addition, I will summarize some of the information available for the ATAC co-activator complex in response to stress and other pathways. Moreover, I will discuss how Fluorescence Recovery After Photobleaching (FRAP) and Fluorescence Correlation Spectroscopy (FCS) have been applied to investigate the dynamics of transcription factors in living cells and summarise important findings and implications of such studies. Last but not least, as many of the transcription related complexes are composed by multiple subunits, I will also discuss the findings and the implications of studies investigating different aspects of complex assembly regulation in living cells.

## 1.1 Core promoter elements in Pol II transcription initiation

Pol II transcription can be divided in three phases: i) initiation, that involves the recruitment of Pol II at the promoter of a target gene by a set of proteins, called general (or basal) transcription factors (GTFs) as well as the Mediator complex; ii) elongation, during which the RNA transcript is synthesized by the polymerase, and iii) termination, that results in the release of both Pol II and the RNA product from the transcribed DNA template.

Initiation is considered to be the liming step of eukaryotic transcription. Particularly, it is the formation of a productive preinitiation complex (PIC), consisting of Pol II and the general transcription factors allowing basal levels of transcription, that can be rate limiting. It has been shown that both the sequence and the chromatin environment of the core promoter of a gene both serve as a platform for the recruitment of GTFs and the nucleation of the PIC (Thomas & Chiang, 2006; Muller & Tora, 2014).

One could assume that since the efficiency of PIC formation seems to be a universal mechanism of Pol II transcription initiation regulation, the core promoter sequences that serve as PIC assembly sites would also be conserved. Indeed, Pol II core promoters may contain characteristic, highly conserved DNA motifs. Highly conserved core promoter elements are the TATA box, that is recognised by TBP (TATA-binding protein) a component of the TFIID (Transcription Factor II D) complex and BRE, recognised by TFIIB (Transcription Factor II B)(Deng & Roberts, 2007; Fuda *et al*, 2009). In total, seven distinct types of core promoter elements have been identified in eukaryotes. It must be noted that these core promoter elements are not present in every core promoter. There are core promoters that lack any of the identified conserved motifs indicating that various other sequences and chromatin features are involved Pol II initiation regulation (Thomas & Chiang, 2006; Sandelin *et al*, 2007; Juven-Gershon *et al*, 2008; Fuda *et al*, 2009).

Multiple core promoter elements can be present upstream or downstream of the TSS (Transcription Start Site) of a transcription unit. The combinatorial effect of these sequences, exhibited by the recruitment of different sets of GTFs, in combination with interaction with distant regulatory sites (i.e. enhancers) results in the very accurate but also diverse transcription initiation regulation (Thomas & Chiang, 2006; Fuda *et al*, 2009).

## 1.2   RNA pol II: a structure-function paradigm

RNA polymerase II (RNA pol II/ Pol II) is the multisubunit complex that is responsible for the transcription of most eukaryotic protein-coding genes. It interacts with the general transcription factors the template DNA and multiple other factors that orchestrate its function during transcription initiation, elongation and termination. The characterisation of Pol II structure has provided a better understanding of the way those interactions are achieved.

Yeast and human Pol II are composed of 12 highly conserved subunits (in human RPB1 to RPB12) (Edwards *et al*, 1991; Young, 1991). Among those, RPB5, 6, 8, 10 and 12 are shared between the other two classes of eukaryotic DNA dependent RNA polymerases (RNA pol I and RNA pol III). On the other hand, RPB4, 7, 9 and the unstructured Carboxy-Terminal Domain (CTD) extension of RPB1 are specific to RNA pol II (Woychik *et al*, 1990; Carles *et al*, 1991; Young, 1991; Hampsey, 1998; Wild & Cramer, 2012). It is RPB1 together with RPB2 that form the active site of the enzyme and bear the catalytic activity of the complex (Lee & Young, 2000).

X-ray structure analysis of the *S. cerevisiae* Pol II holocomplex has shown that Rpb1 and Rpb2 are at the centre of the complex forming the active site. Rpb5, Rpb6 and Rpb8 bind directly to Rpb1 whereas Rpb9 binds to Rbp2. Rpb7 links Rpb4 to Rpb6. The remaining subunits (Rpb3, Rpb10, Rpb11, and Rpb12) create a bridge that links Rpb1 to Rpb2 (Figure 1.1, A). Four mobile modules termed core, clamp, shelf and jaw-lobe, have been described at the central part of the complex (Figure 1.1, B). The active site is found at the base of a cleft at the centre of the complex. The clamp module has a broad range of movement that can swing over the active centre region and holds the DNA-RNA hybrid in place. It is has been suggested that after interactions of promoter DNA with GTFs (especially TFIIB), only the template strand of DNA can slip in the cleft to be scanned for a transcription start site (Cramer, 2000; Cramer *et al*, 2001; Kostrewa *et al*, 2009).
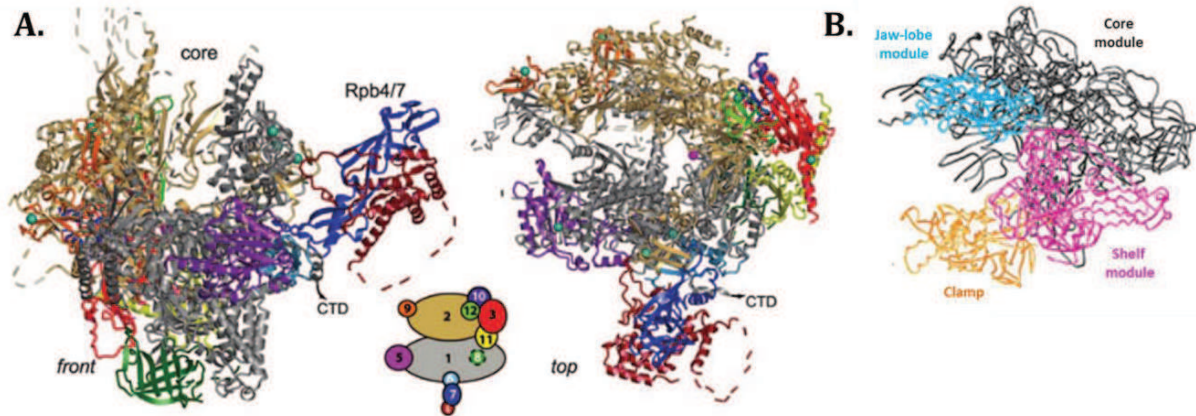
**Figure 1.1. RNA polymerase II structure and the localisation of its four mobile modules.**

**A.** Front (left) and top (right) view of the complex shown as a ribbon diagram. The 12 subunits Rpb1-Rpb12 are coloured according to the key between the views. Dashed lines represent disordered loops. **B.** Backbone traces of the four mobile modules of the Pol II structure core. Jaw-lobe, clamp, and shelf modules structure shown in grey, blue, yellow and pink, respectively. Modified from (Cramer *et al*, 2001) and (Armache *et al*, 2005).

Pol II exhibits a broad range of interactions with different sets of factors at each phase of transcription (initiation, elongation, termination). For efficient PIC formation, Pol II has to directly or indirectly interact with the GTFs. However, transcription initiation and start of RNA synthesis require disengaging of Pol II from the promoter (promoter clearance) and association with different sets of factors that regulate transcription elongation and termination. Functional and structural studies have shown that the CTD of RPB1, the largest subunit of Pol II, is particularly important for the regulation of these interactions and pre-mRNA processing.

As mentioned above the CTD of RPB1 is an unstructured sequence that in human cells contains 52 tandem repeats of the consensus heptapeptide Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS). The CTD is a dynamic structure that functions as a scaffold for the interaction with regulators of Pol II function. Two functionally distinct forms of Pol II can be distinguished depending on the level of CTD phosphorylation. Hypophosphorylated CTD is characteristic of the IIA form of Pol II which is related to early steps of transcription initiation and Pol II recruitment to the PIC (Lu *et al*, 1991; Serizawa *et al*, 1993). During transcription elongation and termination, a cycle of phosphorylation and dephosphorylation of S2 and S5 result in a hyperphosphorylated CTD that characterises the IIO form of Pol II. Specifically, phosphorylation of S5 by cyclin-dependent kinase 7 (CDK7), associated with the GTF TFIIH, is a signal for PIC destabilisation and is a key step for the transition from initiation to elongation (Serizawa *et al*, 1993). Subsequently, CDK9 present in positive transcription elongation factor b (P-TEFb) phosphorylates S2, required for productive elongation and recruitment of mRNA 3'-end processing factors (Komarnitsky *et al*,

2000; Ahn *et al*, 2004). Apart from directly targeting CTD serine 2, CDK9 in the context of P-TEFb has another important role that regulates productive elongation. It has been shown that at many metazoan genes, right after initiation and the synthesis of a short (20-65 nt) RNA molecule, Pol II is poised at promoter proximal positions (Muse *et al*, 2007; Core *et al*, 2008; Nechaev *et al*, 2010; Rahl *et al*, 2010). This paused Pol II form is promoted by the interaction of DRB sensitivity inducing factor (DSIF) and Negative Elongation Factor (NELF) that inhibit further elongation. Phosphorylation of these two factors by P-TEFb results in Pol II release and allows productive elongation (Peterlin & Price, 2006; Cheng & Price, 2007).

The level of S2 and S5 phosphorylation is also regulated by protein phosphatases, some of which have been shown to interact with GTFs. For example, TFIIF-associated CTD phosphatase 1 (FCP1) interacts with TFIIB (Chambers *et al*, 1995), RPB4 (Kimura *et al*, 2002b), and TFIIF subunit RAP74 (Chambers *et al*, 1995; Kobor *et al*, 2000), and dephosphorylates S2. Interestingly, the interaction TFIIF with FCP1 has been shown to enhance S2 dephosphorylation suggesting that this stimulation may regulate the transition of Pol II from the IIO (elongating) to the IIA (initiating) form (Chambers *et al*, 1995). Interestingly, Ssu72 integral component of the CPF mRNA 3′-end processing complex, has been shown to interact with general transcription factors and affects Ser5-P and Ser7-P depending on the orientation relative to the backbone polarity of the CTD (Rosado-Lugo & Hampsey, 2014). The so called CTD phospho-code has been expanded with studies showing that Tyr1-P, Thr4-P, and Ser7-P are also phosphorylated at distinct steps of RNA pol II transcription underlining the regulatory plasticity of this structure (Heidemann *et al*, 2013).

Apart from phosphorylation, the CTD can be extensively modified by glycosylation. Glycosylation occurs mainly on the fourth position of the heptapeptide with the glycosylated residues being mainly a threonine but also a serine (in non-consensus sequences) (Kelly *et al*, 1993). Since glycosylation has been found to be mutually exclusive to phosphorylation (Comer & Hart, 2001; Iyer *et al*, 2003) it has been suggested that it could prevent kinases from phosphorylating the CTD during early initiation (Thomas & Chiang, 2006). Lastly, extensive Pol II ubiquitination has been linked to DNA damage signalling pathways (Woudstra *et al*, 2002; Bucheli & Sweder, 2004; Gillette *et al*, 2004). Arrested Pol II at sites of DNA lesions gets polyubiquitinated and among other functions it may act as a signal for the recruitment of nucleotide excision repair (NER) factors (Bregman *et al*, 1996).

## 1.3 General Transcription Factors and Pre-Initiation Complex (PIC) formation

Accurate, site-specific recruitment and engagement of RNA pol II in transcription necessitates the presence of a specific set of factors, which are collectively called general transcription factors (GTFs). These factors are Transcription Factor (TF) IIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. They were initially identified as enzymatically active fractions of subcellular extracts under the presence of which, purified Pol II could transcribe *in vitro* a DNA template (Weil *et al*, 1979). Among the GTFs there are two large multisubunit complexes: TFIID that composed of TBP plus 13 TBP Associated Factors (TAFs) and TFIIH that has10 subunits. TFIIB is the only GTFs that is a single polypeptide (Thomas & Chiang, 2006).

### 1.3.1 General Transcription Factors

**TFIID**

TFIID is a multisubunit complex that was identified as a factor required for RNA pol II transcription initiation. It is composed of TBP plus 13 TAFs. TFIID mediates promoter recognition via interactions with elements around the TSS (Fire *et al*, 1984; Young, 1991).

Key component of the complex is TBP which in the sequential assembly model is the factor that in the context of TFIID recognises the promoter to start nucleation of the PIC. TBP is highly conserved in eukaryotes with a size of 339 aa in human. The importance of this factor for the regulation of eukaryotic transcription is also apparent from the fact that apart from TFIID, TBP is also present in the core-promoter binding factors of RNA pol I and RNA pol III, (SL1 and TFIIIB respectively). However, although there is a single type of TBP in yeast, one or two TBP related (TRF1 and TRF2, also called TLF) are present in metazoans. In addition, vertebrates contain a third TRF, called TRF3 or TBP2 (Gazdag *et al*, 2007). Interestingly, in mammals TRF2 and TRF3/TBP2 are involved in the regulation of gonad-specific genes, without the simultaneous presence of TBP (Davidson, 2003; Müller *et al*, 2010)(Gazdag *et al*, 2007). TBP has a bipartite structure which is highly conserved particularly at the carboxy-terminal part, which interacts directly with promoter sequences inducing a 90° bending of DNA. This, results in the formation of an asymmetric platform that facilitates PIC assembly (Kim *et al*, 1993; Kim & Burley, 1994; Nikolov *et al*, 1996).
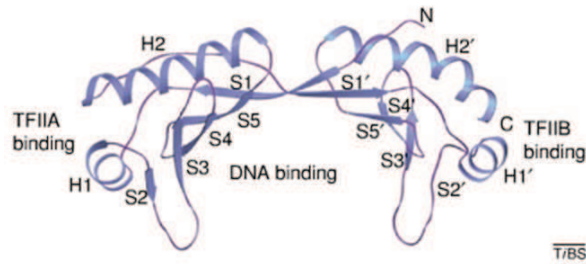
**Figure 1.2. 3D structure of the TATA-binding protein (TBP) core domain**
Ribbon representation of the 3D structure of the TATA-binding protein (TBP) core domain
determined by X-ray crystallography. The regions of TBP contacted by transcription factor
TFIIA, TFIIB and the DNA are indicated. Adapted from (Davidson, 2003)

DNA-TBP interactions can be highly regulated in a positive or negative way. Negative regulation
is required as it has been suggested that non-specific binding to A/T rich regions may lead to the
formation of non-productive transcription complexes (Coleman & Pugh, 1995). Several protein-
protein interactions have been involved in this type of regulation which inhibits promoter
recognition and/or TBP interactions with TFIIA and TFIIB. Association of TBP with the amino-
terminal region of BTAF1 in the B-TFIID complex blocks TBP-promoter binding (Pereira *et al*,
2001). In addition, the carboxy-terminal region of BTAF1 disrupts TBP-DNA complex formation
in an ATP dependent manner (Auble *et al*, 1997; Chicca *et al*, 1998; Pereira *et al*, 2001; Andrau *et
al*, 2002). It has to be noted that apart from their inhibitory role on TBP activity, these functions
of BTAF1 facilitate the release of TBP from non-specific sites, thus allowing its redistribution to
actual promoter regions (Collart, 1996; Li *et al*, 1999; Geisberg *et al*, 2002). Another mechanism
of negative regulation is based on the affinity of specific domains of TAF1 (the largest TFIID
subunit) with functional regions of TBP. Interaction of the two factors can block TBP-DNA
complex formation but can also compete for TFIIA binding (Kokubo *et al*, 1993; Kokubo *et al*,
1998; Liu *et al*, 1998). Similarly, negative cofactor 2 (NC2) can inhibit the interaction of TBP with
TFIIA and TFIIB by blocking the respective interaction surfaces of TBP. The complexity of TBP
activity regulation is elevated from the interaction between NC2 and BTAF1. Particularly, the
NC2α but bot NC2β component of human NC2, was found associated with BTAF1. Interestingly,
it was shown that this interaction promotes BTAF1 association with TBP (Klejman *et al*, 2005).
The interplay between these factors, together with the potential of TBP to form homodimers
unable to bind DNA (Coleman *et al*, 1995; Jackson-Fisher *et al*, 1999), illustrates the complex
network of protein interactions that regulate TBP/TFIID mediated promoter recognition. Two
GTFs that promote TBP-DNA complex formation are TFIIA and TFIIB. There are three major
mechanisms by which TFIIA enhances TBP activity. The first include the dissociation of TBP
dimers that increase the availability of monomeric TBP (Coleman *et al*, 1999). Second, TFIIA
competes with the negative regulator TAF1 for their common binding regions on TBP (Kokubo

*et al*, 1998). Third, its association with TBP-promoter complex prevents BTAF1 from exhibiting its destabilising role (Auble & Hahn, 1993). TFIIB has also similar effects as it enhances and stabilizes the formation of TBP-DNA complex (Imbalzano *et al*, 1994a; Zhao & Herr, 2002).

TBP association with TAFs forms the TFIID complex that recognises a broader range of sequences. This binding spectrum is particularly important since a recognisable TATA element is absent from the majority of vertebrate promoters. Different TAFs mediate the recognition of distinct promoter elements expanding possible TFIID-DNA contacts. Thus the presence of TAFs is essential for transcription from these promoters that lack canonical TATA-box (Martinez *et al*, 1994; Huisinga & Pugh, 2004). Moreover the dynamic association of TAFs in TFIID has been suggested to provide increased potential for cell type specific transcription initiation regulation. This is particularly important for metazoans and the regulation of gene expression at different developmental stages (Müller *et al*, 2010).

Additionally, the fact that TAFs have been found to interact with transcriptional activators suggests that TFIID has a co-activator function. Interestingly, it has been demonstrated that different domains of the same activator can interact with different TAFs or that the same TAF can interact with many activators (Gill *et al*, 1994; Chiang & Roeder, 1995; Verrijzer & Tjian, 1996; Rojo-Niersbach *et al*, 1999; Asahara *et al*, 2001). Thus, activator mediated TFIID recruitment to TATA-less promoter elements can regulate productive PIC assembly at these sites. TAFs were also found to interact with several GTFs like TFIIA (Yokomori *et al*, 1993), TFIIE (Ruppert & Tjian, 1995), TFIIB (Goodrich *et al*, 1993), but also with the catalytic subunits of Pol II (Wu & Chiang, 2001). The extensive network of protein contacts that TAFs are involved could play a key role in facilitating the nucleation of PIC in different cellular background.

Apart from the interactions with activators, TAFs can direct the recruitment of TFIID to promoters via interactions with covalently modified histones. It is well established that H3K4m3 and H3K9ac are enriched at active promoters (Wang *et al*, 2008b). Certain TAFs contain structural domains which interact with histone modifications. Particularly, TAF1 has a double bromodomain that specifically recognises H4K5/12 of H4K8/16 but also H4K16 with a lower affinity (Jacobson *et al*, 2000). As these marks are found in nucleosomes within actively transcribed chromatin regions, TAF1 could act as reader that recruits TFIID at active promoter sites. In addition, a chromatin interacting domain that has been found at the C-terminal region of TAF3 is the plant homeodomain (PHD). Proteomic screening based on stable isotope labelling by amino acids in cell culture (SILAC) showed that TAF3 binds specifically, with high affinity to H3K4m3 via the PHD domain. Interestingly, the simultaneous presence of acetylated lysines

H3K9 and H3K14 further enhanced the observed crosstalk between H3K4m3 and TFIID (Vermeulen *et al*, 2007).

Overall the broad range of TAF mediated interactions summarised above, allow the regulation of recruitment and stabilisation of TFIID at active promoter without the requirement of a canonical TATA element.

**TFIIA**

Human TFIIA contains three subunits (α, β and γ). Early studies showed that TFIIA is essential for transcription (Reinberg & Roeder, 1987), while later, it was shown that basal transcription does not necessitate its presence (Van Dyke *et al*, 1988; Wu & Chiang, 1998). These contradictions on the role of TFIIA probably result from the differences of complexity in the systems by which transcription efficiency was studied (partially or highly purified system). As mentioned above, it is believed that TFIIA acts as an antirepressor of inhibitors, that are present for example in a partially purified system, by facilitating the interaction of TFIID, TBP and DNA (Imbalzano *et al*, 1994b; Kang *et al*, 1995). There are two particularly important functions of TFIIA related with PIC formation. First, it promotes the dissociation of TBP dimmers increasing the efficiency of promoter elements recognition (Coleman *et al*, 1999). Second, it competes with TAF1/BTAF1 inhibition, stabilising the formation of the PIC (Auble & Hahn, 1993; Kokubo *et al*, 1998). In addition, the interactions of TFIIA with several components of TFIID (i.e. TAF1, TAF4, TAF11, TBP), activators, and GTFs (e.g. TFIIE) indicate that it could also have coactivator function (Kobayashi *et al*, 1995; Langelier *et al*, 2001; Solow *et al*, 2001; Dion & Coulombe, 2003).

**TFIIB**

TFIIB is composed of 316 aminoacids in human (345 in yeast) with 5 functional domains along its sequence. It is a key factor for the recruitment of Pol II to the preinitiation complex. The interaction between its amino-terminal zinc ribbon domain (B-ribbon) is involved in the recruitment of RNA pol II to the promoter and its carboxy-terminal domain (B-core) binds TBP and DNA at the core promoter (Nikolov *et al*, 1995; Kostrewa *et al*, 2009) (Figure 1.3). Its B-linker domain plays a role in TSS selection whereas the B-reader domain contributes to the opening of DNA at the promoter. In addition TFIIB interacts directly with RAP30 and RAP74 (subunits of TFIIF) and TAF9 (Ha *et al*, 1993; Fang & Burton, 1996; Deng & Roberts, 2007).
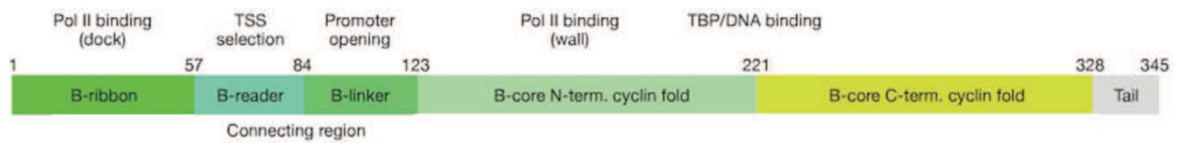
**Figure 1.3. Functional domains on yeast TFIIB sequence.**
Organization of the distinct functional domains of TFIIB along its sequence. Known interactions and transcription initiation steps related to each domain are mentioned above the scheme. Adapted from (Kostrewa *et al*, 2009).

According to the proposed model for the role of TFIIB in initiation-elongation transition, it is the B-ribbon that recruits promoter DNA to Pol II and positions the active site above the cleft of the polymerase, resulting in the closed complex conformation (Kostrewa *et al*, 2009). Subsequently, the B-linker facilitates the melting of 20 bp downstream of the TATA box and the template strand can slip in the cleft and be positioned in the template tunnel. The role of B-reader is important at this point to stabilise the generated bubble near the active centre. The transition to the open complex formation is completed by the loading of the DNA duplex into the downstream cleft. Next, DNA template strand is scanned for a start site (Inr motif), a process in which again the B-reader is involved. The RNA chain synthesis starts with the first phosphodiester bond formation when the first two nucleotides are opposite the Inr motif. The following phase, termed abortive transcription, includes the generation of short DNA-RNA hybrids that are not stably bound, resulting in the ejection of short RNAs. In the final phase (promoter escape), once the synthesised RNA becomes longer than seven nucleotides TFIIB is released and the elongation complex is formed (Kostrewa *et al*, 2009).

**TFIIF**

TFIIF is formed by the heterodimerisation of its two subunits RAP30 and RAP74. It regulates PIC formation at multiple levels. It directly interacts with Pol II facilitating its association with promoter bound TFIID-TFIIB complex (Flores *et al*, 1991). It has been suggested that TFIIF induces topological changes in the promoter DNA that is wrapped around Pol II (Robert *et al*, 1998). Thus, the affinity of Pol II with the TBP-TFIIB complex is increased and the factors assembling into PIC are stabilising at the promoter region. As TFIIF is closely interacting with TFIIB and Pol II, it is thought that it affects the specificity of the selection of the transcription start site (Hampsey, 1998). An additional way by which TFIIF is increasing Pol II transcription specificity is by inhibiting and/or reversing the binding of Pol II at non-specific sites but also functioning as an elongation factor (Cojocaru *et al*, 2008). Lastly, the direct interaction between TFIIF and TFIIE is important for the next step of PIC assembly as it recruits TFIIE and TFIIH (Maxon *et al*, 1994).

**TFIIE**

In human, TFIIE is a heterooctamer of TFIIEα and TFIIEβ. According to the sequential PIC assembly pathway, TFIIE and TFIIH are recruited at the promoter after the assembly of the TFIID-TFIIB-Pol II/ TFIIF complex. TFIIE plays a crucial role in promoter melting as it interacts directly with TFIIB, TFIIF, TFIIH, Pol II and promoter DNA. Particularly, it stimulates all the enzymatic activities of TFIIH (ATPase, CTD kinase, helicase) facilitating the transition from initiation to elongation (Ohkuma *et al*, 1991; Ohkuma & Roeder, 1994; Watanabe *et al*, 2003).

**TFIIH**

As mentioned above, TFIIH is recruited to the promoter together with TFIIE. TFIIH is a complex of 10 subunits that via its multiple enzymatic activities plays an important role both in transcription initiation and in nucleotide excision repair (NER) DNA damage response (Drapkin *et al*, 1994). TFIIH consists of two functionally distinct subcomplexes. The core subcomplex of TFIIH contains XPB, p62, p52, p44, and p34. The DNA-dependent ATPase activity of this subcomplex is required for stable promoter opening and transcription initiation (Conaway & Conaway, 1989; Roy *et al*, 1994). Absence of TFIIH leads to promoter stalling of Pol II and abortive transcription. Moreover, the ATP-dependent helicase activity of XPB is required for promoter clearance of Pol II (Kumar *et al*, 1998). The cyclin activating kinase (CAK, composed by CDK7, Cyclin H, MAT1) subcomplex of TFIIH phosphorylates Pol II CTD (Lu *et al*, 1992). It has been demonstrated that Pol II associates with the pre-initiation complex in a hypophosphorylated form (II A) whereas it has to be hyperphoshorylated (IIO) to escape from the promoter and switch to the elongating form. TFIIH has a crucial role in this transition via its CDK7 subunit that phosphorylates serine 5 of Pol II CTD (Lu *et al*, 1991; Serizawa *et al*, 1993).

## 1.3.2 Preinitiation complex assembly pathways from a deterministic perspective

The discovery of GTFs initiated extensive investigations, using mainly *in vitro* approaches, to discover the way these factors interact with Pol II and the DNA template to form a stable pre-initiation complex (PIC). Two models derived from these studies:

• The sequential assembly pathway: According to the model describing this pathway, the PIC is assembled at the promoter in a stepwise manner. Particularly, it was described that TFIID is the GTF that binds first at the promoter core element and that this interaction is further stabilised by the presence of TFIIA and TFIIB (Buratowski *et al*, 1989; Van Dyke *et al*, 1989). The next steps involve the recruitment of Pol II together with TFIIF that results in the formation of a

stable complex between TFIID-TFIIA-TFIIB-Pol II/TFIIF and the DNA template at the promoter site. Finally TFIIE are TFIIH are recruited to complete the formation of a stable transcription PIC (Figure 1.4, A).

- The Pol II holoenzyme pathway: The second model came up as a consequence of the observation that Pol II could be purified as a holoenzyme complex containing not only some GTFs but also other factors that included the SWI/SNF chromatin remodelling complex, the acetyltransferase GCN5, and the co-activator mediator (Kim *et al*, 1994; Koleske & Young, 1994). The absence of TFID from these purifications was in favour of the notion that TFIID may act independently as a factor that binds to the core promoter, its binding is stabilised by TFIIA and then allows the recruitment of the Pol II holoenzyme (Thomas & Chiang, 2006) (Figure 1.4, B).
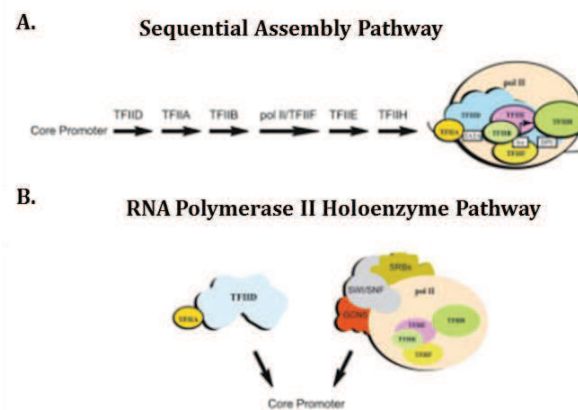


**Figure 1.4. Preinitiation complex assembly pathways.**
According to the two pathways suggested for the recruitment of GTFs in a highly ordered process, preinitiation complex (PIC) formation can occur: (A) by stepwise recruitment of the general transcription machinery (the sequential assembly pathway) or (B) by recruitment of preassembled RNA pol II holoenzyme and TFIID complexes (the two-component pathway).

There is no study providing conclusive evidence to support the validity of either model. However, although the models seem contradictory their co-occurrence in vivo cannot be excluded. Moreover, as it is discussed in chapter 1.6, the findings of studies based on live imaging techniques are challenging this static view of PIC assembly. Indeed the suggested models may represent the two extremes of a dynamic system under different conditions. In addition, an important aspect of eukaryotic genomes has further "complicated" our view about the interactions of the GTFs and promoters, as well as other transcription related regulatory elements. The "complication" being the fact that these regulatory elements are not readily accessible due to the compaction of DNA into chromatin. Thus, the function of additional co-factors is needed to cross this additional barrier.

## 1.4 Chromatin organisation

The eukaryotic genome is tightly packed in the confined space of nucleus in the form of chromatin. Thus, it is chromatin that should be considered as the actual substrate of fundamental DNA dependent nuclear processes like transcription, replication and repair. Chromatin is not just a static DNA-protein complex but rather a macromolecular assembly with a broad range of conformational dynamics *in vitro* and *in vivo* (Figure 1.5).
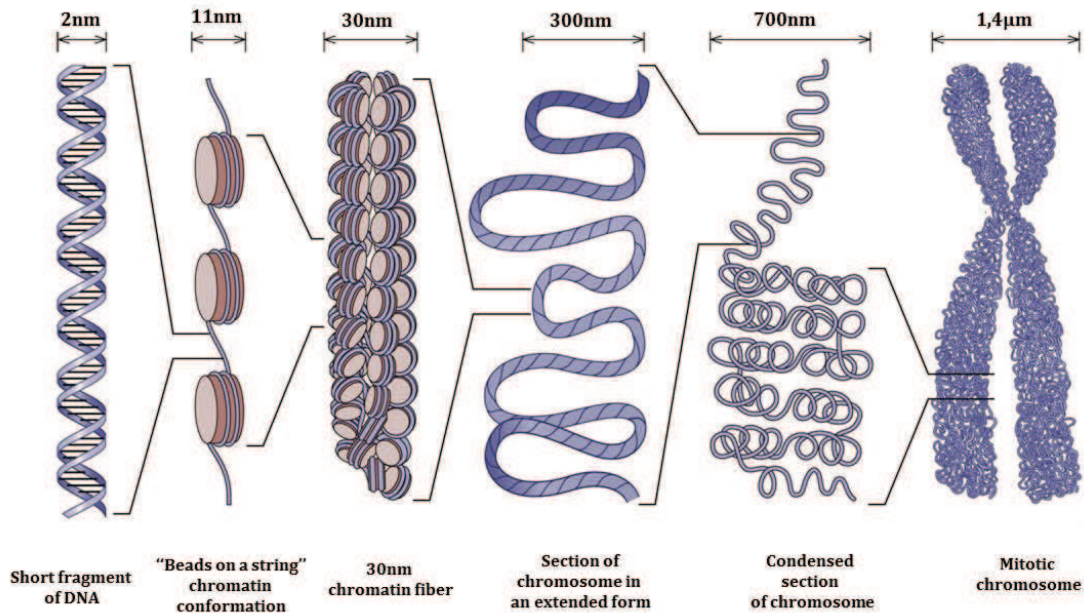


**Figure 1.5.Chromatin organisation.**
Schematic illustrating the different levels of chromatin compaction. DNA is wrapped around a histone octamer to form nucleosomes that are connected by stretches of linker DNA in a ''beads on a strings'' conformation. Nucleosome folding results to the formation of a fiber-like structure of about 30 nm in diameter. Further compaction of the 30 nm fibers into higher-order structures eventually results in the formation of chromosomes. As chromatin is a dynamic structure, those distinct levels of organisation may not be clearly distinguished in living cells. Adapted from (Jansen & Verstrepen, 2011).

The basic structural unit of the chromatin fiber is the nucleosome core, a particle that consists of 147 base pairs of DNA wrapped around a histone octamer (Kornberg, 1974; Arents & Moudrianakis, 1993; Richmond & Davey, 2003). Each histone octamer is composed of two copies of H2A, H2B, H3, and H4 (core histones) assembled in an H3/H4 tetramer and two H2A/H2B dimers (Figure 1.6). All four core histones have a histone fold domain at the carboxy-terminal end. These domains are unique, evolutionarily conserved dimerization motifs. They are particularly significant for the assembly of the nucleosome core since they organise the central 129 of 147 base pairs in left-handed superhelical turns. Consequently, these are the domains that mainly contribute to DNA-protein interactions (Baxevanis *et al*, 1995; Richmond & Davey,

2003). A direct consequence of the nucleosome structure is that it may act as an obstacle for transcription.

In addition, core histones form extensions of disordered tails at their amino-terminal ends. Histones H2A and H2B have a shorter carboxy-terminal tail domain too. Core histone tails protrude from the nucleosome and they are considered as determinants of chromatin fiber dynamics. They interact loosely with nucleosomal DNA hence they are the most accessible and mobile parts of the nucleosome.

A nucleosome is composed by the nucleosome core together with the DNA segments that connect it with the preceding and succeeding nucleosome cores. Those sequences called linker DNA, vary in length among species and cell types, with the average nucleosome length being 200 bp (Schalch *et al*, 2005). The combination of 11 nm nucleosome arrays results in a condense fiber of approximately 30 nm diameter. Note however, that several recent studies questioned the existence of the 30 nm chromatin fiber (Ausio, 2014; Razin & Gavrilov, 2014). The compaction of the nucleosome arrays into the higher order form is facilitated by a fifth histone called H1 (Thoma *et al*, 1979). H1 does not interact with the core histones but it can stabilise the interaction of 20 additional bp with the periphery of the nucleosome resulting in a chromatin particle of approximately 160 bp named chromatosome (Simpson, 1978; Song *et al*, 2014).
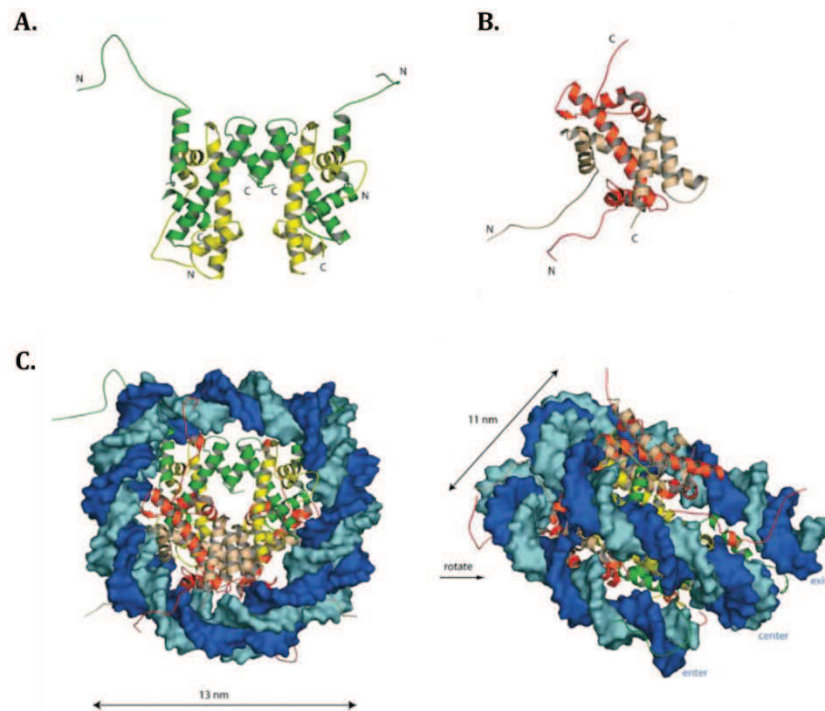
**Figure 1.6. The atomic structure of core histones and the nucleosome core particle.**
**A.**: Core histones H3 (green) and H4 (yellow) form a tetramer. **B.**: A dimer is formed from H2A (red) and H2B (pink). **C.**: The DNA makes 1.7 turns around the histone octamer to form the nucleosome core particle that has a disk-like structure. Each strand of DNA is shown in different shade of blue. Histones are coloured as in (A) and (B). Adapted from (Khorasanizadeh, 2004).

Initially, formation of chromatin was considered just as way of packaging large genomes in the confined nuclear volume. However, increasing evidence shows that the formation of this highly packed yet dynamic structure may also result in the selective restriction of transcription factor movements and thus their accessibility to their regulatory sites.

## 1.4.1   Histone PTMs and modulation of chromatin organisation and function

Chromatin, from the basic building unit (nucleosome) to the complex conformation of chromatin fibers into higher order assemblies (chromosomes), has specific structural characteristics. However, there are several mechanisms that can alter chromatin architecture at different levels. Post-translational modifications (PTMs) of histones, chromatin remodelling complexes and exchange of canonical histones with histone variants have been extensively studied. Histone PTMs have an impact both on chromatin structure but also in transcription regulation. In the following sections the general ways that histone PTMs impact chromatin structure and consequently transcription efficiency are discussed and the functional output of some of the best-characterised histone PTMs is summarised.

Before describing the role of specific histone modification, I will summarise the two general ways in which these modifications affect the accessibility of chromatin-associated factors with their target sites. The first is by directly altering the electrostatic charge of histones and/or the affinity between different nucleosomes. The second, non-mutually exclusive way that chromatin modifications can affect transcription, is by defining the type of chromatin binding proteins that are recruited at regulatory regions.

**1.4.1.2 Direct structural effects of histone modifications on nucleosomes and chromatin**

The first evidence that histone PTMs may affect the properties of nucleosomes came from studies related to the effects of histone tails on chromatin conformation. It was shown that removing the amino-terminal histone tails from nucleosome cores had a negative effect on their further compaction into a 30 nm chromatin fiber (Fletcher & Hansen, 1995). Moreover, removing hyperacetylation from histone tails from nucleosomes was increasing the accessibility of sequence specific factors to their binding sites (Vettese-Dadey *et al*, 1996). On the other hand, removal of histones tails by trypsin digestion did not seem to affect the stability of the nucleosome core (Ausio *et al*, 1989). Thus, the role of histone tails seemed to be more related to the interaction between nucleosomes and the accessibility of transcription factors to their sites than with the stability of the nucleosome core. An example of how histone tail PTMs can affect chromatin structure is the acetylation of H4K16. Experiments in which nucleosomal arrays containing acetylated H4K16 were compared to arrays lacking this PTM or lacking the full H4 tail, showed that H4K16 acetylated arrays did not form the 30 nm fiber as it was the case for non-acetylated arrays (Shogren-Knaak *et al*, 2006).

Apart from the effect of modifications on sites found at the histone tails, nucleosome stability can be affected by modifications that alter histone/DNA interactions. PTMs at the histone fold domains or the additions of larger moieties like ubiquitin on histones are thought to have this type of effects. A relevant example is acetylation of lysine 56 of histone H3 (H3K56ac). It occurs at the region where DNA enters and exits the nucleosome suggesting that acetylation could affect DNA-histone interaction. This modification has been extensively studied in yeast and among others has been shown to play a role in genomic instability, repair and transcriptional initiation and elongation, by altering nucleosome stability (Celic *et al*, 2006; Haldar & Kamakaka, 2008; Williams *et al*, 2008; Värv *et al*, 2010). In human cells, the same modification was identified in a screen for histone PTMs responsive to DNA damage (Tjeertes *et al*, 2009) and has also been found to correlate positively with the binding of Nanog, Sox2, and Oct4 transcription factors at their target gene promoters (Tan *et al*, 2013). This fact indicates the H3K56ac is related to fundamental, highly conserved chromatin regulated processes.

On the other hand ubiquitination of K123 of H2B, although it involves the addition of bulky peptide like ubiquitin (~8kDa) is believed not to disrupt but actually increase nucleosome stability during transcription initiation and elongation (Chandrasekharan *et al*, 2009). However, these effects have to be further investigated. In addition, it has been shown to serve as a bridge for the recruitment of methylases of histone H3 (Lee *et al*, 2007).

**1.4.1.3 Histone PTMs as docking sites for effector proteins**

An alternative way, in which posttranslational modifications of histones can affect the structure of chromatin, is that they consist marks for histone PTM "readers" that associate with chromatin via domains that recognise particular modifications. In many cases, these "readers" can be subunits of chromatin-remodelling or chromatin-modifying complexes. In contrast to histone modifiers, chromatin-remodelling complexes are ATP-dependent and their effect on chromatin has to do with repositioning/sliding of nucleosomes on DNA, removing histones, or exchanging core histones with histone variants (Lusser & Kadonaga, 2003; Clapier & Cairns, 2009). The activity of these complexes is required, among others, to maintain the compaction of chromatin, to attribute specialised characteristics in well-defined chromatin regions and to regulate the accessibility of DNA binding factors in the compact environment of chromatin. Particularly for transcription activation, the activity of chromatin modellers can expose DNA regulatory elements facilitating the binding of activators to upstream enhancers, or enable the association of the components of the basal transcriptional machinery with promoters (Clapier & Cairns, 2009; Voss & Hager, 2014).

The interaction of nucleosome chromatin-modifying or -remodelling complexes with DNA or modified histones is carried out via specific domains of their subunits. Well-studied examples of a motif common in this type of complexes are the bromodomains that recognise acetylation, the chromo-like domains of the Royal family and the PHD domains that recognise methylation, and the domain within 14-3-3 proteins that recognises phosphorylation (Dhalluin *et al*, 1999; Macdonald *et al*, 2005; Wysocka *et al*, 2006; Mujtaba *et al*, 2007). These types of interactions can tether a chromatin-remodelling or –modifying complex on chromatin regions enriched with histones that bare a particular modification. In this way, the abundance of a histone PTMs can target and/or regulate the activity of a remodeller or chromatin-modifying complex (Suganuma & Workman, 2011).

### 1.4.2 Posttranslational modification of histones

It is more than 50 years since the first evidence for a relationship between the posttranslational modifications of histones and gene regulation was found (Allfrey *et al*, 1964). Since then, several types of modifications have been found. Histone PTMs may involve the addition of small chemical moieties on histones (acetylation, methylation, phosphorylation etc.) or the attachment of large globular proteins like ubiquitin and SUMO (Small Ubiquitin-like Modifier). An important feature of these modifications, that amplifies the number and the complexity of potential molecular signals chromatin can receive and transmit, is the fact that they are reversible.

#### 1.4.2.1 Histone Acetylation

Histone acetylation occurs with the transfer of an acetyl group from acetyl-CoA to the specific lysines on histones and is catalysed by enzymes called histone acetyltransferases (HATs). The reverse process, removal of the acetyl group is performed by histone deacetylases (HDACs). The opposite action of these two enzyme families renders histone acetylation a highly dynamic modification (Yang & Seto, 2007).

The addition of the acetyl group to lysine is neutralising the positive charge of this residue. Thus, acetylated lysines have lower affinity for DNA. When amplified along large chromatin regions this reduction of histone/DNA affinity results in chromatin decompaction of various levels. Hence, it is not surprising that histone acetylation has been positively correlated with active transcription with that corresponding marks found at the sites of active promoters and euchromatic regions (Katan-Khaykovich & Struhl, 2002; Pokholok *et al*, 2005; Roh *et al*, 2005).

In general, acetylated histone residues are considered marks of active transcription. However, as for most histones PTMs, acetylation of specific residues has been found to affect different cellular pathways. Particularly, enrichment in H3K9ac and H3K14ac has been found in chromatin regions that promote gene expression (Pokholok *et al*, 2005; Li *et al*, 2007). H4K16ac can affect the global conformation of chromatin resulting in a less packed fiber that consequently becomes more accessible for nuclear proteins (Shogren-Knaak *et al*, 2006). Acetylation can also act as a signal for the assembly of chromatin as H4K5ac and H4K12ac have been described as marks for the deposition of newly synthetized histones (Sobel *et al*, 1995; Roth *et al*, 2001).

The enzymes that add the acetyl group to lysines are called histone acetyltransferases (HATs) Most nuclear HATs function in the context of larger complexes and exhibit different substrate

specificities (Carrozza *et al*, 2003). Although the first report of histone acetylation and speculation for its effects on transcription was already made 50 years ago (Allfrey *et al*, 1964), the first direct link between proteins that affect the levels of histone acetylation and transcriptional regulation resulted mainly from two studies 20 years later. In the first study, it was found that a HAT from *Tetrahymena* was homologous to Gcn5 (general control non-repressed 5)(Brownell *et al*, 1996), an already known coactivator of transcription in yeast (Georgakopoulos & Thireos, 1992). The second, showed that a histone deacetylase identified from *in vivo* screen was homologous to a yeast transcriptional regulator (Taunton *et al*, 1996). Soon after, it was shown that both the histone acetyltransferase and the deacetylase, were components of larger complexes, in the context of which the enzymes exhibit enhanced catalytic activity, and can be efficiently recruited to their target sites (Grant *et al*, 1997; Kadosh & Struhl, 1997; Roth *et al*, 2001).

### 1.4.2.2 Histone Phosphorylation

Histones can be phosphorylated on serine, threonine and tyrosine residues that mostly localise on their amino-terminal tails. The covalent attachment of a phosphate group, from ATP to the hydroxyl group of the target amino-acid side chain is catalysed by kinases whereas removal of the modification is carried out by phosphatases. The negative charge that is added to the target histone during phosphorylation, classifies this modification in the category of histone PTMs that also affect chromatin conformation. Kinases might be targeted to their chromatin sites either directly by the presence of DNA binding motifs in their sequence or indirectly by associating with chromatin interacting factors (Oki *et al*, 2007; Rossetto *et al*, 2012).

Phosphorylation of histones has been linked both with global changes of chromatin structure but also with specific signalling pathways. Particularly, in mammals, phosphorylation of H3 at Ser10, is mediated in mitosis by the Aurora-B kinase and it has been shown to be involved in the initiation of the chromosome condensation process. However, the presence of the mark is not necessary to maintain the compact chromatin state (Van Hooser *et al*, 1998; Crosio *et al*, 2002; Johansen & Johansen, 2006). A well-studied example of histone phosphorylation involved in signalling, is the case of the histone variant H2A.X (called γH2A.Z when phosphorylated) (Rogakou *et al*, 1998). Two serine residues of its carboxy-terminal are phosphorylated in quick response to DNA double strand break (DSB) damage. γH2A.Z is then recognised by DNA repair enzymes that are recruited on the damage site. It is also recruiting the kinases that catalyse the phosphorylation of H2A.X. In that way, a positive feed-back loop is generated that spreads γH2A.Z to large DNA regions (Kang *et al*, 2005).

### 1.4.2.3 Histone Methylation

As described above, acetylation and phosphorylation of histones is changing the charge of the modified protein. On the contrary, methylation has no such effect. It mainly occurs on the side chains of lysines and arginines of H3 and H4. Lysines can be mono-, di- or trimethylated whereas arginines can be monomethylated and symmetrically or asymmetrically dimethylated. This range of possible methylated states is indicative of the complexity underlying the effects of this mark (Ng *et al*, 2009).

The enzymes depositing methylation marks are collectively called histone methyltransferases (HMTs). Most HMTs have a highly conserved catalytic domain called SET, found in all studied eukaryotes. Many studies have focused on the study of HMTs especially on those that regulate developmental genes as components of larger protein complexes like Polycomb and Trithorax (Dillon *et al*, 2005; Steffen & Ringrose, 2014).

Methylation is a mark that can have diverse effects in chromatin structure and transcription regulation. These effects depend on the methylated residue, the level of methylation and the proteins that recognise the modification. H3K4me3, H3K36me3, or H3K79me3 are marks characterising transcriptionally active regions. Heterochromatic regions and regions that are transcriptionally repressed are found enriched in H3K9me3, H3K27me3, and H4K20me3. However an interesting finding that reveals the complexity but also the regulatory potential of this histone modification is the monomethylated H3K9, H3K27, and H4K20 which are predominantly found in euchromatic regions. Thus, it has been shown that even the degree of methylation of the residue can differentially regulate gene expression (Black *et al*, 2012).

### 1.4.2.4 Histone Ubiquitination

Modifications described above result in the addition of small chemical moieties on a histone. In contrast to those, ubiquitination entails the addition of a bulky polypeptide (~8 kDa). The addition of such a moiety is expected to have significant effects on chromatin structure and transcriptional regulation.

Three enzymes, that have been classified in three types, act sequentially to complete the covalent binding of ubiquitin on lysines of histone or non-histone proteins. First, an E1 enzyme activates ubiquitin in an ATP dependent step. Next, activated ubiquitin is transferred to the active site of a ubiquitin carrier protein enzyme, E2. The last step requires the action of a

ubiquitin-protein ligase, enzyme E3, that is highly substrate specific and is catalysing the linkage of ubiquitin to the lysine of the target protein (Hershko & Ciechanover, 1998).

Two well-studied sites of histone lysine monoubiquitination are the H2AK119 and H2BK120 (in human cells) that are ubiquitinated mainly by co-transcriptional mechanisms but seem to have counteracting effects in transcription activation (Kim *et al*, 2009; Wang *et al*, 2009). H1, H2AZ, H3 and H4 have also been found ubiquitinated but their abundance is low.

H2A is monoubiquitinated at K119 (H2Aub) exclusively in higher eukaryotes and found highly enriched mainly in heterochromatic regions like the inactive X chromosome or the transcriptionally silenced XY body (Smith *et al*, 2004; Baarends *et al*, 2005). The first direct link between H2AK119ub and gene repression was made when it was discovered that the E3 ubiquitin ligase complex of human H2AK119 is composed of four proteins Ring1a, Bmi-1, HPH2 , Ring1b; all members of the Polycomb group with Ring 2 identified as the catalytic subunit (Wang *et al*, 2004). Later studies showed Ring1A and Bmi-1are involved in Hox gene silencing and suggested that H3K27me regulates H2AK119 ubiquitination by PRC1 (Cao *et al*, 2005). According to the suggested regulatory pathway for H2Aub deposition, the binding of PRC2 with its H3K27 HMT activity results in the recruitment of PRC1 that has subunits with E3 ubiquitin ligase activity, that catalyse ubiquitination of H2A. Regarding the mechanistic aspects of the function of H2AK119ub there is evidence from studies on promoters of chemokine response genes, that this mark forms a physical barrier which blocks FACT recruitment. FACT is one of the most thoroughly characterized factors that are required for Pol II elongation on chromatin templates (Orphanides *et al*, 1998). Essentially, it allows elongation on chromatin by binding and displacing the H2A/H2B dimer from the core nucleosomes (Belotserkovskaya *et al*, 2003).Thus the physical barrier of ubiquitin results in Pol lI pausing in the promoter proximal region, stressing the role of H2A deubiquitination for transcription initiation or elongation. As H2AK119ub can be deubiquitinated by the PR-DUB complex the abundance of this mark can be tightly regulated and influence transcriptional activation (Zhou *et al*, 2008).

In contrast to H2AK119ub, monoubiquitinated H2BK120 (H2Bub) is a mark of active transcription. In mammals, RNF20 and RNF40 (homologs of yeast Bre1 ubiquitin ligase), heterodimerize and interact with RAD6 (H2B ubiquitin conjugating enzyme) to monoubiquitinate H2B (Kim *et al*, 2009). H2Bub is a dynamic mark that is removed by conserved deubiquitinases. Ubp8 and USP22 are the main H2Bub deubiquitinases in yeast and human cells respectively. In both organisms, they are subunits of the SAGA co-activator complex and act as components of its deubiquitinating (DUB) module (Henry *et al*, 2003; Zhang *et al*,

2008; Zhao *et al*, 2008). Further details regarding the contribution of SAGA complex DUB module in global H2Bub levels and transcription regulation are described in chapter 1.5.10.

A cross –talk between H2Bub and H3K4 and H3K79 methylation has been described in yeast (Briggs *et al*, 2002; Sun & Allis, 2002). Reduction of H2Bub levels supresses HMTs Dot1 and Set1 probably by different mechanisms. In the case of Dot1 it has been suggested that the presence of H2Bub acts as an allosteric activator of catalytic activity of the enzyme (McGinty *et al*, 2008). However, in the case of Set1 that acts in the context of a complex (COMPASS), a mechanism in which H2Bub acts as docking site facilitating the recruitment of the complex seems more probable (Lee *et al*, 2007; Vitaliano-Prunier *et al*, 2008). There is some evidence that this type of crosstalk may also occur, although not fully conserved, in higher eukaryotes (Kim *et al,* 2009; Wang *et al*, 2009; Jung *et al*, 2012).

Deposition of H2Bub has been proposed to be mediated by the PAF complex. PAF is known to associate with elongating Pol II and act as transcription elongation regulator. Deletion of two subunits of PAF in yeast decreased H2Bub levels without affecting the recruitment of Rad6-Bre1 (E2-E3) enzymes (Ng *et al*, 2003a; Wood *et al*, 2003). This suggests that PAF activates Rad6-Bre1 and through direct interaction could transfer it along the transcribed region. Other studies have shown that some of PAF subunits are required for H3K4 and H3K79 methylation and also that PAF is required for the recruitment of the COMPASS complex at promoters and 5' regions of genes (Krogan *et al*, 2003; Ng *et al*, 2003b). This evidence, taken together with well supported model of H2Bub-H3K4me/H3K79me cross-talk, link PAF with transcription elongation and H2Bub deposition.

ChIP-on-chip or ChIP-seq analyses of H2Bub distribution in the genome of human, plants and *D. melanogaster* showed that it is enriched exclusively in regions of active genes whereas it seems to be absent at silent genes, intergenic regions and other heterochromatic regions (Minsky *et al*, 2008; Kharchenko *et al*, 2011; Roudier *et al*, 2011). With similar experimental approach, H2Bub in yeast was found uniformly distributed along genic nucleosome positions (Batta *et al*, 2011). Taken together, these genome-wide studies would be in agreement with a model of H2Bub deposition in RNF20/RNF40/Rad6/PAF/RNA pol II elongation dependent way.

Since the link between transcription elongation and deposition of H2Bub is quite well established, several studies have focused on the interactions between the components of elongating RNA pol II machinery and the enzymes involved in H2B ubiquitination. The most direct, mechanistic link between the two processes was made when it was found that the presence of ubiquitinated H2B during elongation is increasing the efficiency of nucleosome reassembly but also facilitates the eviction of H2A/H2B dimer (Pavri *et al*, 2006; Fleming *et al*,

2008). There is strong evidence that the combinatorial activity of FACT, PAF and monoubiquitinated H2B machinery is necessary for these processes. The simultaneous presence of these factors is allowing Pol II to move along the chromatin template overcoming the nucleosomal barrier (Bonnet *et al*, 2014a).

## 1.5 The transcriptional coactivators SAGA and ATAC

Apart from the general transcription factors described above in chapter 1.3.1, there are additional classes of transcription factors. The first includes the sequence specific DNA binding factors that may act either as activators or as repressors. These factors can specifically affect the rate of transcription of a particular gene either by binding to the promoter proximal region of that gene or by interacting with distant sequences relative to TSSs, which are characterised as enhancers or silencers depending on their effect on transcriptional activation. The action of activators (or repressors) is mainly exhibited in the regulation of genes that are related to specific developmental stages or responses to environmental stimuli (Näär *et al*, 2001; Fuda *et al*, 2009).

The second class of factors are generally termed transcription co-factors. These factors, as stated by their name, interact with activators (or repressors) to mediate the signal of the later directly to GTFs and Pol II, and thus drive specific gene expression profiles. Alternatively, an activation or repressor signal, will indirectly affect the transcription rate of a gene via changes that co-activators induce on chromatin environment that is linked to the transcription status of the target gene. Co-activators can be further classified in two general categories depending on their mode of function: The first one includes those proteins that are part of or associate with the basal transcription machinery. The most well characterised examples of co-activators that directly associate with the basal transcription machinery are the TBP-associated factors (TAFs), components of TFIID and the Mediator complex. The second category includes those factors that modify the structure of chromatin (Thomas & Chiang, 2006; Fuda *et al*, 2009). As mentioned in chapter 1.4, chromatin can be modified either by ATP dependent chromatin remodelling complexes that replace or alter the positions of nucleosomes or by enzymes that modify histones post-translationally. SAGA and ATAC, the histone acetyltransferase containing complexes investigated in the present study belong in the latter category of co-activators.

### 1.5.1  The SAGA complex in *S. cerevisiae*

SAGA (Spt-Ada-Gcn5-Acetyltransferase) is a 1,8 MDa complex which was first identified in yeast (Grant *et al*, 1997). Data from genetic interactions gradually revealed that certain genes were related to similar phenotypes resulting from defects in transcriptional activation. Particularly it was observed that *gcn5* and *ada* mutants showed growth defects together with lower level of activation by different activation domains like GCN4 and VP16 (Marcus *et al*, 1996). In parallel, direct interactions between GCN5 and ADA2 were demonstrated with genetic and biochemical approaches and ADA3 was described as the additional component of the heterotrimeric complex GCN5-ADA2-ADA3 (Horiuchi *et al*, 1995). A few years later, the discovery that GCN5 has an *in vitro* acetyltransferase activity specific for histone H3 and H4 residues, further supported the hypothesis that the identified complex could be involved in chromatin structure-dependent transcriptional activation (Brownell *et al*, 1996; Pollard & Peterson, 1997). Apart from yADA2 and yADA3, other adaptor proteins like yADA1 and yADA5 were also identified as components of the complex (Marcus *et al*, 1996; Horiuchi *et al*, 1997). As ADA5 was found to correspond to SPT20, the assumption that a functional link may exist between Ada and Spt proteins was made (Roberts & Winston, 1996). Interestingly, Gcn5 and Ada2 were found to co-fractionate with Spt3, Spt7 and Spt20 in a complex that was named SAGA (Spt-Ada-Gcn5-acetyltransferase). In addition, it was observed that Spt7 and Spt20 are important for the integrity of the complex (Grant *et al*, 1997).

Mass spectrometry analyses led to discovery that several TBP-associated factors (TAFs) (TAF5, TAF6, TAF9, TAF10, TAF12) are components of the SAGA complex, some of which also present in TFIID. These subunits have both structural and regulatory roles. Their association with a complex different than TFIID was a strong indication for the diverse roles of these factors in transcription regulation (Grant *et al*, 1998). The largest of the SAGA subunits, Tra1 (400 kDa) was identified as a stably associated component of the purified complex. Tra1, and its highly conserved human homolog TRRAP, are thought to mediate interactions between the SAGA complex and activator proteins (Grant *et al*, 1998; Brown *et al*, 2001; Bhaumik *et al*, 2004).

The discovery of Ubp8 as a putative ubiquitin protease and the validation of its presence in ySAGA initiated the characterisation of the deubiquitination (DUB) module of SAGA. Ubp8 in the context of SAGA removes ubiquitin from monoubiquitinated H2B (H2Bub) (Gavin *et al*, 2002; Sanders *et al*, 2002; Henry *et al*, 2003; Daniel *et al*, 2004). Additionally, novel components of SAGA (Sgf11, Sus1) were discovered and their presence was shown to be required for the SAGA-dependent Ubp8 activity (Ingvarsdottir *et al*, 2005; Lee *et al*, 2005; Köhler *et al*, 2006). The

fourth subunit of the DUB module is Sgf73 that via its N-terminal domain links this module with the rest of the SAGA complex (Köhler *et al*, 2008; Lee *et al*, 2009).

### 1.5.2 The ADA complex in *S. cerevisiae*

Apart from the large multisubunit SAGA complex a smaller GCN5 containing complex has been identified in yeast. Initially, it was found that this small complex shares Ada2, Ada3, Gcn5 subunits with SAGA. Early studies had identified an ADA complex specific subunit, Ahc1 (ADA HAT complex Component 1) which was found to play a role in the stability of the complex (Eberharter *et al*, 1999). More recently, Sgf29 (also present in SAGA) was also characterised as a bona fide subunit of the complex together with the novel component Ahc2 that was found in close proximity with Ahc1 and co-purified with the other components of the complex (Lee *et al*, 2011).

It was shown that both purified SAGA and ADA HAT complexes acetylate H3K14 and H3K18 peptides but only SAGA could acetylate H3K9 (Grant *et al*, 1999). Moreover, another study has shown that deletion of GCN5 bromodomain has an effect on the nucleosome acetylating activity of SAGA complex but no detectable effect was observed on the activity of Ada complex (Sterner *et al*, 1999). Thus, although the information on the function of ADA complex is not extensive, there is evidence that the presence of GCN5 within this complex alters its enzymatic properties in a way that is distinct from SAGA.

### 1.5.3 The SAGA-LIKE (SLIK) complex of in *S. cerevisiae*

As the name of this complex suggests, its subunit composition is related to that of the SAGA complex. The SLIK complex contains Gcn5, has similar substrate specificity with SAGA (towards histone H3) and is involved to some extent in gene expression regulation. The two complexes share at least 13 subunits but there are some differences in their composition that are considered particularly important for their specific roles (Pray-Grant *et al*, 2002).

First, unlike SAGA, SLIK does not contain Spt8 that is important for the SAGA-TBP interactions. Notably SAGA purified from an *spt8Δ* does not interact with TBP *in vitro* (Sterner *et al*, 1999). A second characteristic of the SLIK complex is that it contains a truncated version of Spt7, resulting from proteolytic cleavage (Spedale *et al*, 2010). This version of Spt7 lacks the C-terminal Spt8 binding domain that explains the absence of Spt8 in SLIK. This feature of SLIK makes it more similar to the metazoan SAGA complex. Higher eukaryotes have a truncated homolog of yeast Spt7 and no gene ortholog of SPT8 is found in their genome (Spedale *et al*, 2012). Third, Rtg2 is a SLIK specific component essential for the structural integrity of the

complex. This protein is involved in the regulation of genes whose expression is altered in cells with abnormal mitochondrial function (Pray-Grant *et al*, 2002). Genetic data show that SLIK and SAGA have partially overlapping functions. It has been demonstrated that the phenotype observed in *stp7Δ* mutant (encoding Spt7 that is shared between the complexes) is similar but more severe than those of individual *rtg2Δ* or *spt8Δ* mutants which show milder and partially overlapping phenotypes. Nevertheless, phenotypes related exclusively to *rtg2Δ* or *spt8Δ* mutants were also identified suggesting that the two complexes are not functionally redundant (Pray-Grant *et al*, 2002).

### 1.5.4 The SAGA complex in metazoans

Information for SAGA complex composition and function in metazoans comes mainly from studies conducted in *Drosophila* and human cells. As shown in Table 1.1, the composition of the complex between yeast, fly and human is similar. However, there are distinct features which reflect the functional divergence of the complex during evolution.
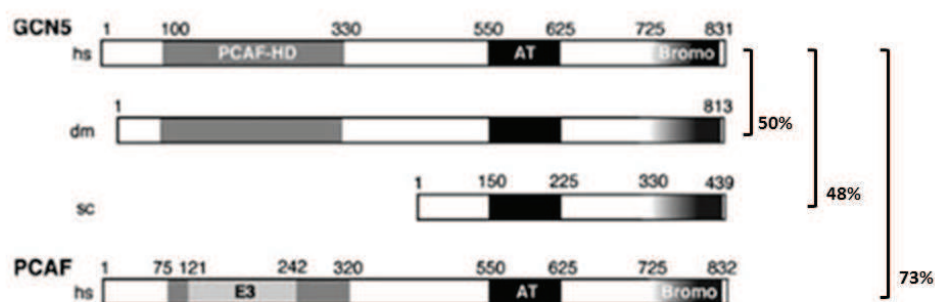


**Figure 1.7. Domain organization of the GCN5 and PCAF enzymes in human, *Drosophila* and yeast.**

Schematic representation of domains organization of the GCN5 and PCAF proteins from human (hs; *Homo sapiens*), *Drosophila melanogaster* (dm) and yeast (sc; *Saccharomyces cerevisiae*) (only GCN5 is present). The PCAF homology domain (PCAF-HD) is shown in grey, the acetyltransferase (AT) domain is shown in black and the bromodomain (Bromo) is shaded. E3 is a ligase domain that has been described for PCAF (Linares *et al*, 2007). Amino-acid positions are marked over the boxes. The % of identity between the different factors is indicated on the right of the horizontal lines, representing the pair wise comparisons. Modified from (Nagy & Tora, 2007).

Two important differences have to do with the catalytic subunit of the complex in this group of organisms. First, metazoan GCN5 bares an amino-terminal extension, called the PCAF homology domain, that almost doubles its size in comparison to yGcn5. Second, specifically in vertebrates, *GCN5* has duplicated and diverged to the *PCAF* gene which encodes a protein that is ~70%

identical to GCN5 (Figure 1.7) (Yang *et al*, 1996). PCAF can also incorporate into SAGA, in a mutually exclusive way to GCN5 and has a HAT function similar to GCN5 (Krebs *et al*, 2010; Nagy *et al*, 2010). Thus, in contrast to *Drosophila*, in human two SAGA complexes can be formed, one containing GCN5 and the other containing PCAF.

Several studies have investigated the composition of hSAGA complex. Initially, minor discrepancies were observed in the composition of the complex between different labs (Martinez *et al*, 1998; Ogryzko *et al*, 1998; Wieczorek *et al*, 1998). Most of these have been eventually attributed to the different methodological approaches used in each case (e.g. purification by immunoprecipitation against endogenous subunit vs purification against overexpressed tagged subunit). Currently, apart from Spt8, homologs of every other ySAGA subunit have been found in the respective metazoan SAGA complexes ( Table 1.1).

**Table 1.1. Subunits and functional modules of SAGA complex in yeast, fly and human**

| Module | *S. cerevisiae* | *D. melanogaster* | *H. sapiens* |
|---|---|---|---|
| **HAT** | Gcn5 | Gcn5 | GCN5/PCAF |
|  | Ada2 | Ada2b | ADA2b |
|  | Ada3 | Ada3 | ADA3 |
|  | Sgf29 | Sgf29 | SGF29 |
| **DUB** | Ubp8 | Nonstop | USP22 |
|  | Sgf73 | CG9866 | ATXN7/-L1/-L2 |
|  | Sgf11 | Sgf11 | ATXN7L3 |
|  | Sus1 | E(y)2 | ENY2 |
| **Link with activators** | Tra1 | Tra1 | TRRAP |
| **TBP regulation** | Spt3 | Spt3 | SPT3 |
|  | Spt8 | - | - |
| **Structural core** | TAF5 | WDA | TAF5L |
|  | TAF6 | SAF6 | TAF6L |
|  | TAF9 | TAF9 | TAF9/TAF9b |
|  | TAF10 | TAF10b | TAF10 |
|  | TAF12 | TAF12 | TAF12 |
|  | Ada1 | Ada1 | ADA1 |
|  | Spt7 | Spt7 | SPT7 |
|  | Spt20 | Spt20 | SPT20 |

Subunit composition of SAGA complex in *S. cerevisiae*, *D. melanogaster*, and *H. sapiens*. Subunits are grouped according to the function of the complex in which they are mainly involved.

### 1.5.5  The modular organisation of SAGA reflects its multiple functions

SAGA complexes from yeast to human, have multiple but conserved functions that can be attributed to the different activities of the complex. These include the histone acetyltransferase activity, H2B deubiquitinase activity and interactions with activators and TBP. Those functions seem to be highly related to the organisation and interaction of the subunits within the complex. Thus, a summary of the available information about the structure of SAGA and the organisation of its components can provide useful insights on how the different activities are regulated. To date, information about the structure of SAGA complex comes mainly from studies, which have been conducted from SAGA isolated from *S. cerevisiae* and human cells (see below). Since yeast and human SAGA complexes are highly conserved, it would be reasonable to assume that the complex must have very similar organisation and function in the two organisms.

The first Electron Microscopy (EM) study having as a central goal to gain insights in the structure of SAGA involved purification of the complex from human cells (Brand *et al*, 1999a). Later, ySAGA was purified by TAP-tagged Spt7 pull-downs in yeast strains (Wu & Winston, 2002), and the relative position of several SAGA components was shown in a 3D structure produced by EM after immuno-labelling (Wu *et al*, 2004) (Figure 1.8, A). These two studies provided direct information about the actual structure of the complex at 3.5-nanometer resolution by electron microscopy and digital image analysis of single particles. On the other hand, the limited number of targeted subunits and the effect of the position of antibody epitopes in the final yeast SAGA structure did not allow precise estimation of interactions between multiple subunits. More recently, an approach based on quantitative proteomics, took advantage of several deletions of non-essential SAGA components to create a model for SAGA subunits interaction and described its modular organisation (Lee *et al*, 2011) (Figure 1.8, B). The latest model for yeast SAGA architecture, comes from a study based on crosslinking of the purified complex followed by mass spectrometry analysis, that provides information on SAGA subunits interactions at the amino acid level (Han *et al*, 2014).

The common ground of the conclusions made from these structural-functional studies of SAGA architecture investigation is that the complex is organised in modules consisted of subunits related to a specific activity. An important observation is that the subunits with the two enzymatic activities of the complex are localised in two distinct modules the histone acetyltransferase and the deubiquitinase module (HAT and DUB module respectively). The HAT module includes Gcn5 together with Ada2, Ada3 and Sgf29, regulators of its catalytic activity and specificity. The DUB module consists of the ubiquitin protease Ubp8 (hUSP22) together with Sgf11 (hATXN7L3), Sgf73 (hATXN7/-L1/-L2) and Sus1 (hENY2) proteins that regulate the

catalytic activity of the enzyme. The core structural module consists of Ada1, Spt7, Spt20 and several TAF proteins (Table 1.1). In yeast, TAF5, TAF6, TAF9, TAF10 and TAF12 are shared between SAGA and TFIID, in contrast in metazoans only TAF9, TAF10 and TAF12 are shared between TFIID and SAGA complexes. Yeast TAF6 and TAF9, Spt7 and TAF10, and Ada1 and TAF12 have histone fold domains, and have been shown to play key structural role in the complex (Gangloff *et al*, 2001; Selleck *et al*, 2001). The importance of Ada1, Spt7, Spt20 and TAFs for the integrity of the complex has been demonstrated by deletions or mutations of these proteins that affect SAGA complex composition and integrity (Horiuchi *et al*, 1997; Grant *et al*, 1998; Sterner *et al*, 1999; Kirschner *et al*, 2002; Wu & Winston, 2002). Apart from their structural role, TAF subunits also play important transcription related regulatory roles in the context of the complex (Hall & Struhl, 2002; Klein *et al*, 2003). Another group of subunits link SAGA either with components of the general transcription machinery or with specific activators that could recruit it to specific genomic loci. Particularly, Tra1 (transcription activator-binding protein) has been shown to directly interact with transcription activators, a function that is probably related to the recruitment of SAGA at specific loci (Brown *et al*, 2001; Klein *et al*, 2003). In addition, Spt3 and Spt8, (note that Spt8 is absent in fly and human) are known to genetically and biochemically interact with TBP linking SAGA function with the basal transcription machinery (Eisenmann *et al*, 1992; Basehoar *et al*, 2004; Sermwittayawong & Tan, 2006; Mohibullah & Hahn, 2008).
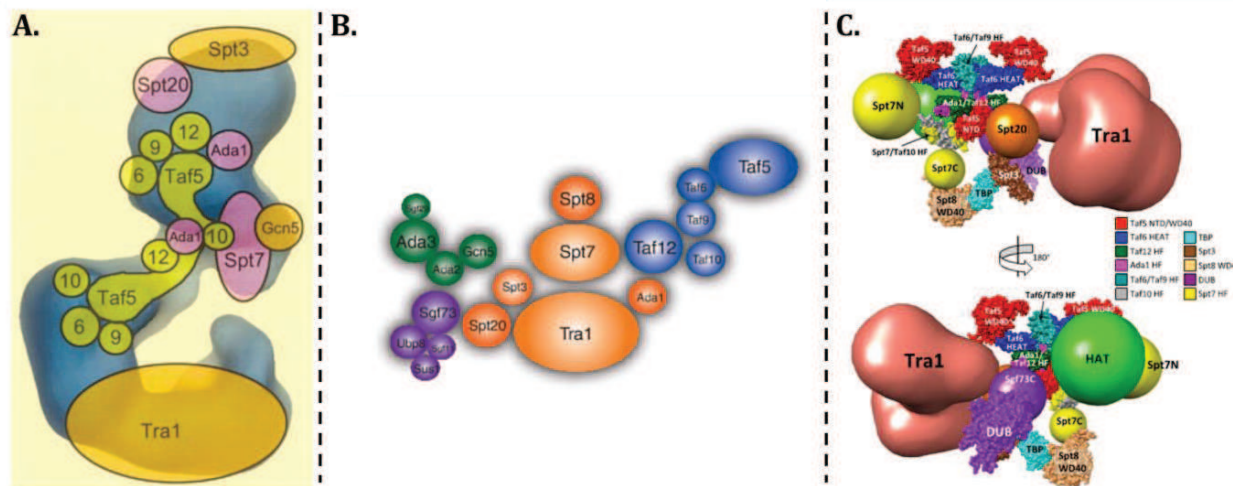


**Figure 1.8. Proposed models of yeast SAGA structure and subunits interaction.**

**A.:** Model proposed in the study of Wu *et al*, (2004) based on the 3D structure obtained by electron microscopy. **B.:** Suggested organisation of ySAGA subunits based on quantitative proteomics analysis in the context of combinatorial subunit depletions (Lee *et al*, 2011). **C.:** Structural model supported from the study of Han *et al*, (2014), based on chemical crosslinking and mass spectrometry analysis.

As described above, there is a general picture for the structure of SAGA and the organisation of its subunits. The latest model suggested by Han *et al* (2014), although it may not correspond to the precise structure of the complex, is interesting as it depicts the complexity of the interaction network amongst the subunits and the modules of the yeast complex (Figure 1.8, C). According to that model, the core module of SAGA consists of two copies of each of Taf6-Taf9, and Ada1-Taf12 (Taf-like histone-fold containing pairs) together with two copies of Taf5. This structure is supported by the intermolecular crosslinking map and is proposed as an analogous to what has been found for the core TFIID structure comprised of two copies of Taf6-Taf9 and Taf4-Taf12 together with two copies of Taf5 (Bieniossek *et al*, 2013). The addition of one copy of Taf8-Taf10 in TFIID or Spt7-Taf10 to this core structure is suggested to create the platform of interaction with the other modules of the complex (Han *et al*, 2014).

The histone fold domains of Ada1 and Taf12 but also the N-terminus of Taf5 crosslinked extensively with Sgf73 indicating that it is this subunit which links the DUB module (Ubp8 , Sgf11, Sgf73, Sus1) with the TFIID-like core module. Sgf73 was also found linked with Ada2 and Ada3, components of the HAT module. This confirms the observation of previous studies showing that Sgf73 plays crucial role in linking the DUB module to the rest of SAGA complex (Köhler *et al*, 2008; Lee *et al*, 2009). In addition, the histone fold domains of Ada1 and Taf12 were found crosslinked with Ada3, component of the HAT module (Gcn5, Ada3, Sgf29, and Ada2). Thus, from this model it is suggested that Ada3 links this module with the core of the complex. The crosslinks of Sgf73 and Ada3 suggest that HAT and DUB module are in close proximity. *In vitro* functional tests performed in the same study, support this observation showing that mutations of Sgf73 may affect the activity of the HAT module and the efficiency of its association in the SAGA complex (Han *et al*, 2014). In contrast, the study by Bonnet *et al*. (2014b) suggested that the two activities may function independently from each other. *In vitro* and *in vivo* studies have observed the inhibitory effect of poly-glutamine expansion in Ataxin7 (human homologue of Sgf73) on the catalytic activity of GCN5. Moreover, it has been shown that hGCN5 can stabilise the DUB module in human (McMahon *et al*, 2005; Palhan *et al*, 2005; Atanassov *et al*, 2009; Burke *et al*, 2013). On the other hand, it has been described that the poly-glutamine expansion in Ataxin7 has no effect on the HAT activity of SAGA (Helmlinger *et al*, 2006). Thus these studies, although controversial to a certain extent, may suggest potential links between structural conclusions derived from the crosslinking mapping and their functional consequences.

Spt20, Tra1 Spt3 also showed extensive crosslinks with the histone fold domains of Ada1 and Taf9, Taf10 and Taf12. This further supports the role of the core module as the platform for the anchoring of the other modules. Interestingly, both in the 3D structure of SAGA as described by

Wu *et al*, (2004), and the crosslinking mapping of SAGA subunits carried out by Han *et al*, (2014), TAFs are mapped at the central part of the complex. Moreover, in both studies Spt3 is in close proximity with Spt20 and Spt7 appears close to Gcn5/HAT module. The two studies also agree on the mapping of Tra1, the largest subunit of the SAGA complex that in both cases seems to be at the periphery of the complex. On the other hand, these two studies are not in agreement with the model suggested by Lee *et al*.(2011) (see Figure 1.8). According to the results of the later study, the systematic analysis of interactions in combinatorial depletion background showed that the Spt subunits do not cluster with TAF proteins and that the latter group of proteins is not localised in the centre of protein-protein interactions within the complex.

Another outcome of the study of Han *et al* (2014) regards the investigation of SAGA-TBP interaction. It was shown that TBP interacts mainly with Spt3 and Spt8 but it was also found that TBP crosslinks with Spt20 and Spt7. Although the genetic interaction of Spt3 and Spt8 with TBP is known, only Spt8 has been found to be able to interact with TBP in a SAGA independent manner (Eisenmann *et al*, 1992; Warfield *et al*, 2004; Sermwittayawong & Tan, 2006; Laprade *et al*, 2007). The explanation provided, is that Spt3 requires Spt20 to facilitate its interactions with TBP.

The information gathered so far from different studies investigating SAGA structure is precious for the interpretation of the function of the complex. It can be used as a guideline to design more targeted functional experiments that could reveal unknown aspects of the regulatory role of SAGA in transcription. Nevertheless, in most of these studies the complexes that are used for the analysis are purified and or analysed in conditions (e.g. crosslinking, recombinant complexes) that do not fully recapitulate the cellular/nuclear environment. Thus, it is probable that there are properties of the complexes that this type of analyses cannot access. For example the existence of sub-complexes could be masked by the abundance of holocomplexes under the selected conditions or the dynamic interactions between components of the complex may not be captured. Eventually, other complementary experimental approaches should be applied to obtain information for the modular organisation a complex in living cells.

### 1.5.6 Regulation of SAGA recruitment

Most studies investigating the role of SAGA in transcription activation have focused on the how the complex is recruited to gene promoters. Experimental data from different eukaryotes has shown that the presence of SAGA at the promoters of certain genes is needed to facilitate Pol II recruitment and PIC formation (Wyce *et al*, 2007; Nagy *et al*, 2009; Helmlinger *et al*, 2011; Lang *et al*, 2011). It seems that there are different mechanisms regulating this process with several subunits of SAGA being implicated.

As expected, activator mediated recruitment is one of the proposed mechanisms. One of the first subunits identified to function in this context was Tra1 (hTRRAP). As shown by several studies in *S. cerevisiae*, interacts with activators to regulate SAGA recruitment at the promoters of specific sets of genes (Brown *et al*, 2001; Bhaumik *et al*, 2004; Fishburn *et al*, 2005; Reeves & Hahn, 2005), (Figure 1.9, A). Consistent with these observations, TRRAP, the human ortholog of Tra1 had been found to regulate the activity of c-Myc and E2F (two oncoproteins involved in cell cycle regulation) before even being identified as a SAGA component (McMahon *et al*, 1998). Activator dependent interactions may involve other SAGA subunits too, particularly the TAF proteins. An example is yTaf12 that was found to interact with different acidic activators (Fishburn *et al*, 2005; Reeves & Hahn, 2005). Interactions with the basal transcription machinery components can be another way of SAGA recruitment as it is also supported by experimental data. Particularly, earlier studies have shown the genetic interaction between Spt3 and TBP (Larschan & Winston, 2001). More recently, it has been shown that Gal4 can recruit SAGA which subsequently directs the recruitment of TBP. Spt3 was identified as the subunit regulating this interaction (Laprade *et al*, 2007). Moreover the direct interaction of TBP with Spt3 and Spt8 has been demonstrated by in vitro and in vivo approaches (Mohibullah & Hahn, 2008) (Figure 1.9, B).
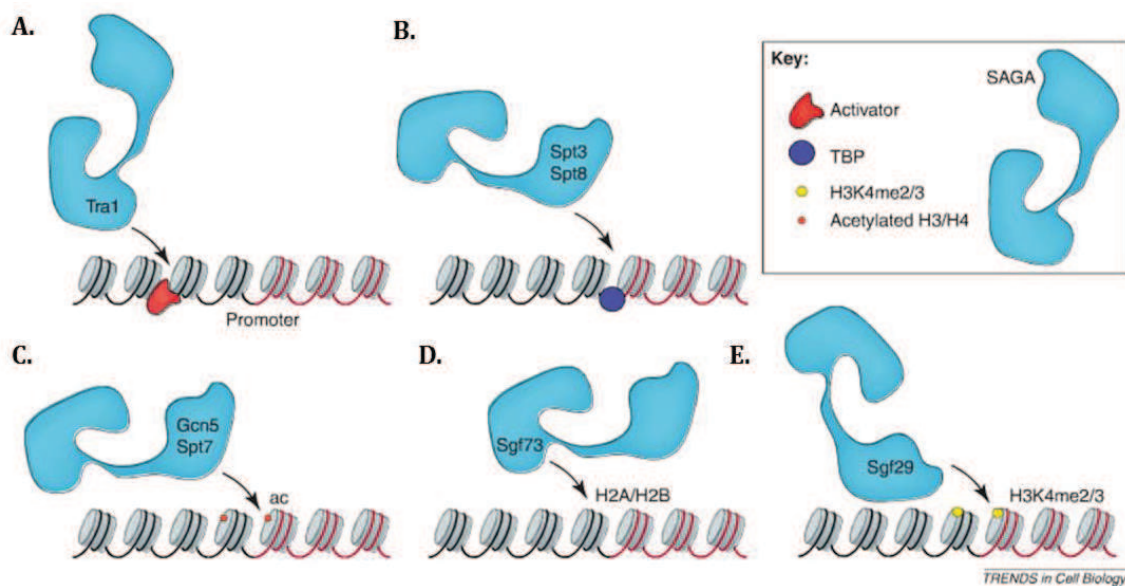


**Figure 1.9. Different modes of SAGA recruitment require distinct subunits.**
**A.:** Interactions with activators are mediated by Tra1 (hTRRAP) the largest subunit of the complex. **B.** Spt3 and Spt8 (the latter absent in human) are required for interactions with TBP and potentially the basal transcription machinery. **C.** Bromodomains found on Gcn5 and Spt7 function as "readers" of acetylated nucleosomes. **D.** The zinc finger domain of Sgf73 is responsible for the binding to H2A–H2B dimers. **E.** The tandem C-terminal tudor domains of Sgf29 mediate binding to H3K4me2/3 residues. Adapted from (Weake & Workman, 2012).

Several SAGA subunits have domains that have been shown to associate with specific chromatin components. For example, Gcn5 has a bromodomain known to associate with acetylated lysines of histones. It has been shown that in the absence of activators, the stability of SAGA on promoters depends on this domain of Gcn5. The same domain is present on ySpt7, and although it does not seem to be required for ySAGA recruitment at the promoter, it can exhibit such function if swapped into Gcn5 (Hassan *et al*, 2002)(Figure 1.9, C). Another SAGA subunit having domains with chromatin 'reading' properties is Sgf29. It has tandem Tudor domains at the carboxy-terminal part of its sequence that have been found to specifically bind H3K4me2/3 marks. The presence of Sgf29 at promoters and its overlap with H3K4me3 mark has been verified by ChIP-sequencing (Vermeulen *et al*, 2010). Moreover, SAGA H3K4m3 binding has been shown to be Sgf29 specific suggesting that it is this subunit which links the complex with the particular modification (Bian *et al*, 2011)(Figure 1.9, E). Another, interaction of SAGA subunits with chromatin components involves the SCA7 domain of Sgf73 and ATXN7 (human ortholog). This domain has been shown to bind H2A-H2B dimers but not H3-H4 tetramers (Figure 1.9, D). Thus, the SCA7 domain of Sgf73/ATXN7 is thought to mediate the interaction of the DUB module with chromatin, affecting the positioning of the module relative to its substrate A similar domain is also present in ATX7L3, another subunit of the SAGA DUB module, but in the context of the latter protein it does not have the same binding affinity for H2A-H2B (Bonnet *et al*, 2010).

It is clear that the different ways of SAGA recruitment described so far should not be considered as mutually exclusive mechanisms. On the contrary, as different subunits of the complex seem to be involved in each of these types of recruitment regulation, it is highly probable that it is the combinatorial effect of these interactions that determines the specificity of SAGA recruitment at promoters (Figure 1.9).

### 1.5.7 ATAC, the second GCN5 containing complex present in metazoans

The SAGA complex seems to be conserved in terms of subunit composition and structure from yeast to humans. However, it has been more than a decade since the identification of yADA2 paralogs in *Drosophila* (ADA2a and ADA2b) indicated that an additional Gcn5 containing complex might be present in metazoans (Kusch *et al*, 2003; Muratoglu *et al*, 2003).

This complex, called ATAC (Ada-Two-A-Containing) was initially characterised in *D. melanogaster*. The first identified interacting partners within this complex were ADA2a, Gcn5 and ADA3 (Muratoglu *et al*, 2003). Later on, the analysis of purified ATAC complex by MudPIT (MultiDimensional Protein Identification Technology) led to the identification of novel subunits (HCF and Atac1) of the complex that were not components of the SAGA complex (Guelman *et al*, 2006). Thus, it became clear that dADA2a and dADA2b participate in a distinct HAT containing complex. A proteomic approach was followed in a more systematic way to eventually identify a number of additional ATAC specific components. Among the identified subunits, dATAC2 was shown to have an acetyltransferase domain. Therefore, the ATAC complex may exhibit dual HAT activity deriving from two independent subunits (Suganuma *et al*, 2008).

Several studies examined the existence and composition of ATAC complex in mouse and human (Wang *et al*, 2008a; Guelman *et al*, 2009; Nagy *et al*, 2010). Apart from minor discrepancies, it is considered that these studies have revealed the components of ATAC complex in mammals. Overall it has been found highly conserved in fly, mouse and human.

ATAC has a HAT core module which, like in SAGA, is responsible for the HAT activity of the complex by GCN5. In human, GCN5 has duplicated and diverted to PCAF. As mentioned in chapter 1.5.4, the respective proteins are very similar and can both incorporate, in a mutually exclusive way (Krebs *et al*, 2010; Nagy *et al*, 2010), into SAGA and ATAC histone acetyltransferase modules (see Table 1.2). As these two similar proteins associate strictly with the catalytic modules of the related complexes, it is thought that they play an important role in the regulation of the HAT specificity and activity of each complex. The existence of PCAF as an alternative catalytic component of the HAT module in vertebrates, indicates that complexes containing PCAF (SAGA or ATAC) may have some specific functions. Thus, the number of SAGA-like complexes seems to increase in parallel with the developmental complexity of the organisms in which these complexes are found. Thus it is tempting to assume that some of their functions are directly related with the regulation of specific developmental pathways (Spedale *et al*, 2012).

**Table 1.2. Subunit composition of ATAC and SAGA complexes in *H. sapiens***

| hATAC | | hSAGA |
|---|---|---|
| **GCN5/PCAF** | | **GCN5/PCAF** |
| **ADA3** | | **ADA3** |
| **SGF29** | **HAT module** | **SGF29** |
| - | | **<span style="color:red">ADA2b</span>** |
| **<span style="color:red">ADA2a</span>** | | - |
| - | | TRRAP |
| - | | ADA1 |
| - | | SPT3 |
| - | | SPT7L |
| - | | SPT20 |
| - | | TAF5L |
| - | | TAF6L |
| - | | TAF9 |
| - | | TAF10 |
| - | | TAF12 |
| - | | ATXN7L3 |
| - | | USP22 |
| - | | ENY2 |
| - | | ATXN7 |
| ZZZ3 | | - |
| YEATS2 | | - |
| MBIP | | - |
| ATAC2 | | - |
| NC2β | | - |
| WDR5 | | - |

hATAC and hSAGA have three common subunits in their HAT module: GCN5/PCAF, ADA3 and SGF29 (in bold black). ADA2a and ADA2b (in bold red) are the complex specific subunits of hATAC- or hSAGA-HAT module respectively.

ATAC and SAGA HAT modules have three common subunits SGF29, ADA3 and GCN5/PCAF (Table 1.2). Both in fly and human, ADA2a is the ATAC-specific HAT module subunit and ADA2b is found only the respective module of SAGA. The identification of two HAT complexes with a common catalytic subunit raised questions on the catalytic specificities of each complex *in vitro* or *in vivo*.

When the specificity of purified yeast or human GCN5 was tested *in vitro* on free histones it was found that it acetylates mainly H3K14 and to a lesser extent H4K8 and H4K16 (Brand *et al*, 1999b; Grant *et al*, 1999). However, GCN5 acetylates mainly H3K9 and K14 residues when instead of histones the substrate is changed to mononucleosomes. Within the context of SAGA complex in *Drosophila*, GCN5 acetylates mainly H3K9, K14, K18 and K23. Purified dATAC complex, exhibits similar specificity for H3 and H4 acetylation when free histones are the substrate. However a clear difference between the specificity of the two complexes is revealed when the substrate is changed to mononucleosomes as dATAC acetylates more efficiently H4 (Guelman *et al*, 2006; Suganuma *et al*, 2008). *In vivo* results are generally in agreement with the above observations. It has been shown that the levels of H3K9ac and H3K14ac are specifically affected by mutations of dADA2b but not dADA2a (Pankotai *et al*, 2005) In support of these results, it was shown that dADA2a mutants have reduced levels of H4K5 and 12 (Ciurciu *et al*, 2006). Regarding the specificity of the second HAT subunit, *Drosophila* ATAC2, it was shown that is related to the level of H4K16ac without affecting the acetylation of H3K9, H3K14, H4K5, and H4K8 (Suganuma *et al*, 2008). Note however, that the human ATAC2 subunit did not exhibit any HAT activity *in vitro* (Nagy *et al*, 2010).

In mouse and human the specificity of ATAC is not completely clarified, but there seems to be a consensus for its function that differs from what is observed in fly. When the stability of ATAC is impaired by knocking down mATAC2, global levels of H3K9, H4K5, H4K12, and H4K16 are affected (Guelman *et al*, 2009). On the other hand ADA2a knockdown in human cells was shown to specifically affect H3K9 and H3K14 (Nagy *et al*, 2010). Although there is some level of discrepancy between these two studies, the important, common observation is that ATAC function in mammals has expanded its specificity to at least one H3 lysine residue (K9).

Altogether these studies show that GCN5 can acetylate different types of substrates and/or histone lysine residues depending on the complex in which it is incorporated. This change in specificity is believed to be the effect of the interactions between the catalytic subunit of the complexes and the other subunits with which it associates

### 1.5.8 Structural organisation of ADA2a and ADA2b and their role in GCN5 activity regulation

Since ADA2a and ADA2b constitute the only different subunit between the HAT modules of ATAC and SAGA, it has been hypothesised that their interaction with the other components of these modules (GCN5/PCAF, ADA3, SGF29) define aspects of the catalytic activity and specificity of each module. Both ADA2a and ADA2b have a ZnF (zinc finger) and a SANT domain at their amino-terminal end, and a SWIRM domain at their carboxy-terminal end (Figure 1.10). Indeed, there is significant evidence that these domains play a role both in the function of the HAT modules, but also in the interactions between subunits.
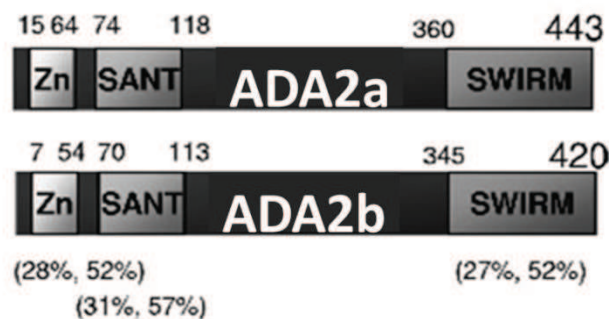


**Figure 1.10. Domain organisation and sequence relationships of human ADA2a and ADA2b.**
Scheme illustrating levels of sequence identity and homology (identity %, homology %) between the three domains of ADA2a and ADA2b. Numbers above the scheme correspond to amino acid positions. Overall sequence identity and homology is 26% and 50% respectively. Zn: zinc finger. Adapted from (Gamper *et al*, 2009).

In yeast, the SANT domain of ADA2 is required for the activity Gcn5 on free histones (Sterner *et al*, 2002). On the other hand, human GCN5 can acetylate free histones without any enhancement of this activity by the presence of ADA2a or ADA2b. However, the presence of ADA2b but not ADA2a seems to be required for the activity of GCN5 on mononucleosomes (Gamper *et al*, 2009). It is suggested that the SANT domain of ADA2b is involved in the recognition and/or interaction with nucleosomes, a function that cannot be performed by the respective domain of ADA2a. This difference, at present, is not well understood. Moreover, it has been shown that the interaction of ADA2b with GCN5 requires mainly the amino-terminal region of ADA2b that includes most of its SANT domain (Gamper *et al*, 2009)(Figure 1.11).
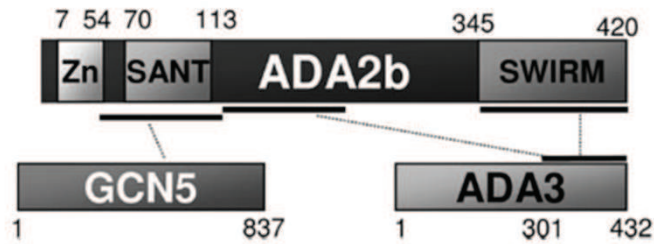
**Figure 1.11. Schematic summary of ADA2b domain interactions with GCN5 and ADA3.**
Ada2b SANT domain is important from the interaction with GCN5. The SWIRM domain but also part of the central region of ADA2b were found to be required for interaction with ADA3. Adapted from (Gamper *et al*, 2009).

The SWIRM domain of ADA2b was found to mediate interactions with ADA3. It was shown that ADA2b-SWIRM is able to bind ADA3 but ADA2b lacking a SWIRM domain is still able to interact with ADA3. Regarding the effect of the ADA2b -SWIRM on the catalytic properties of GCN5, it has been shown that removal of the domain does not affect the level of GCN5 activity both in the case of mononucleosomes and chromatin (nucleosomal arrays) (Gamper *et al*, 2009). However, there seem to be specific aminoacids in the SWIRM domain that are particularly important for the effect of ADA2b on catalytic activity of GCN5. Mutations of these residues reduce the efficiency of GCN5 acetylation on mononucleosomes. It has been suggested that this may be due conformational changes that disrupt interaction of ADA2b regions with the nucleosome or directly affect the interaction of SWIRM with the nucleosome (Gamper *et al*, 2009).

Several studies have shown that SWIRM domains are involved in chromatin interaction by binding to DNA or histone tails and can be found in proteins with very different functions (Qian *et al*, 2005; Da *et al*, 2006; Tochio *et al*, 2006). It seems that despite the overall structural similarity the range of DNA/chromatin binding affinity among different SWIRM domains can be broad (Yoneyama *et al*, 2007). Thus, the interactions between proteins with SWIRM domain(s) and chromatin will also vary. Particularly for the SWIRM domain of ADA2a, it has been shown that it binds to double-stranded DNA but not to mononucleosomes. On the other hand it binds to dinucleosomes which may be an indication that it recognises linker DNA (Qian *et al*, 2005; Gamper *et al*, 2009).

A study in *D. melanogaster* has also attributed a role in the C-terminal region of dADA2a and dADA2b for the regulation of the HAT activity of the respective dATAC and dSAGA complexes. dADA2a and dADA2b chimeras were used to test the effects on the function of these domains in transgenic *Drosophila* lines (Vamos & Boros, 2012). Among the different combinations of dADA2a/dADA2b hybrids that were used in this study, it was found that the only hybrid

proteins that could partially rescue dADA2a or dADA2b mutant flies and restore H3 and H4 acetylation were the ones that included the C-terminal region of the respective mutant. For example a dADA2a hybrid in which the whole C-terminal region of dADA2b was added could rescue partially a dADA2b deletion. It must be noted that bioinformatics analysis, performed by the authors of the same study, failed to identify SWIRM related structural motifs in dADA2b. Nevertheless, the results of this study indicate that the C-terminal regions of the two proteins, contribute to a certain extent, and more than the other regions, to the functional specificity of the complexes in which they incorporate.

### 1.5.9 Evidence for distinct functional roles of ATAC and SAGA in gene specific transcription regulation

The differences in subunit composition between two distinct HAT complexes that share a catalytic subunit (GCN5/PCAF) must be reflected to some extent in the specificity of their function. Several studies have attempted to discover, describe, and clarify those differences in the context of gene expression regulation.

In human, SAGA has been implicated in the regulation of genes that are induced under endoplasmic reticulum (ER) stress conditions (Nagy *et al*, 2009). Particularly, increased recruitment of hSAGA subunits (SPT20, SPT3, ATXN7L3) together with Pol II was measured at the promoters of ER stress responsive genes in cells treated with thapsigargin (ER stress inducing reagent). Notably, SPT20 siRNA knockdown that affects the integrity of hSAGA complex significantly reduces the expression of the same genes in stress conditions. A similar link has been demonstrated in yeast where Spt20 deletion prevents yeast cells from activating pathways related to ER stress (Welihinda *et al*, 1997; Welihinda *et al*, 2000).

In addition, a different function in the activation of immediate early (IE) response to 12-O-tetradecanoyl-phorbol-13-acetate (TPA) was suggested ATAC and SAGA in *Drosophila* and human cells. After TPA treatment, increased colocalisation between ATAC subunits (hYEATS, hMBIP homolog) with Pol II was observed on polytene chromosomes of *Drosophila* salivary glands indicating elevated interactions with chromatin. The same was not observed for the SAGA subunit (ADA2b) tested under the same conditions. Similarly, increased recruitment of ATAC, but not SAGA-specific subunits was detected at the promoters of immediate early (IE) genes in HeLa cells after TPA treatment (Nagy *et al*, 2010). Moreover it was found that ATAC was required for the regulation of transcription of these IE even without induction, particularly by affecting H3 acetylation at their promoters.

An interesting example of transcription associated complex specific function involves the interaction of ATAC with the Mediator complex. The Mediator complex (MED) is a large multisubunit complex that acts as transcription coactivator. It is believed that it mediates molecular interactions between transcription activators/repressors, coactivators, general transcription factors and Pol II to regulate the stability of PIC formation and transcription initiation (Conaway *et al*, 2005; Taatjes, 2010). It has been shown that the effect of MED in transcription can be either activating or repressive, depending on its composition that defines its two major forms. Incorporation of the MED kinase module in the complex defines TRAP form that represses transcription, whereas the presence of MED26 defines PC2 form that has an activating role (Malik & Roeder, 2005; Paoletti *et al*, 2006). Mass spectrometry based analysis of ATAC interactome in mouse embryonic stem cells (mESc) showed that ATAC and the activating form MED (PC2) can form a stable complex (Krebs *et al*, 2010). SAGA was not found to be involved in a similar interaction. Moreover, PC2-ATAC meta-coactivator complex (MECO) specifically forms between GCN5 but not PCAF containing ATAC. The same type of ATAC –MED interaction was found in mouse fibroblasts and HeLa cell (although weak). In addition, it was shown that transcription activation of several noncoding RNA genes requires the presence of the identified MECO (Krebs *et al*, 2010). Overall, this study provides strong indications that functional interactions of coactivators can occur in a cell-type specific manner, to achieve gene specific transcription regulation. It also revealed a clear functional difference not only between ATAC and SAGA but also between ATAC complexes that contain GCN5 or PCAF as a catalytic subunit.

Another type of stress-induced pathways in which ATAC and SAGA components seem to be involved are those related to DNA-damage. Depletion of ADA3 in mouse embryonic fibroblasts (MEFs) is sufficient to increase the levels of DNA damage related marks like phosphor-ATM, γH2AX, even in the absence ionising radiation. Additionally, the increase in chromosomal aberrations in the case of ADA3 deficient fibroblasts indicates a role for ADA3 in genomic stability (Mirza *et al*, 2012).  In human cells ADA3 is required to stabilise p53 via acetylation and activation of p53-regulated genes upon DNA-damage. However it has not been specified if these functions are SAGA- or ATAC- dependent (Nag *et al*, 2007). These data were supported from the observation that in human ADA2b but not ADA2a, associates with promoters of p53 dependent genes (Gamper *et al*, 2009). Similarly, in *Drosophila* dADA2b (dSAGA component), but not dADA2a (dATAC component), was described to participate in a Dmp53 (p53 homolog) regulated pathway of gene expression (Pankotai *et al*, 2005).

The studies summarised above illustrate some examples of ATAC and SAGA specific functions in gene specific activation. Moreover the function of SAGA and, to a lesser extent, ATAC has also

been investigated using genome wide approaches. Those studies have elucidated some functional aspects of the complexes but also generated new questions.

## 1.5.10 Genome-wide distribution of human SAGA and ATAC complexes

The information from studies that focus on the roles of SAGA in transcription regulation on limited, selected sets of genes can provide detailed mechanistic information about the function of the complexes binding at these particular genomic sites. To assess the role of SAGA in gene expression regulation at genomic regions other than the promoter but also in a genome wide scale, different types of classic or more recently developed biochemical approaches have been used. There is increasing evidence that the function of SAGA complex, instead of being limited to its interaction with specific promoters and the modification of chromatin state in a local scale is rather global. Thus SAGA may have functional characteristics that fit more to a general co-factor of RNA pol II transcription rather than to a co-activator that regulates a limited set of genes.

The first evidence that Gcn5 may act in genomic regions much broader than a gene promoters came almost 15 years ago. At that point, it was shown that Gcn5 can affect H3 acetylation levels in a region of several kilobases (kbs), flanking a gene that had been previously described as Gcn5 regulated This observation implied that apart from regions that Gcn5 acts to regulate specific genes it can also catalyse histone acetylation in a global scale without specific regulation (Vogelauer *et al*, 2000).

Additional evidence for a broader function of SAGA came from a study that showed that the complex associates with the coding regions of *S. cerevisiae* genes and that its presence in the ORFs (open reading frames) of these genes depends on transcription and phosphorylation of Ser5 of RNA pol II CTD. Results from the same study, supported the idea that SAGA is important for transcription elongation and that is required for efficient nucleosome eviction from highly transcribed genes, facilitating this process indirectly by H3 acetylation (Govind *et al*, 2007). Observations pointing in the direction of Gcn5 acting on the coding regions of highly transcribed genes were also obtained from a genome-wide study (ChIP-chip) on *S. pombe* (Johnsson *et al*, 2009). Other studies, based on ChIP-seq analysis, were conducted in human cells and showed that SAGA has a relatively low number of binding sites genome-wide and that these sites were identified as promoters or enhancer sequences (Vermeulen *et al*, 2010; Krebs *et al*, 2011). Particularly, in the study of Krebs *et al*, (2011), global distribution of ATAC and SAGA binding site was assessed by ChIP-seq. Chromatin IP was performed with antibodies targeting ATAC or SAGA-specific subunits ZZZ3 and SPT20 respectively. Although some patterns of differential binding were identified between the two complexes, the total number of high confidence binding sites for each complex was not more than 400. These results indicated that the role of the

complexes is exhibited at only a few genomic loci. However, the possibility that very transient complex-chromatin interactions, that are below the sensitivity limit of this approach were missed, was not excluded (Krebs *et al*, 2011).

Evidence for a role of the SAGA complex in post-initiation regulation of transcription came from ChIP-seq experiments that examined that localisation of SAGA in muscle cells of *D. melanogaster*. It was found that for the majority of SAGA-bound genes the complex is localised at the promoters. However, in 14% of these genes, SAGA was also found enriched in the coding regions and its distribution was positively correlated with the presence of Pol II (Weake *et al*, 2011).

The idea of SAGA being involved in transcription activation mainly through interactions with promoter elements was supported by a study based on gene expression level analysis in the context of single or combined deletions of SAGA and TFIID subunits in yeast. It was proposed that SAGA dominates the expression of 10-15% of *S. cerevisiae* genes including most of stress-induced genes that have a TATA like sequence in their promoter. The rest 85-90% of yeast genome was characterised as TFIID dominated, including mainly housekeeping genes (Huisinga & Pugh, 2004). The concept of SAGA and TFIID dominated genes was also supported by the more recent ChIP-chip based analysis of 200 transcription regulators performed in yeast (Venters *et al*, 2011). Here, the authors distinguish between the TFIID and SAGA regulated pathways of transcription that seem to have certain characteristics. The SAGA pathway appears to be more specialised, with a more diverse range of interacting factors and distinct mobilisation of the transcription machinery under stress conditions as compared to TFIID.

Nevertheless, it was later shown that there is only a small overlap between genomic locations that the subunits of these complexes bind (as determined by genome wide approaches) and the effect of the deletion of these factors on the expression level of the bound genes (Lenstra & Holstege, 2012). The outcome of the later study suggests that absence of binding of a factor at a particular genomic location may not have a detectable effect to the expression level of the corresponding gene. This might be either because the factor, although it binds, has no functional output at the particular site or due to functional redundancy between different factors (Lenstra & Holstege, 2012). Similar effects have also been demonstrated in studies evaluating the effects of the deletion of specific histone modification in transcription regulation. Like in the case of transcription factors, although a mark may be correlated to active transcription globally, the effects of its deletion affect the expression levels of limited sets of genes (Lenstra *et al*, 2011). It has to be noted that the transcriptomic data used in the comparative study of Lenstra *et al,* (2012), resulted from quantification of total steady state mRNA. As it has been shown recently, total mRNA levels changes can be masked by compensation effects between mRNA decay and

synthesis rates (Dori-Bachash *et al*, 2012; Sun *et al*, 2013). Therefore the above studies cannot be conclusive regarding the effect of the deletions of these transcription factors and particularly SAGA subunits, on the global level of transcription efficiency.

Evidence that clarifies certain aspects of the function of SAGA in a global scale came from the study of Bonnet *et al*, (2014b). The authors of this study used a ChIP-seq approach to investigate the genome-wide function of the complex, but this time it was not the individual SAGA subunits that were used as the targets for chromatin immunoprecipitation. Instead, the distribution of the complex was tracked indirectly by following its two enzymatic activities in *S. cerevisiae* and human cells. Particularly, since the HAT module acetylates H3K9 and the DUB module deubiquitinates H2B, chromatin pull down was performed against H3K9ac and H2Bub marks in wild type cells and cells in which the one or the other SAGA activity was impaired. The comparison of the distribution of H2Bub between cells with active and inactive DUB revealed no change of H2Bub level at the promoters but increased levels of the mark in transcribed regions of almost all expressed genes, both in yeast and human (Bonnet *et al*, 2014b). In addition, it was shown that the level of SAGA-driven H2Bub deubiquitination along genomic regions is homogenous and proportional to H2Bub levels. This observation indicated that the presence of H2Bub per se may directly or indirectly drive SAGA recruitment in these regions. H3K9ac was found enriched at the promoters of all actively transcribed genes and its density was significantly reduced in SAGA-HAT impaired yeast and human cells (Bonnet *et al*, 2014b). Another important finding of the same study was that in conditions when the SAGA complex was disrupted in yeast, the levels of nascent mRNA were significantly reduced for all tested genes and also RNA pol II occupancy was reduced as shown by ChIP-seq analysis conducted on the same yeast strains. These results indicate that SAGA is involved in the regulation of RNA pol II recruitment in a genome wide scale, although the mechanism of this interaction has to be investigated. To test if the global DUB activity of SAGA is dependent on Pol II transcription the removal of ubiquitin from H2Bub was observed in conditions of transcription inhibition under actinomycin D treatment. It was found, that under drug treatment conditions, removal of the marks occurs in a SAGA dependent manner, in a few minutes. This finding provides strong evidence that the interaction of SAGA with its substrate must be very dynamic, since in a very short time its activity seems to affect dramatically the genome-wide abundance of the mark (Bonnet *et al*, 2014b). The results of Bonnet *et al*, (2014b) are also is in agreement with previous studies which showed that deletion of Ubp8 increases the global levels of H2Bub activity  and that depletion of SAGA HAT components decreases the  global levels of  H3K9 acetylation in yeast and in human cells (Henry *et al*, 2003; Bian *et al*, 2011). Although numerous studies have investigated the gene specific functions of ATAC and SAGA by using subunit-specific ChIP approaches, and also the function of SAGA has been studied with a different approach by

following the action of SAGA in a genome wide scale, the systematic evaluation of the dynamics of the complexes in living cells has not been performed.

## 1.6 Investigating mechanisms of action of transcription factors in living cells

In the previous chapters, the role of some of the most important components related to RNA pol II transcription initiation regulation was summarised and a general description of transcription-associated functions of SAGA and ATAC HAT containing complexes was provided. The information regarding the role of these factors comes mainly from experimental approaches that could be grouped in two general categories. The first, includes those studies in which the interaction of regulatory factors was investigated with biochemical assays *in vitro*, using different sets of highly purified proteins, protein complexes or cellular extracts. This approach was extremely valuable, as many aspects of endogenous transcription regulation and complex subunit interactions were successfully reconstituted and dissected in the simplified and controlled environment of the test tube. The second broad category could include all the studies that exploited the findings of *in vitro* approaches to design functional experiments that in the majority of cases were conducted on cell populations (e.g. ChIP-seq, mRNA measurements etc.). Studies of the later type were particularly valuable as they revealed that within the cellular environment transcription regulation involves a complex network of protein interactions.

The interpretation of data derived from both types of approaches seemed to support the idea that molecular events related to transcription regulation occur in a sequential, highly ordered way. A characteristic example of such kind of interpretation regards the pathway of PIC assembly. As described in chapter 1.3.2, according to the classical view individual PIC components cannot be recruited to this multiprotein assembly unless certain components have assembled before. According to that notion, each protein and/or protein complex can be recruited only at a very specific step within a series of molecular events. Consequently, protein-protein and protein-DNA complexes can be formed by the stable and well-ordered binding of the factors that constitute them.

However, advances in fluorescence microscopy techniques, together with the establishment of green fluorescent protein (GFP) (Tsien, 1998) as a broadly used genetically encoded fluorescent tag, allowed the direct observation of protein movement in single living cells and revealed the dynamic nature of protein-protein and/or protein-DNA interactions.

## 1.6.1 Live-cell fluorescence microscopy techniques reveal transcription factors dynamics

Advanced live-cell fluorescence microscopy techniques have allowed the study of kinetics of transcription factors interactions with chromatin and also comparison of dynamics between components of multisubunit complexes. Moreover, endogenous or artificial gene arrays have been proved valuable tools for the study of transcriptional output in real time. Importantly, with the application of proper live microscopy techniques the above processes can be resolved in a much shorter time scale than what is achieved with most biochemical approaches. Thus the combination of biochemical and live imaging approaches can provide a more complete picture of the function of transcription associated complexes. Fluorescence Recovery After Photobleaching (FRAP) and Fluorescence Correlation Spectroscopy (FCS) are two techniques applied in the present study and have also been used in many other cases to assess the dynamics of proteins in living cells. In the following section, their basic principles are described. Right after certain key findings of studies investigating dynamics of transcription factors in living cells with the above techniques will be summarised.

### 1.6.1.1 Basic principles of FRAP and FCS

Many of the studies investigating the dynamics of nuclear proteins involved in transcription regulation and related cellular processes (i.e. replication, DNA repair), have employed FRAP and its variations. In FRAP, a laser pulse of high and short duration is applied in a well define region of interest (ROI) of the cell to irreversibly bleach fluorescent molecules (e.g. GFP-tagged transcription factors). The quantification of fluorescence recovery within the selected ROI over time provides information about the mobility of the protein of interest (Figure 1.12, B). Two major factors affect the recovery rate of a protein within the nucleus. The first is the diffusion of the protein and the second its binding interactions with other components within the selected region of the nuclear environment (Sprague & McNally, 2005).

In the case of proteins involved in transcription regulation, at least two types of binding could affect their mobility: the first involves interaction with components of complexes in which the fluorescently tagged factor will incorporate and the second could include binding events on chromatin/DNA. Qualitative comparison of FRAP curves between different proteins can provide information about their relative mobility (Figure 1.12, A). In this case a prerequisite is that the FRAP curves have to be obtained using the same imaging protocol. The half-time of recovery ($t_{1/2}$) and the mobile fraction ($F_M$) are two parameters deriving from the semi-quantitative analysis of FRAP curves. $F_M$ represents the fraction of bleached molecules that is replaced by

unbleached molecules during the time that fluorescence recovery is recorded. High $F_M$ indicates that the protein does not interact stably with any other component within the selected environment. For transcription factors studied by FRAP in the nucleoplasm, high $F_M$ most probably indicates transient interactions with chromatin. In FRAP $t_{1/2}$ is the time needed for fluorescence recovery to reach 50% of the maximum intensity. Thus, $t_{1/2}$ is indicative of the rate of fluorescence recovery and can be related both to the size of the tagged protein but also to the interactions of the given protein with the nuclear environment.
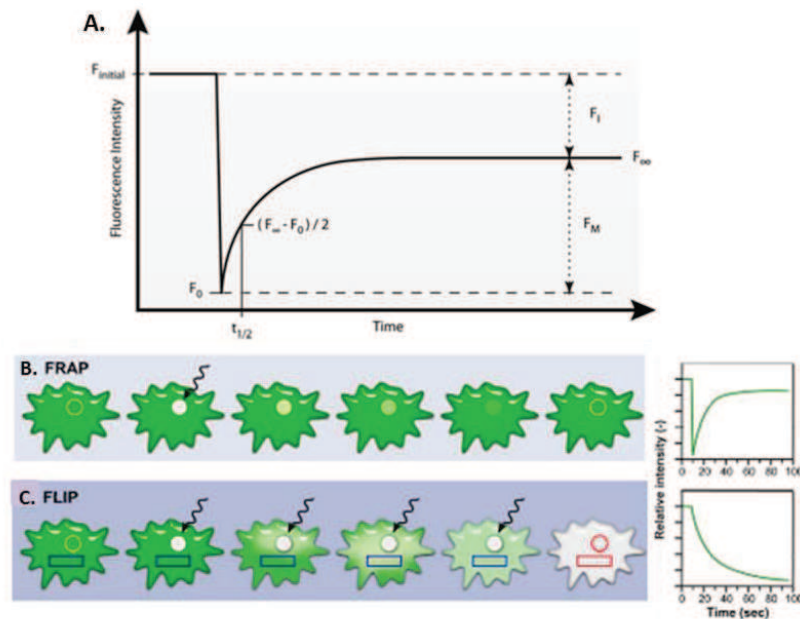


**Figure 1.12. Schematic representation of a FRAP and FLIP experiment.**
**A:** In FRAP, fluorescence intensity is monitored over time to estimate the half time of recovery ($t_{1/2}$), and the mobile and immobile fractions ($F_M$ and $F_I$, respectively). **B.**: In FRAP, a region of interest (ROI) is selected, bleached with an intense laser pulse, and the fluorescence recovery in the ROI is measured over time. **C.:** In FLIP, experiments involve the selected ROI is bleached repeatedly during the entire monitoring period and the fluorescence intensity in a region(s) outside the selected ROI is quantified. Modified from (Carisey *et al*, 2011) and (Ishikawa-Ankerhold *et al*, 2012).

$F_M$ and $t_{1/2}$ are two parameters that can be used only for the comparison of data obtained under the same imaging conditions, to provide information about the overall mobility of a factor (Sprague & McNally, 2005; Carisey *et al*, 2011). To obtain more detailed quantitative parameters from FRAP curves the appropriate mathematical models have to be used. In that way, protein mobility parameters like association/dissociation rates from DNA or DNA bound complexes, and diffusion constants ($D$) can be calculated. The estimation of diffusion constant (measured in $\mu m^2 . s^{-1}$) of a protein is particularly useful as it is a parameter independent of the experimental set up and allows comparison of data obtained in different studies. Several variations of FRAP

(iFRAP, FDAP, FLAP, FLIP) have been described (Ishikawa-Ankerhold *et al*, 2012). Among these FLIP (Fluorescence Loss In Photobleaching) has been used in several studies both to determine the mobility of transcription factors in the nucleus but also to study protein shuttling between different cellular compartments (e.g. transport between the cytoplasm and the nucleus)(Köster *et al*, 2005). In FLIP, a ROI is bleached repetitively and changes in fluorescence intensity are monitored at a distinct region (Figure 1.12, C). Like in FRAP, qualitative or quantitative analysis of FLIP curves is possible. However, in this case decay rather than recovery of fluorescence signal is observed.

Another technique that has been used in numerous studies to investigate the mobility of protein complexes in living cells is FCS (Magde *et al*, 1972; Haustein & Schwille, 2007). FCS is based on the analysis of small fluctuations of fluorescence that occurs in a focal volume of measurement (Figure 1.13). The most common FCS protocols use confocal or two-photon excitation microscopy to define a volume (at the order of femtoliter) in which the diffusion of fluorescent proteins creates characteristic patterns of fluorescence intensity fluctuations. These fluctuations can be quantified by autocorrelating the recorded fluorescence signal in time. A mathematical transformation of the signal allows one to draw an autocorrelation curve G(t), from which the concentration of fluorescent particles and their average diffusion coefficient can be extracted (Haustein & Schwille, 2007; Bulseco & Wolf, 2013).
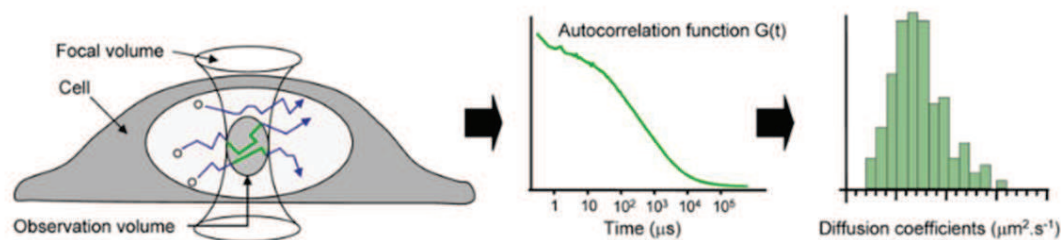


**Figure 1.13. Principle of fluorescence correlation spectroscopy**
Fluorescence fluctuations due to diffusion are recorded in a very small volume. Diffusion of mobile molecules can be measured in the microsecond range. Mathematical transformation allows the generation of an autocorrelation curve G(t). Mobility parameters, like diffusion coefficients, derive from the fitting of autocorrelation curves to the appropriate diffusion models. Adapted from (Gelman *et al*, 2006).

It is important to note that as the concentration of fluorescent molecules in the volume of measurement increases, the relative fluctuation of fluorescence intensity is decreased. Thus, it has been suggested that to sufficiently resolve fluorescence fluctuations no more than 1000 molecules should be present in the femtoliter focal volume (Haustein & Schwille, 2007). Using the appropriate mathematical modelling FCS autocorrelation curves can be used to determine several mobility parameters of the examined protein. Quantitative information includes the

estimation of diffusion time (and diffusion constant), concentration of molecules, and stoichiometry of interactions. In FRAP, due to averaging, a population of molecules with distinct diffusing characteristics might be analysed as a single species, and the calculated mobility parameters will represent an average that might hinder biologically meaningful information. On the contrary, FCS can resolve multiple diffusing components and this property makes this technique particularly useful for the study of very dynamic complexes and their interactions with the nuclear environment. If a fluorescently labelled protein binds to a larger structure or molecule that diffuses slower, binding interactions can be detected. In such types of interactions the obtained FCS data are fitted by a model that includes two diffusing components a fast and a slow. The former will correspond to the free molecules (unbound molecules) whereas the later fraction corresponds to bound molecules (Bulseco & Wolf, 2013).

### 1.6.2   Functional dynamics of transcription factors revealed by FRAP and FCS

The application of FRAP for the study of nuclear proteins involved in transcription initiation, DNA repair and gene splicing revealed the highly mobility of these factors (Houtsmuller *et al*, 1999; McNally *et al*, 2000; Phair & Misteli, 2000). Soon after, different components of the transcription machinery were studied by FRAP in different cell types. Although variations between studies exist, the general observation is that overall, transcription associated proteins are more mobile than initially thought, but a broad range of dynamics has been observed (Figure 1.14).

For example, TFIIB a key component of the PIC has been described in different studies in mammalian cells as a factor which, in FRAP experiments, recovers in the order of seconds (Chen *et al*, 2002; Ihalainen *et al*, 2012). Similar observations were made in yeast where TFIIB was also shown to be highly mobile (Sprouse *et al*, 2008), supporting that the dynamic behaviour of the factor is related to its highly conserved function. TFIIH, a key factor for transcription initiation and also DNA repair, was also found to exhibit short recovery times in cultured human cells expressing XPB-GFP. On the contrary, the interaction of the same factor with DNA lesions at which it is recruited is much more stable. The estimated residence time for TFIIH in the latter case, is in the order of minutes (Hoogstraten *et al*, 2002). Interestingly, when the same factor was studied by FRAP in cells of a knock-in mouse model expressing fluorescently tagged TFIIH (again by XPB-GFP ) a much higher fraction of the factor was found to be immobile, indicating more stable binding interactions (Giglia-Mari *et al*, 2009). The last two studies stress the fact that a given factor may exhibit much different mobility depending on the cellular context and the molecular process it is participating.
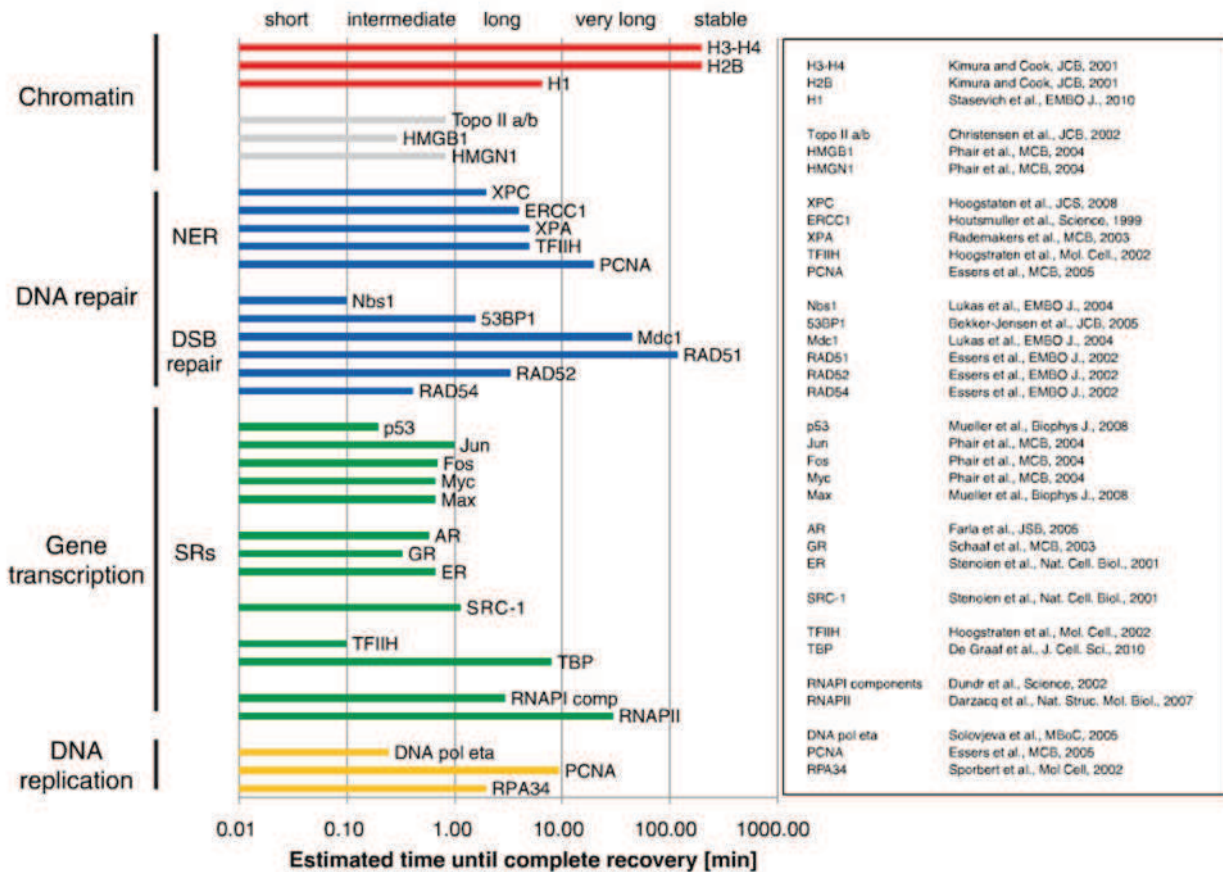
**Figure 1.14. FRAP studies reveal a broad range of dynamics for nuclear proteins.** Estimated time until complete recovery of different nuclear factors in FRAP experiments (log plot). Note that among transcription associated proteins, TBP and RNA pol II have the longest times of full recovery. Adapted from (van Royen *et al*, 2011).

FRAP studies have determined several aspects of the dynamics of TBP, key component of the TFIID complex, that from biochemical studies is known to stably interact with DNA *in vitro*. Experiments performed in yeast (Sprouse *et al*, 2008), and human cells (Chen *et al*, 2002; de Graaf *et al*, 2010) have shown that interactions of TBP with chromatin are much more stable than other PIC components like TFIIB. Initially, results obtained from yeast (Sprouse *et al*, 2008), and human cells (Chen *et al*, 2002), showed that TBP in human is much less mobile than in yeast. However, a comprehensive description of TBP dynamics by FRAP (de Graaf *et al*, 2010), showed that the largest fraction (70%) of human TBP exhibits highly dynamic interactions with chromatin. Nevertheless, the estimated binding time of the immobile fraction remained high, presumably due to stable interaction of this population of molecules with chromatin. The authors of the later study included in their experiments two more subunits of TFIID, namely TAF1 and TAF5. Interestingly, FRAP analysis showed that both TAF1 and TAF5 recover faster than TBP, and have a much lower immobile fraction than TBP. This finding supports the idea

that the components of a given complex (functional or assembling) may exhibit different dynamics in the nucleus.

Another surprise came from the investigation of the mobility of Pol II core subunits which was expected to be slow, as they interact very stably with DNA. FRAP curves of these subunits correspond not only to a slow but also to fast diffusing population of molecules. Investigation of RNA pol II dynamics was first performed by FRAP on a GFP tagged RPB1, the largest catalytic subunit of the enzyme (see 1.2). Two distinct fractions of molecules were identified. The first, which is 75% of the total, moves rapidly whereas the remaining 25% represents a transiently immobile fraction. The fraction of molecules that contribute to the fast part of RPB1 FRAP curves corresponds to RNA pol II molecules during transcription initiation. On the other hand, the slower fraction could represent molecules that are either engaged in elongation or exist in a paused state on the DNA template (Kimura *et al*, 2002a). These observations are similar to what was shown in later studies where the different populations of RNA pol II were further investigated (Hieda *et al*, 2005; Darzacq *et al*, 2007). In the study of Darzacq *et al*. (2007), an artificial array with Tet-inducible promoter in human cells was used to study Pol II kinetics. FRAP analysis of kinetics of RNA pol II recruitment in combination with live imaging of nascent mRNAs, confirmed that the majority of the enzyme interacts rapidly with the promoter whereas only a small fraction is efficiently engaged in elongation (Darzacq *et al*, 2007). Results in the same direction, have also been obtained from the investigation RNA pol I subunits dynamics. First, FRAP experiments showed that recovery kinetics between different subunits of the enzyme differ significantly when measured on endogenous ribosomal genes (Dundr *et al*, 2002). Later, it was also shown that RNA pol I subunits localise differentially during mitosis when rDNA transcription is repressed (Chen *et al*, 2005). In addition, more evidence obtained by FRAP and confirmed by chromatin immunoprecipitation showed that upon transcriptional stimulation of rDNA genes, RNA pol I subunits are more efficiently recruited to promoters (Gorski *et al*, 2008).

Apart from the FRAP experiments described above, several research groups have employed FCS based approaches in living cells to study the interactions of transcription factors with the nuclear environment. As FCS requires more complex instrumentation and computational analysis compared to FRAP, this technique is not as broadly applied as FRAP. In certain cases, FCS has been combined with FRAP since the combination of two approaches can provide complementary information (with different time resolutions) about the dynamic properties of a protein. Moreover, the combination of the two approaches can be used to validate the conclusions made from each individual method (Yao *et al*, 2006; Stasevich *et al*, 2010; Mazza *et al*, 2012). For example, Yao *et al*, (2006), combined FRAP with FCS in a study that investigated the dynamics of association of transcription activator HSF (heat shock factor) with the native

loci of its target gene (*hsp70*). Measurements with both techniques were performed in polytene nuclei of living *Drosophila* salivary glands. FRAP analysis showed that in non-heat shock (NHS) conditions HSF is transiently interacting with the chromosomal loci of *hsp70*. After heat shock (HS), FRAP of HSF was much slower in the same loci. FCS was used to decipher whether the change observed by FRAP, is due to different diffusion mobility of HSF or due to increased association with its target loci. Thus, FCS was performed to assess the diffusion of HSF in the nucleoplasm in NHS and HS conditions. Indeed, the authors found that HS only slightly affects the diffusion of HSF. Thus, slower FRAP recoveries do indicate more stable association of HSF with its target loci (Yao *et al*, 2006). More recently, FCS was used to study the chromatin binding properties of retinoic acid receptor (RAR) in living cells. In this study two distinct eGFP-RAR species were identified based on their different mobility parameters. The fast one, was suggested to correspond to freely diffusing, unbound receptor molecules and the slow one, to the fraction of molecules that transiently interact with chromatin. Importantly, after agonist treatment the authors found an increase of the percentage of slow population of eGFP-RAR molecules, something that could not be detected in previous experiments based on FRAP (Maruvada *et al*, 2003; Gelman *et al*, 2006). Thus, the authors suggested that in their case the difference in eGFP-RAR diffusion was detected due to the increased sensitivity of FCS. Since FCS provides quantitative information on the diffusion properties of an GFP-tagged protein in the microsecond time range, it can reveal changes which may be undetected in an averaging technique like FRAP.

Studies like the ones summarised above have challenged the view that transcription machinery is stable or that transcription factors have long residence times on their genomic sites of action. If that was the case FRAP recoveries of all complex subunit should be in the range of several minutes. However, FRAP measurements for the majority of tested factors, clearly show that rapid binding is a common characteristic of transcription factors which results in fast recoveries. On the other hand many of the factors studied so far appear in the nucleus in the form of at least two distinct populations: one fast, transiently interacting with chromatin and a slower involved in more stable binding events within protein-protein or protein-DNA complexes. These stable binding events are particularly informative for the function of these factors. Moreover, the broad differences in the mobility of different factors show clearly that their dynamic properties are related to the function of each protein or protein complex. In addition, live-imaging approaches allow the investigation of regulatory interactions that cannot be easily revealed with techniques that resolve molecular events at longer time scales. The genome wide patterns of the functional output of transcription factors, seems to be reproducibly exhibited in population studies like ChIP-seq, DamID, etc. Thus it would be interesting to elucidate the regulatory mechanisms which assure a reproducible functional output from rather stochastic and dynamic interactions.

## 1.7 Paradigms and implications of transcription-associated complexes assembly regulation

Some of the FRAP studies described above revealed very different nuclear dynamics between components of the same complex. This was particularly obvious among components of TFIID (de Graaf *et al*, 2010), RNA Pol I (Dundr *et al*, 2002), and RNA pol II (Darzacq *et al*, 2007; Sprouse *et al*, 2008). Such types of observations indicate that differential exchange on a promoter could be related to on site assembly of complex subunits or subassemblies. Thus the assembly regulation of the complexes and function could be tightly linked. Although a plethora of studies has been conducted to elucidate the function and the structure of transcription-associated multisubunit complexes, the information on the regulation of their biogenesis and assembly routes in living cells is not as extensive. Nevertheless, there is evidence that important regulatory steps of complex assembly can be related to the nuclear import of complex subunits that have been preassembled in the cytoplasm as subcomplexes or full complexes.

RNA polymerase II (Pol II), is probably the transcription-associated complex that has been studied the most in term of its biogenesis. The main reasons for that are: its apparent role in transcription, its multisubunit composition, the large amount of structural and biochemical data about the complex and also its high degree of conservation between lower and higher eukaryotes (Cramer *et al*, 2008). Over the last few years, different labs have exploited the advancements in quantitative mass spectrometry (MS) based proteomics, that in combination with the appropriate biochemical and imaging experiments have revealed various aspects of the regulation of RNA pol II assembly (Jeronimo *et al*, 2007; Boulon *et al*, 2010; Forget *et al*, 2010; Czeko *et al*, 2011; Forget *et al*, 2013).

An interesting finding came from the study of Boulon *et al*, (2010) where the authors showed in human cells that RPB1, the largest catalytic subunit of RNA pol II, can be imported in the nucleus only when incorporated to fully assembled complexes. siRNA mediated knock down of any of Pol II subunits resulted in the cytoplasmic accumulation of RPB1. Moreover, α-amanitin treatment, which leads to RPB1 degradation, was found to promote RPB3 cytoplasmic accumulation. These results suggest that the polymerase is assembled in the cytoplasm and that it can be imported into the nucleus only as a fully assembled complex. In addition, the comparison of quantitative MS-based proteomics data between α-amanitin treated (RPB1 degradation and disassembly of Pol II) and control cells, revealed the presence of a cytoplasmic Pol II assembly intermediate formed by RPB2, RPB3, and RPB10-12. It was also shown that cytoplasmic RPB1 interacts with HSP90 R2TP/Prefoldin-like cochaperone hSpagh and GTPase RNA pol II-associated protein 4 (RPAP4)/GPN1. HSP90 and hSpagh stabilise unassembled RPB1 (Boulon *et al*, 2010). The

regulatory role of RPAP4/GPN1 in regulation of Pol II assembly was further elucidated in a second study that showed that downregulation or nuclear export inhibition of this GTPase, that shuttles between the nucleus and the cytoplasm, resulted in cytoplasmic accumulation of RPB1 and RPB2 (Forget *et al*, 2010). However, RPAP4/GPN1 silencing did affect the localisation of the smaller Pol II subunits that were tested. GPN3, a GTPase of the same highly conserved of GPN1 family was also found to interact with RPB4, RPB7, and RPB1 and regulate the import of the later (Carre & Shiekhattar, 2011). In combination, the results of the these studies support the idea that two Pol II separated subcomplexes may exist in the cytoplasm one having as central component RPB1 and the other RPB2. Nevertheless, the possibility that the smaller subunits of Pol II can be independently imported into the nucleus is cannot be excluded. A different role has been suggested for Iwr1 (interacts with RNA polymerase II 1) a factor identified in yeast. Iwr1 contains a nuclear localisation signal (NLS) signal and was shown to bind to Pol II after the RPB1 and RPB2 subcomplexes have assembled. Thus, it has been suggested to act as a molecular sensor that facilitates Pol II holocomplex import through the nuclear pore complex (NPC). Association of Pol II with transcription initiation factors and particularly TFIIB triggers the release of Iwr1 that is subsequently exported to the cytoplasm (Czeko *et al*, 2011).
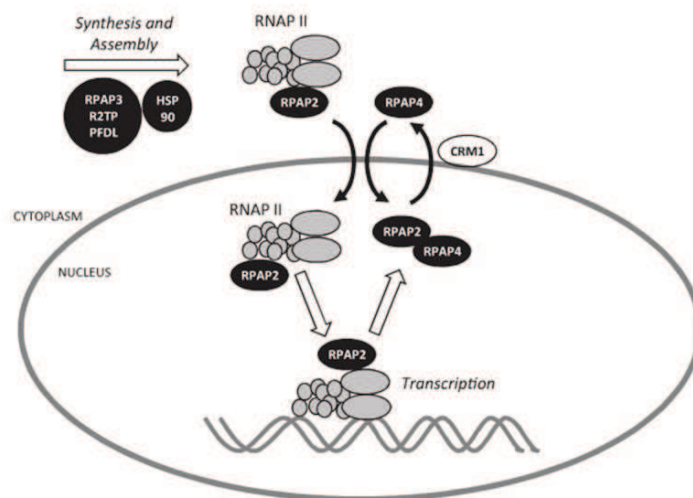


**Figure 1.15. Model for nuclear import of RNA pol II.**
Assembly of RNA pol II subunits is regulated from HSP90 and the RPAP3 (hSpagh)/R2TP/PFDL complex. RNA pol II is imported to the nucleus with the CTD phosphatase RPAP2 that is required for RNA pol II nuclear import. RPAP2 returns to the cytoplasm in a CRM1-dependent nuclear export pathway by associating with the GTPase RPAP4/GPN1. The exact step of RPAP2 dissociation from RNA pol II is not known. It could occur immediately after import or at a later step after transcription termination. RPAP2 regulates transcription both by acting as an RNA pol II CTD phosphatase during the transcription reaction and as an RNA pol II nuclear import factor. Whether RNA pol II is imported to the nucleus in the form of the fully assembled enzyme or sub-assemblies is unclear. Adapted from (Forget *et al*, 2013).

The function of another RNA pol II-associated protein, RPAP2, identified in the affinity purification coupled with mass spectrometry (AP–MS) based analysis of Jeronimo *et al*, (2007), has been investigated and found to be linked with RPAP4/GPN1. RPAP2 shuttles between the nucleus and the cytoplasm with a predominant cytoplasmic localisation. Like its yeast homolog (Rtr1), it has been shown to have a phosphatase activity against RPB1 CTD (Mosley *et al*, 2009; Egloff *et al*, 2012). It interacts both with RNA pol II and RPAP4/GPN1 with distinct domains. Importantly, downregulation or leptomycin-B (LMB) induced nuclear retention of RPAP2 results in RNA pol II accumulation in the cytoplasm (as detected by RPB1). Moreover RPAP4/GPN1 silencing leads to RPAP2 nuclear accumulation indicating that it is exported together with the GTPase. Bioinformatic analysis indicated that RPAP2 structure resembles that of an importin. Therefore, RPAP2- RPAP4/GPN1 were suggested to interact in a importin-GTPase complex that is required for efficient RNA pol II import (Forget *et al*, 2013)(Figure 1.15). These results show a potential link between the regulation of Pol II CTD phosphorylation status and the regulation of the assembly and nuclear import of the enzyme.

The studies described above have revealed that complex and highly regulated interactions between RNA pol II subunits and various regulatory factors are required for the efficient assembly and import of this multisubunit complex. However, due to the complexity of the system, direct information clarifying whether RNAP II is imported into the nucleus exclusively in the form of the fully assembled complex or in the form of sub-complexes has not been provided so far.

At this point, it has to be noted that observations made in two earlier studies suggest an additional level of regulation for RNA pol II productive assembly and import. The first study describes the generation and functional characterisation of a mammalian cell line stably expressing eGFP tagged RPB1 (Sugaya *et al*, 2000). The authors mention, that they observed various patterns of intracellular distribution of eGFP tagged RPB1 among the obtain clones. Importantly, they correlate the expression level of the transgene with its cellular localisation and function. In a clone that a significant fraction of RPB1-eGFP was found engaged in transcription, fluorescence signal was mainly nuclear. On the contrary, in clones highly overexpressing the fusion protein bright nuclear and cytoplasmic signal was observed. In this later case, it is suggested that most RPB1-eGFP represents an inactive pool (Sugaya *et al*, 2000). In the second study, clones of mouse cells stably expressing an α-amanitin resistant form of RPB1 were used to compare the fractions of hyperphosphorylated and hypophosphorylated CTD of the overexpressed RPB1 (Custódio *et al*, 2006). Interestingly, it was found that the higher the expression level of the clone the higher the proportion of hypophosphorylated (unengaged) RPB1 is and also the more it accumulates in the cytoplasm. On the contrary, the level of

hyperphosphorylated (engaged) RPB1 remained stable and localised in the nucleus. Moreover, as it was shown by immunofluorescence against RPB1 in cells overexpressing this subunit, upon LMB treatment a reduction of the cytoplasmic signal occurred and a significant increase in RBP1 accumulated in the nucleus was observed. The results of this study, suggest that overexpressed RPB1 although accumulated in the cytoplasm can indeed enter the nucleus but it is the excess that is exported to the cytoplasm. Similar shuttling behaviour is indicated for endogenous RPB1 (Custódio *et al*, 2006). Thus, when studying the assembly of multisubunit complexes, the possibility that individual subunits or subcomplexes can be imported into the nucleus independently of the holocomplex should also be taken into account and tested with the proper assays.

TFIID is another example of a transcription associated multisubunit complex that has been thoroughly investigated in terms of function and structure (see 1.3.1). However, to date, only a limited set of data regarding the regulation of its assembly is available from living cells. In a study investigating interactions between TFIID subunits, was found that TAF10 nuclear import can be regulated by its histone-fold containing interaction partners, TAF3, TAF8 and SPT7L (Soutoglou *et al*, 2005). It was shown that although endogenous TAF10 is nuclear, when the protein is exogenously expressed it accumulates mainly in the cytoplasm of the cells. Interestingly, it was demonstrated that the simultaneous overexpression of any of its four interacting partners (TAF3, TAF8 and SPT7L) was sufficient for the localisation of overexpressed TAF10 in the nucleus. This event was found to be dependent on the presence of NLS signal found in the sequences of TAF8 and SPT7L. Moreover, FRAP was used to compare the dynamics of overexpressed TAF10 in the cytoplasm and the nucleus. It was found that TAF10 is freely diffusing in the cytoplasm whereas when imported into the nucleus, in the presence of exogenous TAF8 or via fusion with heterologous NLS signal, the recovery rate becomes slower indicating increased association with nuclear structures. Moreover, photobleaching of the whole nuclear pool of exogenously expressed TAF10 showed that no significant fluorescence recovery occurred for 30mins. The latter observation shows that overexpressed TAF10 cannot diffuse in the nucleus and stably accumulates in the cytoplasm. *In vitro* experiments supported this view as it was shown that the interaction of TAF10 either with TAF8 or TAF3 is necessary for binding to importins. This set of data suggests that the formation of the above heterodimers is required for the nuclear import of the TAF10 (Soutoglou *et al*, 2005).
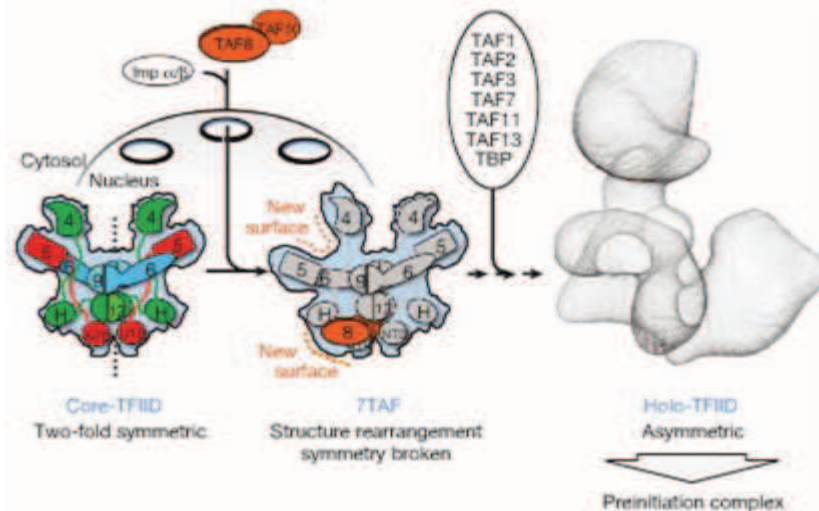
**Figure 1.16. Model for holo-TFIID assembly.**
Core-TFIID consists of two copies of TAF4, TAF5, TAF6, TAF9 and TAF12 that assemble in two-fold symmetric subcomplex (left). The TAF8–TAF10 complex (orange) is imported into the nucleus by importins (Soutoglou *et al*, 2005) (top). Binding of one copy of TAF8–TAF10 breaks the symmetry in core-TFIID. The 7TAF complex (middle) that exhibits two distinct halves, and new binding surfaces important for the incorporation of additional subunits (dashed lines). Accretion of remaining TAFs and TBP in a single copy, results in the formation of asymmetric clamp-shaped holo-TFIID complex (grey mesh) that nucleates the preinitiation complex (right). Adapted from (Bieniossek *et al*, 2013).

Findings from this study have been recently integrated in a model for the assembly of TFIID holo-complex that is mainly based on cryo-EM structural data (Bieniossek *et al*, 2013)(Figure 1.16). Particularly, it has been suggested that it is the binding of TAF10-TAF8 heterodimer that binds to the core-TFIID complex and disrupts its two-fold symmetry. This event seems to be crucial as it creates an asymmetric 7TAF complex providing the molecular surfaces that are required for the interaction and incorporation of TBP and the remaining TAFs in a Holo-TFIID complex (Bieniossek *et al*, 2013). This model signifies the importance of the regulation of nuclear import of individual complex components for the assembly of functional holocomplexes.

Data from the studies summarised above suggest that cytoplasmic assembly mechanisms are highly regulated and probably involve the formation of assembly intermediates. Moreover, the function of several additional factors, that are not part of the final complex, is required for precise regulation of the assembly process. These factors seem to function both during the association of the sub-assemblies of a complex but also to guide the nuclear import and export of individual complex subunits.

An additional level of complexity to these regulatory mechanisms comes from the fact the many of these transcription associated complexes have common subunits. Apart from the subunits shared between the three classes of RNA polymerases there are several other complexes with common components. As it was mentioned in chapter 1.3.1, TBP functions is three different complexes (SL1, TFIID and TFIIIB) involved in RNA pol I, pol II or pol III transcription (Hernandez, 1993). Moreover, several TAFs, like TAF10 mentioned above, are shared between TFIID and the SAGA complex (Nagy & Tora, 2007). As it was mentioned in chapter 1.5.7, SAGA and ATAC complexes have three common subunits (i.e. GCN5/PCAF, SGF29, ADA3) which are all found in the HAT module of each complex. Furthermore, WDR5 subunit found in ATAC complex is also present in histone acetyltransferase and methyltransferase complexes (Shilatifard, 2008; Cai *et al*, 2010; Nagy *et al*, 2010). These examples indicate that apart from the elucidation of mechanism related to assembly of a specific complex, distinct pathways that regulate the allocation of shared components among functionally distinct complexes should be also investigated.

# 2 Materials and methods

## 2.1 Cell culture

U2OS, human osteosarcoma cell line was grown at 37˚C in 5% $CO_2$ in Dulbecco's Modified Eagle's medium (Invitrogen) supplemented with 1 g/l glucose, 10% fetal calf serum (FCS)  and gentamycin (40 μg/ml). HeLa cells used for the preparation of nuclear and cytoplasmic extracts were grown in suspension cultures (cultures of approximately 6 l, $\sim 55 \times 10^8$ cells,) in Minimal Essential Media Spinner Modification (SMEM: M-8167, Sigma Aldrich) supplemented with 7% newborn calf serum, 2 mM Glutamine and 40 μg/ml gentamycin at 37°C in 5% $CO_2$ environment.

## 2.2 eGFP constructs and cDNAs

The construct expressing eGFP-RPB1 was kindly provided by Dr Marc Vigneron (ESBS) and has been described in (Sugaya *et al*, 2000). The following cDNAs were PCR amplified and subcloned in pcDNA™3.1(+) (Invitrogen ) with an N-terminal eGFP fusion: hGCN5 and mADA3 described in (Demény *et al*, 2007), hSPT20desribed in (Nagy *et al*, 2009), hZZZ3 (obtained from OriGene Technologies, Inc), hSGF29 and hYEATS2 previously cloned by Dr Zita Nagy and Dr Anne Riss in our lab. cDNAs of ADA2a described in (vom Baur *et al*, 1998), ADA2b (previously cloned by Dr Zita Nagy and Dr Anne Riss), TFIIB described in (Moncollin *et al*, 1992) were all fused with eGFP at their C-terminal end. pBABE vectors expressing eGFP-TBP and eGFP-TAF5, described in (de Graaf *et al*, 2010) were kindly provided by Prof. Marc Timmers (UMCU, Utrecht) and used as PCR templates to subclone the eGFP fused cDNAs in pcDNA3.1. The eGFP-USP22 construct was created by Dr David Umlauf in our lab and has been described in (Umlauf *et al*, 2013). Finally, ADA2a/b swap mutants were generated by Dr Simon Trowitzsch and kindly provided by Dr Imre Berger (EMBL, Grenoble). The cDNA of the swap mutants was PCR amplified from their original vectors and subcloned in pcDNA3.1 with eGFP fused at their C-terminus. The integrity of all constructs was verified by DNA sequencing, performed at GATC Biotech (Germany).

## 2.3 Transfection

For transfections for FRAP, FLIP and FCS experiments cells were seeded on Ibidi µ (35mm) high glass bottom dishes and transfected before completely attached using FuGENE 6 transfection reagent (Promega) according to manufacturer's instructions. Since the transfected amounts of the construct were much lower than the recommended, total DNA amount was supplemented with empty pcDNA3.1 vector. Transfected amounts of eGFP tagged vectors were different for each technique. FLIP: 100-300 ng/35 mm plate. FRAP: 25-50 ng/35 mm plate. FCS: 5-20 ng/35 mm plate. Cell were imaged 16-24 hours after transfection.

## 2.4 Transcription inhibition conditions

To inhibit transcription, cells were incubated for 1 hour with 5 µg/ml actinomycin D (Sigma-Aldrich, #A9415) or 100 µM DRB (Sigma-Aldrich, #D1916). Evaluation of transcription inhibition was performed by incubating the cells with 5-fluorouridine (Sigma-Aldrich, #F5130) for 15–20 minutes. Both drugs were dissolved in dimethyl sulfoxide (DMSO) that was used to treat the control cells. Incorporation of the modified nucleotide was assessed by staining the cells with an anti-5-FU antibody from Sigma-Aldrich (B2531) followed by followed by staining with Alexa Fluor-488 (green) goat anti-mouse antibody (Molecular Probes, OR, #A-11001). Fixed cells were analysed with a Leica epifluorescence microscope (DM4000) equipped with 40x (NA: 0.75) objectives, with a CCD camera (CoolSnap).

## 2.5 Antibodies

Antibodies used: Anti-γ-Tubulin (T6557) from Sigma-Aldrich, anti-histone H3 core (ab1791) was from Abcam. ADA2b 3122 specific rabbit polyclonal antibody was raised against the C-terminal part of the protein (aminoacids 221-421). Serum was purified using Sulfolink Coupling Gel (Pierce) according to the manufacturer's instructions. All the other antibodies used in this study were previously described: anti-ATXN7L3 2ATX-2B1 (Zhao *et al*, 2008), anti-GCN5 2GC2C11 (Brand *et al*, 2001), anti-ADA2a 2AD2A1, anti-ZZZ3 2616 and anti-mADA3 2678 (Nagy *et al*, 2010), anti-SPT20 3006 (Krebs *et al*, 2011), TAF12 22TA-2A1 (Mengus *et al*, 1995), anti-TAF10 (6TA-2B11) (Wieczorek *et al*, 1998).

## 2.6  Preparation of Nuclear and Cytoplasmic extracts

Extraction was performed from an estimated $9 \times 10^{10}$ frozen HeLa cells (obtained from the IGBMC cell culture facility) growing in suspension. Cells were defrozen in cold water and centrifuged at 3000 rpm at 4°C for 10 min. The supernatant was discarded and the cells were resuspended in a total volume of 600 ml of 1X cold PBS. Packed Cells Volume (PCV) is calculated and the cells were centrifuged at 3000 rpm at 4°C for 10 min. The supernatant was discarded and the washed cell pellet was resuspended in 4xPCV solution I (+protease inhibitor coctail (1x)). After incubation on ice in the hypotonic buffer for 30 min, swollen cells were transferred into the dounce and slow up and down movement of a smooth pestle (10x) was used to lyse the cytoplasmic membranes of the cells. Another centrifugation step was performed at 3500 rpm at 4°C for 10 min with the supernatant corresponding to the cytoplasmic extract. The pellet containing the nuclei was resuspended into 4×PCV solution II (+protease inhibitor coctail (1x)) in a dounce where a tight pestle was used (20x) in an up and down movement. Nuclear protein lysate was transferred in a beaker and stirred for 30 min at 4°C. Then the lysate was centrifuge at 15 krpm at 4°C for 25 min. The supernatant (volume V) was collected and precipitated with the gradual addition of ammonium sulfate powder at 0.3 g/ml. After 30 minutes of incubation, the sample was once more centrifuged at 15 krpm at 4°C for 30 min and the supernatant is discarded. Resuspended pellet in 1/5 volume V of dialysis buffer + 1X PIC consisted the final nuclear extract. The extracts can be dialysed overnight in dialysis membrane with a cut-off of 12-14 kDa at 4°C, frozen in liquid nitrogen and stored at -80°C. Final concentration is expected 4-6 mg/ml. Buffer compositions: Solution I: Tris 50 mM pH 7,9, EDTA 1 mM, DTT 1 mM. Solution II: Tris 50 mM pH 7,9, Glycerol 25%, NaCl 500mM, EDTA 0,5 mM, DTT 1 mM.


## 2.7  Immunoprecipitations

The amount of protein used for immunoprecipitations from HeLa NEs was ~20mg and 50-60mg for CEs. Pre-clearing of the protein samples was performed by mixing NE and CE extracts with 1/10 volume of protein A (if rabbit polyclonal antibody) or G (if mouse monoclonal antibody) sepharose beads that were previously washed twice with 10 volumes of MilliQ $H_2O$ and equilibrated in IP100 mM KCl buffer. The mixture is incubated with gentle agitation, for at least 1hr at 4°C and subsequently centrifuged for 5 min at 1000 rpm at 4°C. The supernatant of this centrifugation is the input (IN) used for the IP. For antibody binding to the beads 1/10 of IN volume of protein A or G sepharose beads is first washed twice with MilliQ $H_2O$ and twice with IP100 mM KCl buffer. Then, approximately 1mg of Ab is added per ml of protein A or G sepharose. IP100mM KCl buffer, equal to the volume of beads, is added and the mixture is

incubated at RT for at least 1hr, under gentle agitation. After Centrifugation (5 min at 1000 rpm) the pellet (beads) is kept and washed 2 times with at least 10 volumes of IP500 mM KCl buffer. Finally the beads were washed 3 times quickly with at least 10 volumes of IP100 mM KCl buffer.

Next the resin bound to Ab was mixed with the pre-cleared NE or CE IN and Incubated incubated O/N with gentle agitation. After centrifugation for 5 min 4°C at 1000 rpm, the resin is washed twice 5-10 minutes with 10 volume of IP 500 mM KCl buffer. Another wash of the resin 3 times quickly with 10 volume of IP100 mM KCl buffer is performed before glycine elution. Glycine buffer (pH 2.8, 0.1 M), of equal to the resin volume, is added and incubated for 5 minutes with gentle agitation. The sample is centrifuged for 5 min 4°C at 1000 rpm. and the supernatant is neutralized quickly with Tris buffer pH 8.8. The eluates are centrifuged at high speed 13000rpm for 20 minutes (4°C) to eliminate all unsoluble material that could plug the the Mass spec column . Buffer compositions: IP100 mM KCl buffer (25 mM Tris Cl pH 7,9, 0.1% NP40, 5mM $MgCl_2$, 10% Glycerol, 5mM $MgCl_2$, 100 mM KCl, 2mM DTT, supplemented with protease inhibitor cocktail). IP500mM KCl buffer (25 mM Tris Cl pH 7.9, 0.1% NP40, 5 mM $MgCl_2$, 10% Glycerol, 500 mM KCl, 2mM DTT, supplemented with protease inhibitor cocktail (PIC))

## 2.8 Western blot

Protein samples were separated by SDS-PAGE, transferred to nitrocellulose membranes (Whatman). Membranes were blocked with 3% skimmed milk in PBS for 1 hour at room temperature. Incubation of the membranes with the different primary antibodies diluted in 0,3% skimmed milk in PBS was usually performed over night at 4°C. Membranes were washed (3 x 10 min) with 0.05% Tween in PBS and incubated for 45 min to 1 hour with anti-mouse (GAMpo) or rabbit (GARpo) horseradish peroxidase linked secondary antibody (1:10000 dilution) from Jackson Laboratories. Enhanced chemiluminescence (ECL) detection system (GE Healthcare) was used to revealing antigen-antibody complexes.

## 2.9 MudPIT mass spectrometry analyses

Immunoprecipitated material was submitted to Multidimensional Protein Identification Technology (MudPIT) (Washburn *et al*, 2001), by the proteomics platform of the IGBMC as described in (Thomas-Claudepierre *et al*, 2013). The abundance of proteins of interest in the obtained MudPIT data was estimated by Normalized Spectral Abundance Factor (NSAF) values as described in (Zybailov *et al*, 2006). This approach accounts for the fact that larger proteins

tend to contribute more peptides/spectra. Briefly, to calculate the NSAF for a given protein k spectral counts (SpC) are divided by protein length (L) to provide a spectral abundance factor (SAF=SpC/L). SAF values were then normalized against the sum of all SAF values for all N proteins in the corresponding run, allowing the comparison of protein levels across different runs.

$$(NSAF)_k = \frac{(SpC/L)_k}{\sum_{i=1}^{N}(SpC/L)_i}$$

For all MudPIT experiments described in this study, the specific proteins present in the IP sample were sorted out, if there were minimum 2 fold enriched in the IP as compared to MOCK sample and, as compared to the input sample (either crude cytoplasmic or nuclear extract). The IP enrichment factor was calculated by dividing the SpC value of each protein in the IP sample by the SpC value in the mock control (i.e. glutathione S-transferase (GST)-IP), performed on the same extracts as the specific protein IP or in the input sample

## 2.10 Microscopy

### 2.10.1 FRAP and FLIP

For all live cell imaging experiments in this study, cells were seeded on Ibidi μ (35mm) high glass bottom dishes in phenol red free medium. FRAP was performed using a Nikon Ti-E inverted microscope (with Perfect Focus System (PFS)) equipped with a CSU X1 Yokogawa confocal spinning disk and a 100× CFI Apochromat TIRF oil NA 1,49 (Nikon) objective. Fluorescence signal was detected with a Photometrics Evolve 512 EMCCD camera. During all FRAP measurements cell were maintained at 37°C in 5% $CO_2$ using the Tokai Hit INUBG2E-TIZ stage top incubator. Roper iLas FRAP unit was used to bleach a circular nuclear ROI of 35 pixels radius corresponding to 65 μm². The selected ROI was monitored every 50 ms. After 5 seconds (100 prebleached collected frames) the circular ROI was quickly bleached at maximum laser power with a pulse of 12 ms. Subsequently, recovery of fluorescence was monitored with the same frequency until a plateau of recovery was reached (100 to 600 seconds). MetaMorph software (molecular devices) was used to process the raw images and obtain the average intensity of three ROIs which were used for further normalization: the bleached ROI, the whole nucleus ROI and a ROI corresponding to background signal. EasyFRAP software (Rapsomaniki *et*

*al*, 2012), was only to normalize the obtained background-substracted fluorescence values. With the latter software, the signal of the bleached ROI is normalised to its average prebleach value, which is set to 1. Next, average fluorescence intensity of the whole nucleus at each timepoint is used to correct for signal loss in this ROI resulting from observational photobleaching and fluorescence loss during photobleaching. For our FRAP experiments, the obtained curves were normalised according to the full scale normalization method (Ellenberg *et al*, 1997), in which the value of fluorescence intensity in the bleached ROI for the first postbleach frame is set to zero. Calculation of half time of recovery ($t_{1/2}$) was performed graphically from normalized FRAP curves by finding the x axis value (time) corresponding to 50% of the average the fluorescence intensity values of of the 20 last timepoints of each curve. Since our FRAP curves were always normalised according to the full scale method (where the first post-bleach value ($F_0$) is normalised to 0) the mobile fraction ($F_M$) was calculated graphically by deducting the average normalized intensity values of the last 20 timepoints (corresponding to the plateau of final recovery $F_\infty$) from 1 (that corresponds to initial recovery, $F_I$) and multiplied by 100 (see also Figure 1.12)(Carisey *et al*, 2011).

FLIP was performed using the same hardware setup as the above but using a 60× CFI Plan Apo VC oil NA 1.40 (Nikon) objective. For the experiments described in this manuscript, a circular cytoplasmic ROI of 20 pixels radius corresponding to $61\mu m^2$ was bleached every 5 sec by a double bleach pulse for a total period of 20 min. Like in FRAP, raw images where quantified by MetaMorph software. Average fluorescence values were extracted from four ROIs that are required for the normalization of the FLIP data: the first corresponds to the whole cytoplasm, the second corresponds to the whole nucleus (analysed ROI), the third corresponding to a second cell that was always included to field of view (used to normalize for observational photobleaching) and a fourth region corresponding to background signal. Normalization of FLIP data was done according to the formulas described in (Nissim-Rafinia & Meshorer, 2011). The main differences with FRAP analysis is that in FLIP the analyzed ROI (in our case nuclear) is other than the bleached, which is not used in the normalization. Moreover, in FLIP it is the average corrected value of the first postbleach frame of the analyzed region that is normalized to 1. For estimation of $t_{1/2}$ of loss illustrated in Figure 4.8, FLIP data were fit with GraphPad Prism (GraphPad Software), using a two-phase exponential decay method for data deriving from the swap mutants and a one-phase exponential decay method for FLIP curves of eGFP.

## 2.10.2 Fluorescence Correlation Spectroscopy

Fluorescence Correlation Spectroscopy (FCS) measurements were performed on a home-build two-photon system set-up based on an Olympus IX70 inverted microscope with an Olympus 60× 1.2NA water immersion objective as previously described(Azoulay *et al*, 2003; Clamme *et al*, 2003) at 37° . Two-photon excitation at 930 nm was provided by a tunable femtosecond laser (Insight DeepSee, Spectra Physics) and photons were collected using a set of two filters: a two-photon short pass filter with a cut-off wavelength of 680 nm (F75-680, AHF, Germany), and a band-pass filter of 520 ± 17 nm (F37-520, AHF, Germany). The fluorescence was directed to a fiber coupled APD (SPCM-AQR-14-FC, Perkin Elmer) connected to a hardware correlator (ALV5000, ALV GmbH, Germany) allowing on-line calculation of the normalized autocorrelation function. In living cells, there is no real steady state for the fluorescence intensity fluctuations. For this reason, FCS measurements were sequentially repeated, typically $15 \times 5$ s with an excitation power adjusted to avoid photobleaching of the fluorophore within the excitation volume. Due to the inherent heterogeneity of the cellular medium, it is difficult to use a model with a well-defined number of diffusing species. To make no assumptions about this number, curves were first fitted by using a model that relies on the Maximum Entropy approach (Sengupta *et al*, 2003, Krieger *et al*, 2013). Individual FCS curves were analyzed in terms of a quasicontinuous distribution of diffusing components allowing to retrieve the number of diffusing species. The robustness of the algorithm was tested with free eGFP diffusing in aqueous solution for which a single peak distribution centered on 96 $\mu m^2.s^{-1}$ was obtained.

For all the proteins studied in this work, bi-modal distributions were obtained, demonstrating that the diffusion process was mainly dominated by the presence of two diffusing species with two different apparent molecular weights. In order to work in conditions where the measured parameters are independent of the expression level, a threshold was used to filter out autocorrelation curves giving a number of molecules above 200. To do so, each individual curves were fitted with according to:

$$G(\tau)=\frac{1}{N}\left[(1-F)\left(\frac{1}{1+\frac{\tau}{\tau_{D1}}}\right)\left(\frac{1}{1+\frac{\tau}{S^2\tau_{D1}}}\right)^{\frac{1}{2}}+F\left(\frac{1}{1+\frac{\tau}{\tau_{D2}}}\right)\left(\frac{1}{1+\frac{\tau}{S^2\tau_{D2}}}\right)^{\frac{1}{2}}\right]$$

where N is the average number of fluorescent species in the focal volume, $\tau$ the lag time, $\tau_{D1}$ and $\tau_{D2}$ the average residence time in the focal volume of the two diffusing species, F the fraction of the high molecular weight diffusing species and S a structural parameter defined as the ratio between the axial and lateral radii of the beam waist. The remaining autocorrelation curves

(N<200) were re-analyzed with the Maximum Entropy model to determine the diffusion constant of each individual species. Characteristic examples of the fitting of the average autocorrelation curves of certain factors are shown in Figure 5.1 (Supplementary Data).

In the frame of the Stokes-Einstein formalism, the diffusion constant of a sphere is related to its hydrodynamic radius according to:

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

where $k_B$ is the Boltzmann constant, T the absolute temperature in K, $\eta$ the viscosity and $r_H$ the hydrodynamic radius of the sphere. In this case, the diffusion constant scales with the molecular weight of the sphere at the power one third. By taking as a reference the molecular weight of the free eGFP ($MW_{eGFP}$=27 kD) it is possible to determine the apparent molecular mass (MW) of the labelled protein by using the following relation:

$$MW = MW_{eGFP}\left(\frac{D_{eGFP}}{D}\right)^3$$

where $D_{eGFP}$ and $D$ are the diffusion constants of the free eGFP and the labelled protein.

## 2.10.3 High content analysis

For the microscopy experiments illustrated in the graphs of DAPI-stained cDNA-transfected cells were grown in 96-well μClear Greiner microplates. Imaging was performed with the high content reader INCell1000 analyzer (GE LifeSciences), using a 20×/0.45 NA PlanFluor (Nikon) objective. Image acquisition of an average number of 3,000 cells per well for each combination of transfected vectors was performed in the two fluorescent channels corresponding to DAPI and eGFP. The Multi Target Analyzer software (GE LifeScience) was used to segment the nuclei of the cells according to the DAPI staining. eGFP signal was then analyzed on this cell population to determine the cytoplasm to nuclei intensity ratio of the tagged factor at a single cell level allowing the determination of a percentage of cells showing positive nuclear translocation.

# 3 Manuscript in preparation I: "ATAC and SAGA coactivator complexes are highly dynamic in human cell nuclei and their mobility is unaffected by transcription inhibition."

## 3.1 Results (I)

### 3.1.1 FRAP based approach to investigate human SAGA and ATAC dynamics

As described in chapter 1.5.10, biochemical approaches applied to study the genome wide function of SAGA and ATAC complexes indicate that these co-activator complexes have a limited number of binding sites on chromatin. However, investigation of the mobility of the complexes directly in living cells has not been described. In addition, as it was summarised in chapter 1.6.2, transcription associated protein can exhibit a broad range of chromatin binding interactions. To answer the question of SAGA and ATAC mobility in the nuclei of living cells and to compare their dynamics with other transcription factor, already studied with live imaging techniques, we first decided to use FRAP. Since FRAP has been applied to study the dynamics of several transcription factors in living cells we used this technique to investigate the global mobility of subunits of the two complexes in the nucleus and compare our findings with what was previously found for other factors. The qualitative comparison of FRAP curves between different factors gives information about their mobility and their interactions with the nuclear environment.

To visualise ATAC and SAGA subunits in living cells we used an approach based on genetically encoded fluorescence labelling. The cDNA of several components of the two complexes was cloned in frame with eGFP (see chapter 2.2). For all FRAP experiments described below the fusion constructs were transiently overexpressed in U2OS cells. In all cases only cells with the lowest expression level possible were selected, provided that the signal to noise ratio (S/N) was high enough for imaging and data analysis.

As described in chapter 2.10.1, all our FRAP data were analysed in a semi-quantitative way. This means that half time of recovery ($t_{1/2}$) of recovery and mobile fraction ($F_M$) were calculated graphically for every FRAP fluorescence recovery curve. The result of a semi-quantitative analysis of FRAP data could vary depending on the selected parameters used for the acquisition of the images. We wanted to be able to compare our results with what has been observed before

with live imaging techniques for other transcription factors and also have a measure of mobility for our proteins of interest. For this reason, other transcription factors (apart from SAGA and ATAC subunits) with already characterised intranuclear dynamics were included in most of our experiments as positive controls. TFIIB and TAF5 were selected as transcription factors for which highly dynamic behaviour has been observed in previous studies (Chen *et al*, 2002; Sprouse *et al*, 2008; de Graaf *et al*, 2010; Ihalainen *et al*, 2012). RPB1, subunit of RNAP II was chosen as a factor that has been shown to recover in two phases, one including the unengaged mobile fraction of molecules and the other the population that is engaged in transcription and due its stable interaction with chromatin has a much slower recovery (Kimura *et al*, 2002a; Hieda *et al*, 2005; Darzacq *et al*, 2007; Fromaget & Cook, 2007). TBP was also used, as it is a component of PICs which also has been shown to have a fraction stably interacting with chromatin (Sprouse *et al*, 2008; de Graaf *et al*, 2010). Unconjugated eGFP was used as a control of a freely diffusing protein that should have no specific interaction with nuclear components. All factors used in the FRAP experiment were tagged with eGFP. For convenience and clarity of representation, in most graphs, figure legends and main text the name of each studied factor is mentioned without including the eGFP moiety.

### 3.1.2 Semi-quantitative FRAP indicates that SAGA and ATAC are highly mobile complexes
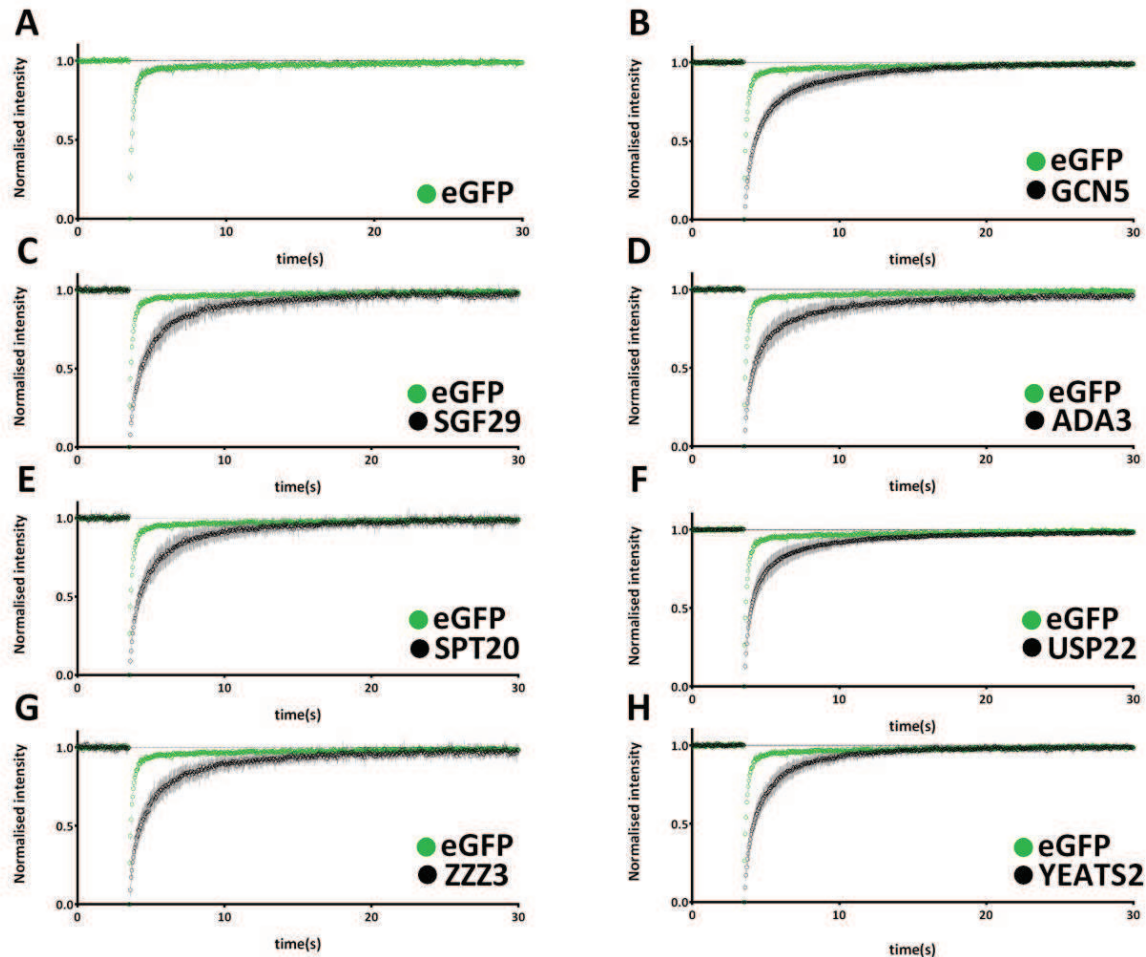


**Figure 3.1. FRAP curves of SAGA and ATAC subunits.**
Graphs show fluorescence recovery kinetics of eGFP, and eGFP tagged human ATAC and human SAGA subunits transiently expressed in human U2OS cells. For all FRAP curves (in this and all FRAP experiments hereafter) the level of prebleach fluorescence is normalised to 1 and bleach depth is normalised to 0. Curves were corrected for background signal and observational photobleaching. For all FRAP graphs the average intensity of the bleached ROI is plotted over time (s). **A:** eGFP FRAP recovery used as a control of freely diffusing protein. The recovery of eGFP is faster than all ATAC and SAGA subunits. **B, C:** recovery curves of GCN5 and SGF29, subunits that are shared between the two complexes. **D, E:** FRAP recoveries of SPT20 and USP22, SAGA specific subunits. **F:** FRAP recovery of ZZZ3, an ATAC specific subunit. In all cases, the cells were imaged for 100 frames before photobleaching the selected nuclear region. Fluorescence recovery was followed for an additional 2000 frames for a total time of approximately 100 sec. To resolve better the first time points of each curve only the first 40 sec are plotted in each graph. Dashed line at Y=1 point indicates level of complete fluorescence recovery. Error bars show ±S.D. from the mean. n>7 for all measurements apart from ADA3 that n=4.

As it has been shown previously, components of the same complex may exhibit different dynamics in the nucleus (Dundr *et al*, 2002; de Graaf *et al*, 2010). We wanted to test whether we can observe such differences among subunits of SAGA and ATAC and also compare the dynamics between subunits of the two complexes. Potential differences could indicate different level of chromatin binding interactions and/or that intranuclear assembly may occur at the sites where the complexes function. Thus, in our first set of FRAP experiments we compared the dynamics of shared and complex specific ATAC and SAGA subunits. As a control representing a freely diffusing protein, unconjugated eGFP was also included in the analysis. As shown in Figure 3.1, we observed that for every tested SAGA and/or ATAC complex subunit the recovery of fluorescence after the bleach occurs within the order of seconds with no detectable immobile fraction. This was the case for common (GCN5, SGF29, ADA3), SAGA specific (USP22, SPT20) and ATAC (ZZZ3, YEATS2) specific subunits. The observation that fluorescence recovery was clearly faster for unconjugated eGFP than any of the tested subunits under our experimental conditions indicated that the ATAC and SAGA subunits are not freely diffusing in the nucleus but rather interacting with chromatin (Figure 3.1).

Apart from comparing qualitatively the recovery curves we wanted to have a measure of comparison for the speed of recovery between the tested factors. This would allow us to detect differences in the mobility between the two complexes or their components. To this end, we calculated graphically the half time of recovery ($t_{1/2}$) for each curve. $t_{1/2}$ is defined as the time between the bleach and the timepoint that fluorescence intensity (F) reaches the half of the final recovered intensity ($F_\infty$). If a given protein A has smaller $t_{1/2}$ in comparison to another protein B, this indicates that fluorescence recovery occurs faster for A. Thus it can be concluded that the mobility of protein A is higher than that of protein B. The comparison of the graphically calculated half times of recovery for our factors of interest is summarised in Figure 3.2. We observed that the estimated range is very similar between shared and complex specific subunits. Components of both complexes appear very dynamic. What is also clear from the same figure is that the average $t_{1/2}$ of recovery for eGFP (~130 ms) is more than 5 times shorter that the average $t_{1/2}$ of recovery (~700 ms)of all tested proteins . This observation further supports the idea that the highly dynamic behaviour we observe is a true characteristic of the function of the two complexes.
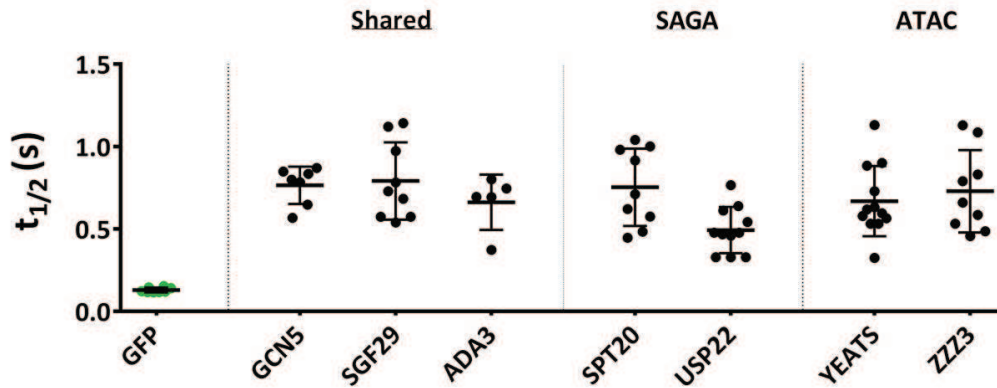
**Figure 3.2. Comparison of half times of recovery between human SAGA and ATAC subunits.**
Scatter plot showing the graphically calculated half times ($t_{1/2}$), of recovery of SAGA and ATAC subunits. $t_{1/2}$ of recovery for all tested proteins was more than 5 fold higher than that of unconjugated, freely diffusing eGFP. No significant differences where observed between shared or complex specific subunits. Mean values are shown by horizontal lines, bars indicate ±S.D. from the mean estimated $t_{1/2}$ of recovery of each subunit.

These results indicate that both SAGA and ATAC complexes are very dynamic in the nucleus. All tested SAGA and ATAC components seem to be highly mobile and apparently they interact transiently with chromatin, at least in a global scale. This is in agreement with biochemical experiments which showed that human ATAC and SAGA have only a limited number (400-500) of high affinity binding sites on chromatin as indicted by genome-wide ChIP-seq analysis (Krebs *et al*, 2011). In good agreement with our results, a study from our lab (Bonnet *et al*, 2014b) recently showed that both in yeast and human cells the SAGA complex acts globally on all actively transcribed genes.

### 3.1.3 The nuclear behaviour of SAGA and ATAC shows that they belong in a class very dynamic transcription factors

As it has been shown, transcription associated factors may exhibit transient or more stable binding interactions with chromatin. Since those interactions can provide significant information for the function of these factors, we decided to further characterise the dynamics of ATAC and SAGA complexes by comparing them with the nuclear dynamics behaviour of other transcription factors. We decided to include in our FRAP analysis other factors for which similar data in living cells was already available. Importantly, our observations regarding the nuclear mobility for TFIIB, TAF5, RPB1 and TBP are in agreement, with previously published FRAP

analyses of the same proteins (see below and Figure 1.14), further supporting the validity of our measurements.

Our FRAP measurements indicated that all ATAC and SAGA subunits show very similar recovery kinetics to TAF5 and TFIIB (Figure 3.3: panels D and F; represented only in comparison with GCN5 for simplicity, compare also with Figure 3.1), which have been described as rapidly diffusing factors only transiently interacting with chromatin (Sprouse *et al*, 2008; de Graaf *et al*, 2010; Ihalainen *et al*, 2012). On the contrary, as shown in Figure 3.3 (panels H and J), FRAP of SAGA and ATAC subunits is faster than that of RPB1 and TBP which have been already described as factors with slow recovery and significant immobile fraction in similar photobleaching experiments (Kimura *et al*, 2002a; Hieda *et al*, 2005; Darzacq *et al*, 2007; Fromaget & Cook, 2007; de Graaf *et al*, 2010).

These observations suggest that the interaction of SAGA or ATAC with components of the transcription machinery and the promoter DNA, may be very transient *in vivo* and/or the promoter/chromatin bound fraction of the SAGA and ATAC complexes represent only a very small fraction of the total pool of these complexes in the nucleus. The comparison of the dynamics of ATAC and SAGA components with the mobility of slow (TBP, RPB1) and fast (TFIIB, TAF5) transcription factors further supports that their function involves very transient interactions with chromatin at a global scale. During our FRAP measurements all subunits of the complexes fully recovered during the post-bleach period, suggesting that almost 100% of the complexes are mobile. However, the existence of an immobile fraction for any of the SAGA or ATAC subunits cannot be ruled out. If immobile fractions exist they would represent only a very low proportion of the total population of molecules that could not be detected with our FRAP protocol.
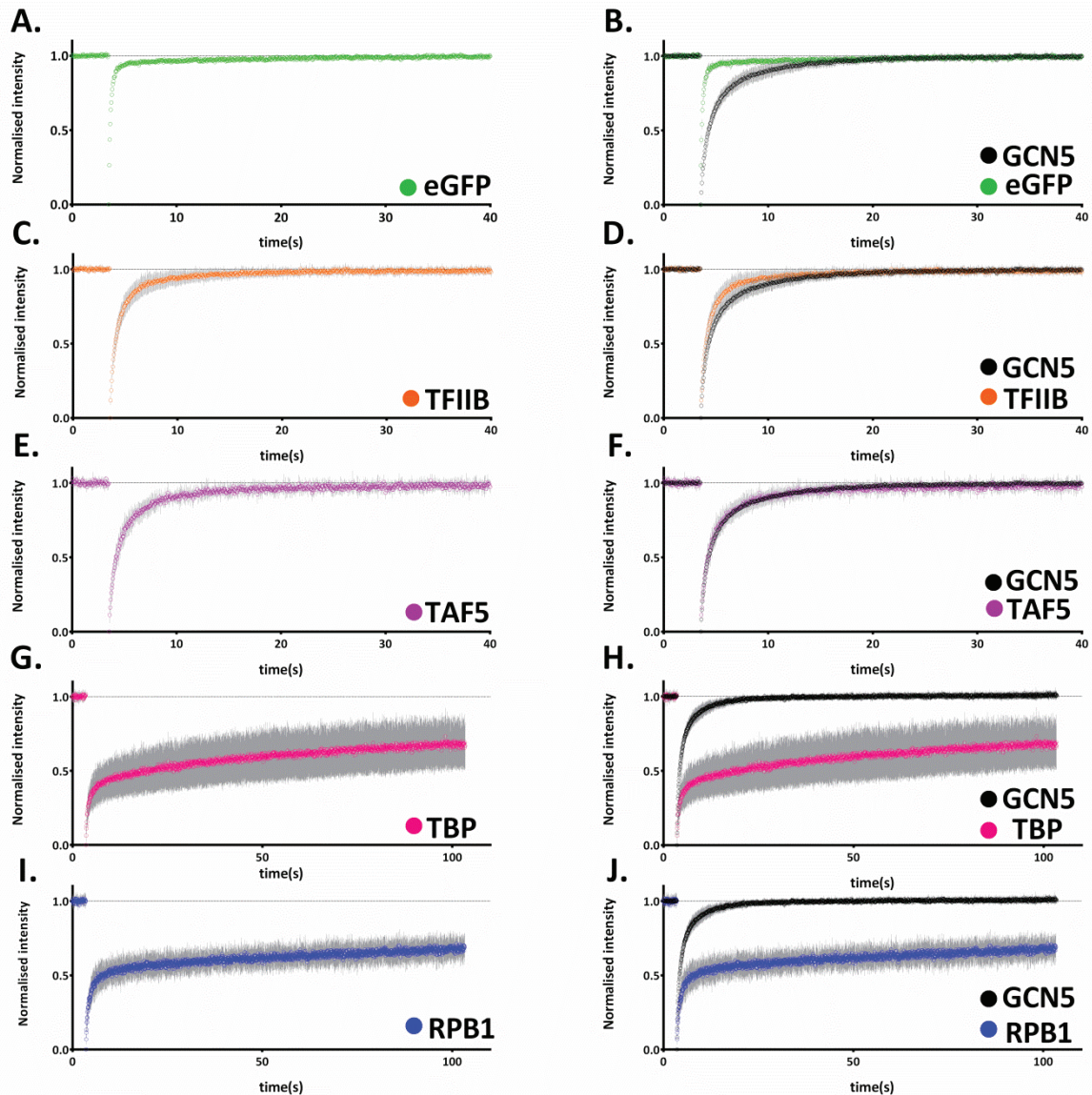
**Figure 3.3 Qualitative comparison between FRAP curves of GCN5 and other GTFs.**
**A, B**: GCN5 recovers in the order of seconds but much slower than the freely diffusing eGFP. **C**: TFIIB FRAP curve show a fast recovery **D:** GCN5 recovery is slightly slower than that of TFIIB. **E, F**: TAF5 FRAP curves and comparison to GCN5 shows similar recoveries**. G, H:** FRAP recovery of TBP subunit and comparison with GCN5. In contrast to GCN5, TBP recovers slowly. **I, J**: FRAP recovery of RPB1, RNAP II subunit and comparison with GCN5. In contrast to GCN5, RPB1 recovers with a slower rate. In all cases the cells were imaged for approximately 100 frames before the bleach and fluorescence recovery was followed for an additional 2000 frames for a total time of approximately 100 sec. Note that in graphs A-F (for eGFP, GCN5, TFIIB and TAF5) only the first 40 sec of recovery are shown to resolve better the first time points. In graphs G-J (TBP, RPB1 and respective comparisons with GCN5) the post-bleach time plotted is extended to 100 sec to better illustrate the differences in the recovery of these factors. Dashed line at Y=1 point indicates level of complete fluorescence recovery. Grey bars indicate SD from the mean.

### 3.1.4 Characterization of the dynamics of SAGA and ATAC complexes after transcription inhibition

We found that SAGA and ATAC dynamics in the nuclei of human cells resemble more to that of highly mobile factors like TAF5 or TFIIB, than that of Pol II or TBP. In the study of Bonnet *et al.* (2014b), the authors tracked the enzymatic activity of the HAT and DUB modules of SAGA by ChIP-seq and other biochemical experiments and found that the complex acts globally on every active gene. Particularly for the DUB function, it was shown that it affects H2Bub levels along the coding region of all actively transcribed genes. Interestingly, the authors demonstrated that transcription inhibition by actinomycin D does not affect the dynamics of the SAGA specific H2Bub deubiquitination. Actinomycin D is a compound that intercalates in double-stranded DNA forming a stable complex and it has been used in many studies to arrest transcribing Pol II. Thus the results of Bonnet *et al.* (2014b) are in good agreement with our findings showing that SAGA is a fast diffusing complex and also show that SAGA retains its enzymatic function independently of active Pol II transcription.

Next we set out to investigate by the dynamic behaviour of SAGA and ATAC coactivator complexes in living cells under transcription inhibition conditions and compare them to the above obtained results. To answer this question we treated human U2OS cells with actinomycin D (5 μg/ml, 1 h), which had been transfected with the eGFP fusion proteins 24-36 hrs earlier. Although the above concentration of actinomycin D is known to efficiently inhibit RNA pol II dependent transcription (Bensaude, 2011), we tested and verified the global inhibitory effect of this drug by 5-FU immunodetection in control and treated cells (Figure 3.4).

After actinomycin D the cells were subjected to FRAP using the same imaging protocol as in the above described FRAP experiments. The comparison of normalised FRAP curves between actinomycin D treated (actD, red curves) and control cells (black curves) is shown in Figure 3.5. Overall, fluorescence recovery kinetics clearly differ between treated and control cells, for every tested protein. Particularly, fluorescence recovery was found slower for TAF5 and TFIIB and all tested ATAC and SAGA subunits. On the contrary we observed a rather faster recovery for RPB1 but as described below the major effect was observed on its mobile fraction.

**Figure 3.4. Actinomycin D treatment efficiently inhibits transcription globally.**
U2OS cells were treated with DMSO (drug vehicle) or actinomycin D for 1 h. Subsequently the cells were incubated with 5-FU for 15-20 min. 5-FU incorporation into nascent transcripts was assessed by indirect immunofluorescence. **A.**: Green signal is indicative of incorporated 5-FU and active transcription. **B.**: After actinomycin D treatment only background signal is detected. Fixed cells were observed in an epifluorescence microscope (DM4000, Leica) with same exposure time (1 sec).

In addition mobile fractions were graphically calculated for every factor (see chapter 2.10.1) and a graph summarising the changes between conditions is shown in Figure 3.6. In the drug treated condition, in most cases, there is a slight change in the immobile fraction of the tested eGFP-fused proteins. Particularly, it seems that for GCN5, SPT20, ZZZ3, YEATS2 and TFIIB, an immobile fraction is appearing in the presence of actinomycin D. On the contrary, in the case of the Pol II subunit (RPB1), the opposite effect is observed; the mobile fraction of RPB1 is increased. This could be the result of an increase in the population of molecules that are not engaged in transcriptional elongation after transcription inhibition. It should be mentioned that in previous studies (Kimura *et al*, 2002a; Fromaget & Cook, 2007) actinomycin D was shown to reduce the mobile fraction of Pol II. Regarding ATAC and SAGA subunits, we found that the mobile fraction of ZZZ3 and YEATS2 (ATAC specific) was more significantly affected than that of SPT20 and USP22 (SAGA specific). The mobile fraction of eGFP was not affected.
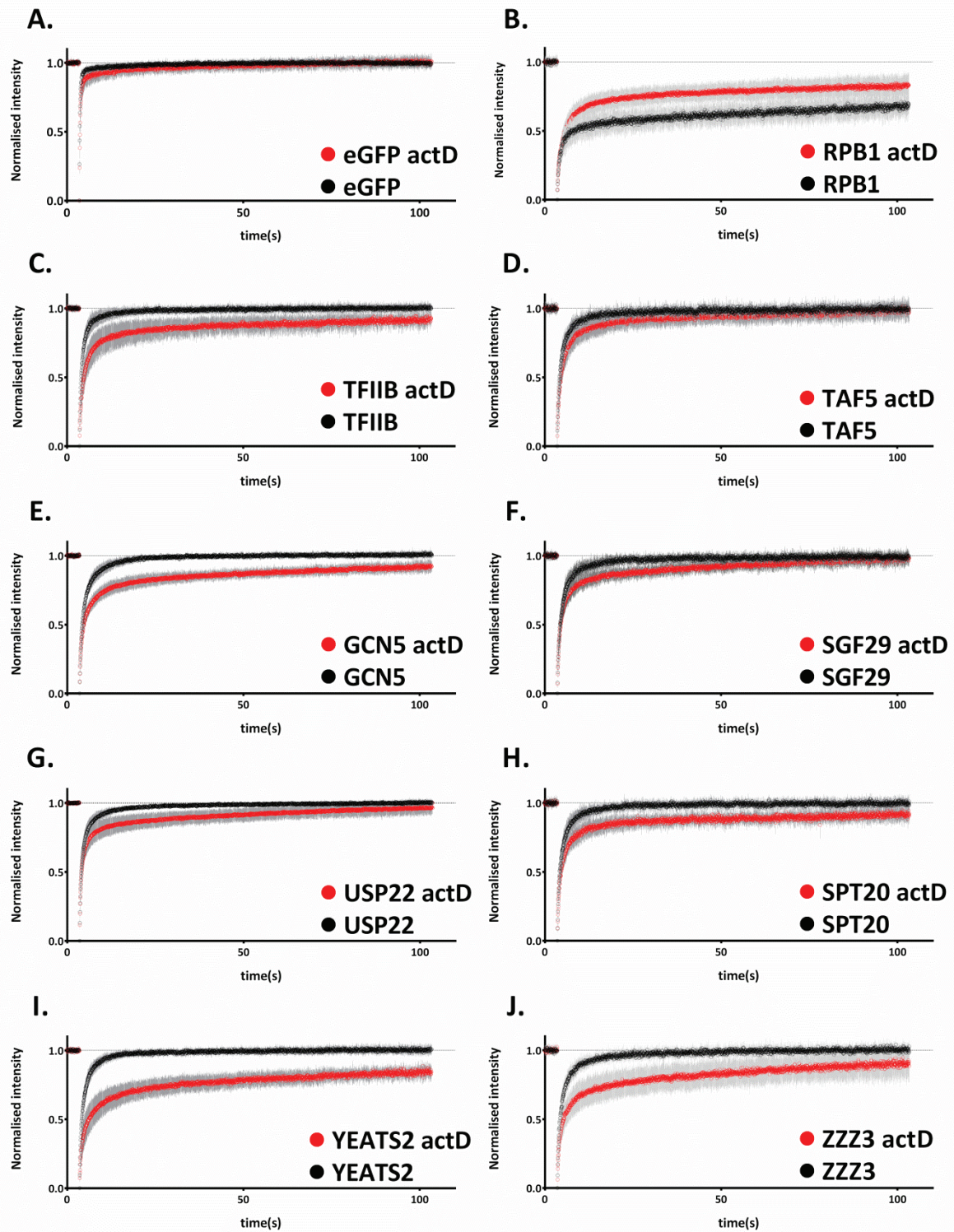
**Figure 3.5. Actinomycin D treatment alters the FRAP curves of all tested factors.**
Qualitative comparison of FRAP curves obtained in control (black circles) or actinomycin D treatment conditions indicates a reduction of mobility for all tested proteins in the drug treated condition. In all cases, except RPB1 (B.), fluorescence recovery is slower and an immobile fraction clearly appears in certain cases (TFIIB (C.), GCN5 (E.), SPT20 (H.), YEATS2 (I.), ZZZ3 (J.)). For RPB1 (B) fluorescence recovery is faster after drug treatment. n>8 for all measured factors.

Next we calculated the half time of recovery ($t_{1/2}$) factors in actinomycin D treated cells and compared them to that obtained in untreated cells. The comparison between conditions is plotted in Figure 3.7. Actinomycin D is significantly increasing the $t_{1/2}$ of recovery of TAF5, TFIIB and RPB1 and has the same effect on the $t_{1/2}$s of the ATAC complex subunits (ZZZ3 and YEATS2). In contrast, there is no detectable effect on the $t_{1/2}$ of recovery of SPT20 and USP22, subunits of the SAGA complex. Although actinomycin D had no effect on the mobile fraction of eGFP, its $t_{1/2}$ was reduced in the presence of the drug. The implications of the later observation are discussed below.
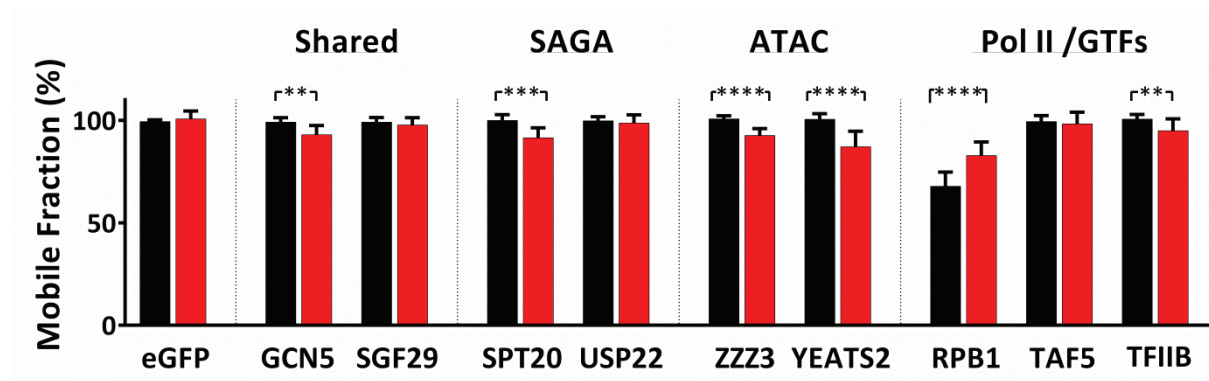


**Figure 3.6. Actinomycin D affects the mobile fraction of most tested SAGA and ATAC subunits.**
Bar chart representing the comparison between the graphically estimated mobile fraction in control (black bars) and actinomycin D (red bars) treated U2OS cells. Unpaired t-test was performed between the indicated groups to estimate the statistical significance of the changes.

Since some factors show changes only in their mobile fraction, but not in their $t_{1/2}$ it seems that only the slower component of each FRAP curve is affected by the drug. This could mean that the mobility of the more dynamic population of the protein is not changing, but the pool of complexes, which were more stably interacting with chromatin are affected by the inhibition of transcription. However the semi-quantitate analysis of the FRAP curves that we performed is not enough to support such a conclusion. To obtain more precise information from our FRAP experiments our FRAP curves should be fitted with an appropriate model to obtain more information about the underlying changes in the mobility parameters of the tested proteins.
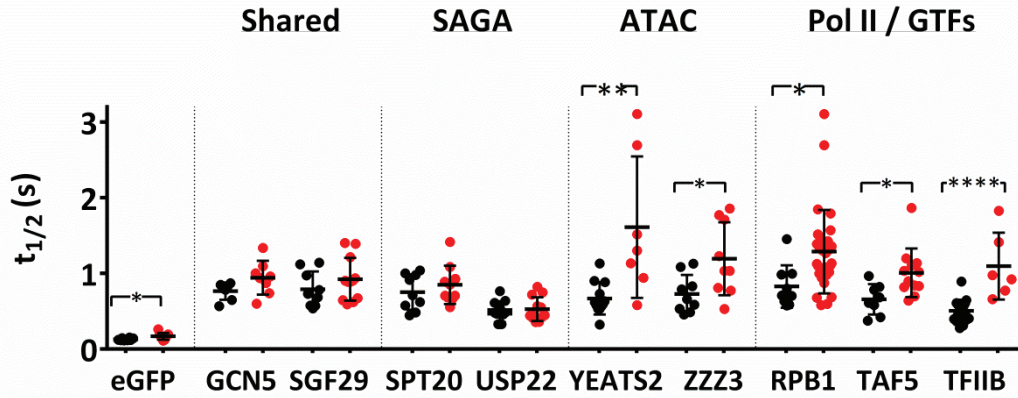
**Figure 3.7. Actinomycin D affects the $t_{1/2}$ of several factors, including that of GFP.** Scatter plot showing the differences between the graphically estimated $t_{1/2}$ in control (black dots) and actinomycin D (red dots) treated cells. Unpaired t-test was performed between groups to estimate the statistical significance of the changes. Significant changes were detected for ATAC specific subunits (YEATS2 and ZZZ3), RPB1, TAF5 and TFIIB. In those cases the increase $t_{1/2}$ indicates a slower recovery rate in the actinomycin D treated cells. Mean values are shown by horizontal lines, bars indicate ±S.D.

The results described above could mean that when transcription is inhibited, SAGA and ATAC exhibit slower mobility due to more stable interactions with chromatin. However, after actinomycin D treatment, a small but detectable reduction in the speed of fluorescence recovery also occurs for eGFP (Figure 3.5 (A), Figure 3.7). We would expect that the drug will not affect the diffusing properties of unconjugated eGFP, since this protein is not expected to have any stable interactions with chromatin, transcription machinery components or actively transcribed loci. A possible explanation for the observed change could be that the intercalation of actinomycin D with the double stranded DNA could change the global structure of the chromatin and thus would affect to a certain extent the mobility of proteins in the nuclear environment. To test whether the differences we observe are a direct effect of Pol II transcription inhibition we decided to carry out a second set of FRAP experiments with another transcription inhibitor.

Thus, next we decided to use DRB (5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside), an inhibitor of RNA polymerase II elongation. This drug is broadly used for transcription inhibition as it specifically inhibits CDK9, the kinase subunit of Positive Transcription Elongation Factor (P-TEFb) (Dubois *et al*, 1994; Nechaev & Adelman, 2011). DRB has no known major effect on chromatin structure, thus any changes of protein mobility could be attributed mostly to a function related with transcription inhibition and particularly elongation. Using this inhibitor we can test to which extent the observed reduction of SAGA and ATAC subunits mobility after actinomycin D treatment was indeed related to transcription inhibition or had to do with changes in nuclear structure.

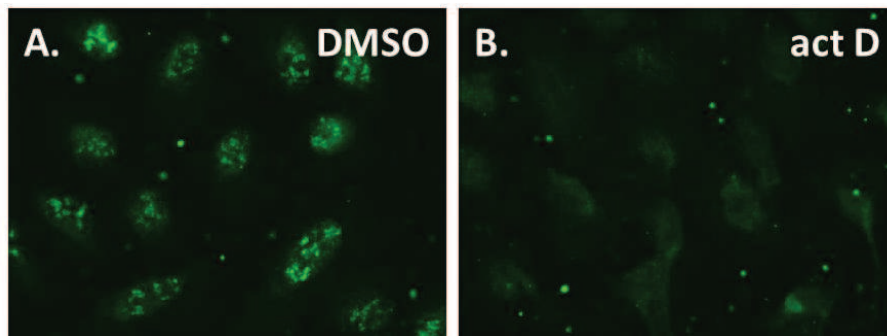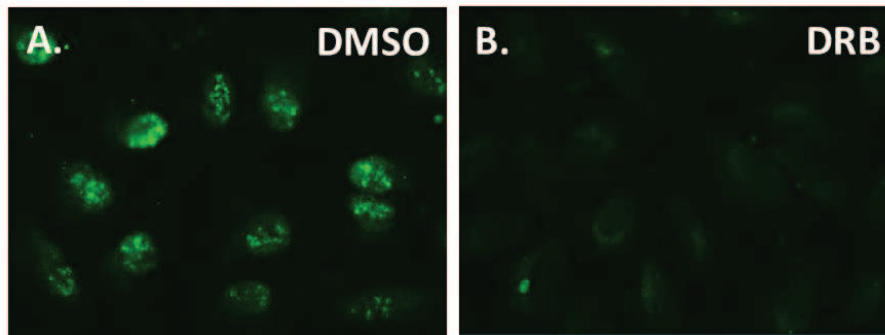**Figure 3.8. DRB treatment efficiently inhibits transcription globally.**
U2OS cells were treated with DMSO (drug vehicle) or 100 μM DRB for 1 h. Like in the case of actinomycin D, assessment of the efficiency of the drug in transcription inhibition was performed by 5-FU incorporation, **A.**: Green signal indicative of incorporated 5-FU and active transcription is detected when cells are treated just with DMSO. B.: DRB treatment results in the disappearance of specific signal. Fixed cells were observed in an epifluorescence microscope (DM4000, Leica) with same exposure time as for actinomycin D (1 sec).

Like for actinomycin D, we first tested and verified the global inhibitory effect of DRB by 5-FU immunodetection in control and drug-treated cells (Figure 3.8). As shown in Figure 3.9, for most tested SAGA and ATAC subunits there is no distinguishable (by FRAP) behaviour change between control cells and those treated with DRB. In contrast to what we observed with actinomycin D, the mobility of eGFP is not affected by the new treatment indicating that DRB is not significantly changing the global structure of chromatin. No effect was observed for TAF5 and TFIIB mobility too. On the contrary, the effect of DRB is clear in the case of RPB1 that was also used as a control for the efficiency of the drug. In agreement with experiments described by (Fromaget & Cook, 2007), we observed increased recovery for Pol II (using a RBP1-GFP read out in our set of experiments) suggesting that during DRB treatment there is a decrease in the Pol II fraction that is engaged in elongation and thus, less associated with the chromatin. We also tested the effect of DRB on TBP and found that it is reducing its recovery rate, however without causing major effect in the mobile fraction (Figure 3.9, Figure 3.10). This observation indicates that TBP is more stably interacting with chromatin when elongation is inhibited. An explanation could be that TBP engaged in PICs formation, is not released due to elongation inhibition and thus the turnover between free and chromatin bound TBP is shifted to the direction of chromatin bound molecules.
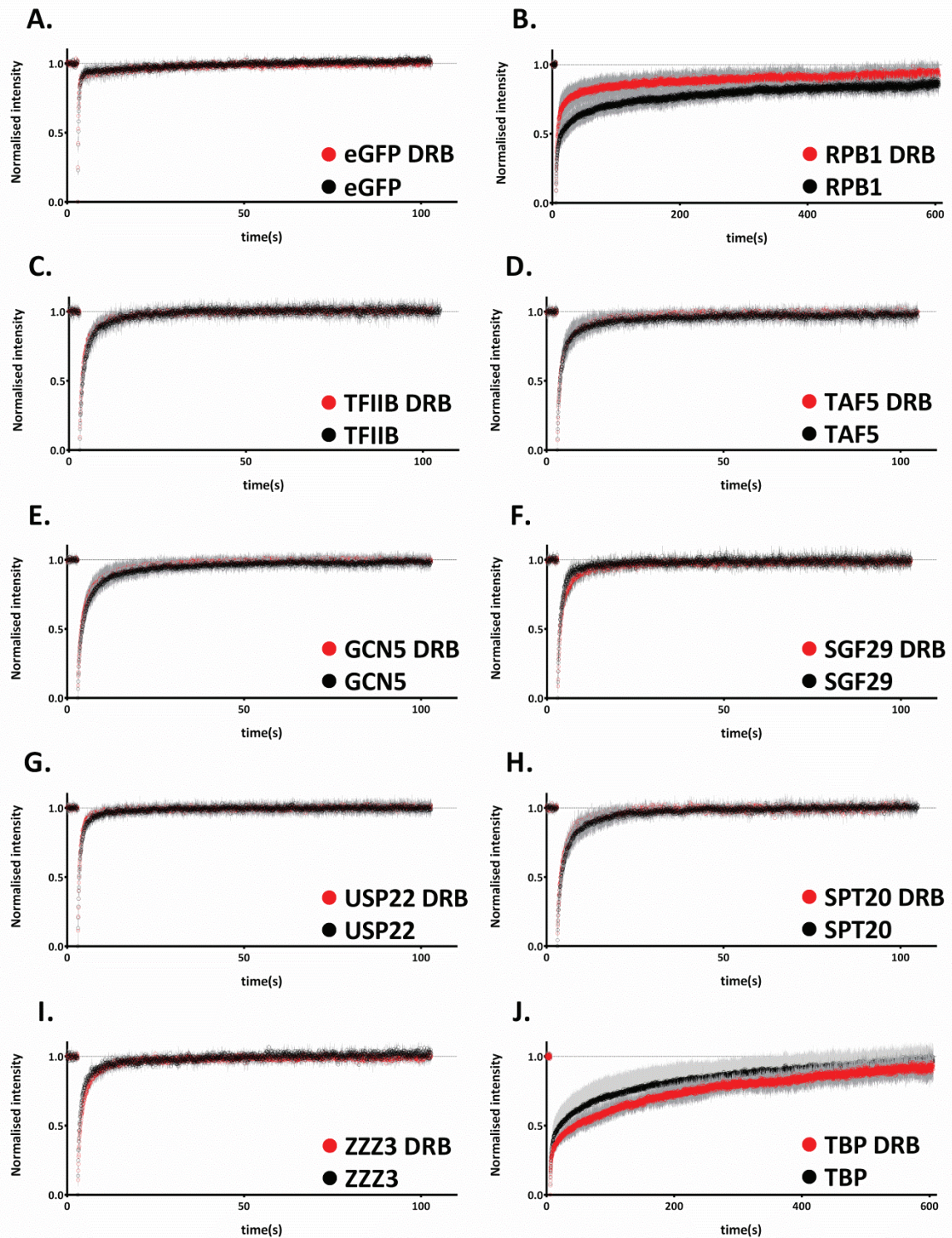
**Figure 3.9. DRB has mild effect on the FRAP curves of most tested factors.**
Qualitative comparison of FRAP curves between control and DRB treated cell shows that for most of the factors tested there is no detectable difference in fluorescence recovery. However DRB results in a faster for RPB1 (B.) and slower for TBP (J.) recovery of fluorescence. Note that there is no detectable effect on the FRAP curve of eGFP.

Among all tested factors, it is only RPB1 for which we could observe a significant change in the mobile fraction (Figure 3.10). Thus, although after actinomycin D treatment we found that the mobile fraction of several factors was affected (reduced in most cases) we did not detect a similar effect caused by DRB.
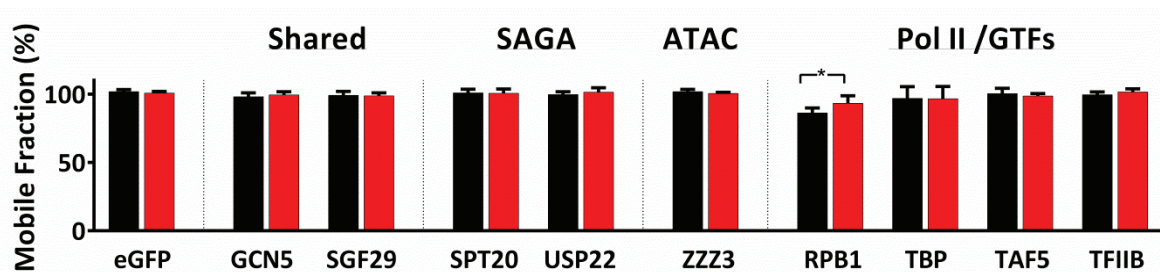


**Figure 3.10. DRB has no significant effect on the mobile fraction of most tested factors.**
Bar chart representing the comparison between the graphically estimated mobile fractions in control (black bars) and DRB (red bars) treated cells. Unpaired t-test was performed between groups to estimate the statistical significance of the changes. No significant changes were detected except for RPB1 that there is a decrease of the immobile fraction.

The comparison of $t_{1/2}$ of the proteins in DRB and control conditions confirmed the absence of significant changes in the mobility of most tested proteins (Figure 3.11). Surprisingly, we found that although no SAGA specific (SPT20, USP22) or ATAC/SAGA (GCN5, SGF29) shared subunit was affected, the ATAC specific subunit (ZZZ3) shows reduced mobility as indicated by the change of its half time of recovery. It has to be noted that the most significant effects of actinomycin D treatment were also observed on ATAC subunits (ZZZ3, YEATS2) used in the previous set of experiments (Figure 3.6 and Figure 3.10). This may be related to a Pol II transcription dependent function of the ATAC complex that would differentiate it from what has been observed for SAGA.
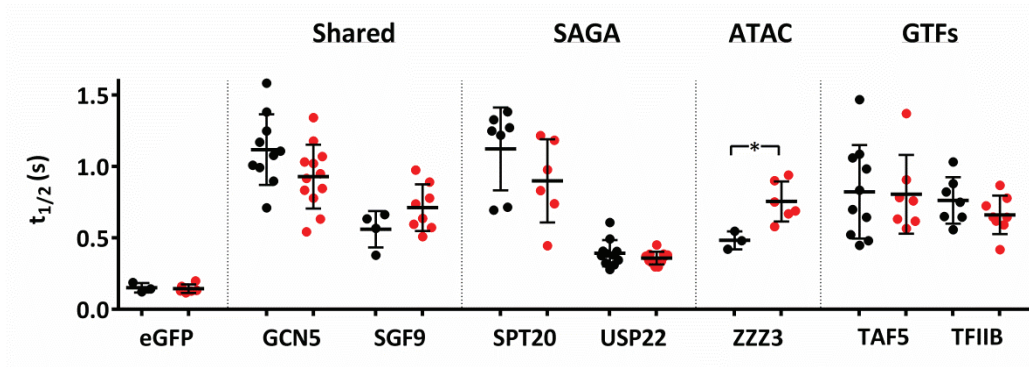
**Figure 3.11. DRB effects on the $t_{1/2}$ of tested SAGA/ATAC subunits and GTFs.**
Scatter plot showing the differences between the graphically estimated $t_{1/2}$ in control (black dots) and DRB (red dots) treated cells. Unpaired t-test was performed between groups to estimate the statistical significance of the changes. No significant changes were detected except for ZZZ3 where an increase of $t_{1/2}$ was detected indicating a slower recovery in the DRB treated condition. Note that GFP $t_{1/2}$ is not affected by DRB in contrast to the significant change observed after actinomycin D treatment. Mean values are shown by horizontal lines, bars indicate ±S.D.

### 3.1.5 Investigation of SAGA and ATAC dynamics by FCS

Our FRAP analysis indicates that SAGA and ATAC complexes are highly dynamic with almost identical fluorescence recovery rates for all subunits. FRAP is a powerful technique to study diffusion properties of a protein, with the technique offering as main advantage the possibility to detect (if present) and quantify the immobile fraction of a protein and investigate potential changes under different conditions. As it can be seen in Figure 3.1, using FRAP we showed that in normal conditions SAGA and ATAC components are highly mobile without a detectable immobile fraction. Moreover, FRAP analysis showed that the average mobility of SAGA specific subunits is not significantly altered under DRB induced transcription inhibition. A potential effect of transcription inhibition of ATAC has to be further investigated as the $t_{1/2}$ of ZZZ3 was found increased after DRB treatment (Figure 3.9 and Figure 3.11).

Despite its variable advantages, FRAP remains an averaging technique which may not detect and resolve changes in the diffusion of multiple components especially in the case of highly mobile proteins (see chapter 1.6.1.1). Since we discovered that SAGA and ATAC belong in a group of very dynamic factors, we wanted to ensure that we do not miss important information about their mobility behaviour in the nucleus. To surpass FRAP limitations in the study of very dynamic proteins, and to verify our observations, we decided to use a technique that is particularly suitable to estimate mobility parameters of fast diffusing proteins. To this end, we established a protocol for FCS (Fluorescence Correlation Spectroscopy) in the nucleus of living cells and performed measurements to obtain further information about the mobility of ATAC and SAGA subunits. As described in chapter 1.6.2, FCS has been used in several studies to estimate the diffusion properties of proteins based on the analysis of fluorescence fluctuations in a femtoliter volume. FCS data analysis results in estimations of several mobility parameters of the examined proteins.

In all experiments described below, our eGFP-tagged proteins of interest were overexpressed in U2OS cells and the selected points for the FCS measurements were always in the nucleus. Nucleoli, nuclear periphery and bright spots in the nucleus of the cells were excluded from our measurements. It must be noted that for reliable FCS measurements, the expression level of the eGFP tagged protein has to be particularly low in order to resolve fluorescence fluctuation at the single molecule level.

The first goal of our FCS experiments was to obtain information about the diffusion properties of SAGA and/or ATAC components in normal conditions and to test whether the highly dynamic nature of these factors, as observed by FRAP, could be confirmed with this technique. To do that,

the autocorrelation data collected during our measurements had to fit to an appropriate model by which the mobility parameters of each protein will be calculated.

One of the most informative mobility parameters of a protein that can be obtained from FCS experiments is its diffusion constant ($D$, expressed in $\mu m^2.s^{-1}$). $D$ is independent of the experimental set up, thus being an absolute parameter to compare the mobility between different proteins. The higher the value of $D$ for a given protein, the faster this protein is diffusing in the observation volume of measurement. In this chapter, most of the results obtained by FCS are described on the basis of this parameter. Since for all tested proteins, our analysis indicates the existence of two distinct diffusing species we will refer to them as *fast* and *slow*, as this reflects the obtained $D_{fast}$ and $D_{slow}$ characterising these species. The *fast* species will have a higher mean $D$ value than the *slow* species (Figure 3.12).

### 3.1.6 Unbiased fitting of FCS data reveals two distinct diffusing species for all studied transcription factors

As mentioned above, to resolve fluorescence fluctuations at the single molecule level, the concentration of fluorescent molecules in the selected volume of FCS measurements has to be very low. Otherwise, above a certain threshold of molecular density, the estimated mobility parameters (e.g. $D_{constant}$) correlate with the number of molecules. We performed several experiments to determine this threshold in our conditions and make sure that the expression level in the selected cells does not affect the estimation of mobility parameters. Eventually, we determined that concentrations above 200 molecules in the femtoliter (fL) volume of measurement should be filtered out from our analysis. As described in chapter 2.10.2, all our results derive from measurements performed in a volume where the total number of molecules (N) is lower than 200.

Our FCS measurement methods are described in chapter 2.10.2. Since the nuclear environment is very heterogeneous, choosing a model with a well-defined number (one, two etc.) of diffusing species cannot be done easily, especially in an unbiased way. Thus, to avoid choosing a model in which we would have to predetermine the number of diffusing components, autocorrelation curves were fitted using a model based on the Maximum Entropy approach (Sengupta *et al*, 2003). As a control, we used the same approach to determine the diffusion constant of monomeric free eGFP diffusing in aqueous solution. In this case, a single peak distribution centered on 96 $\mu m^2.s^{-1}$ was obtained. This value is in good agreement with previous estimations of the diffusion constant eGFP (Potma *et al*, 2001).
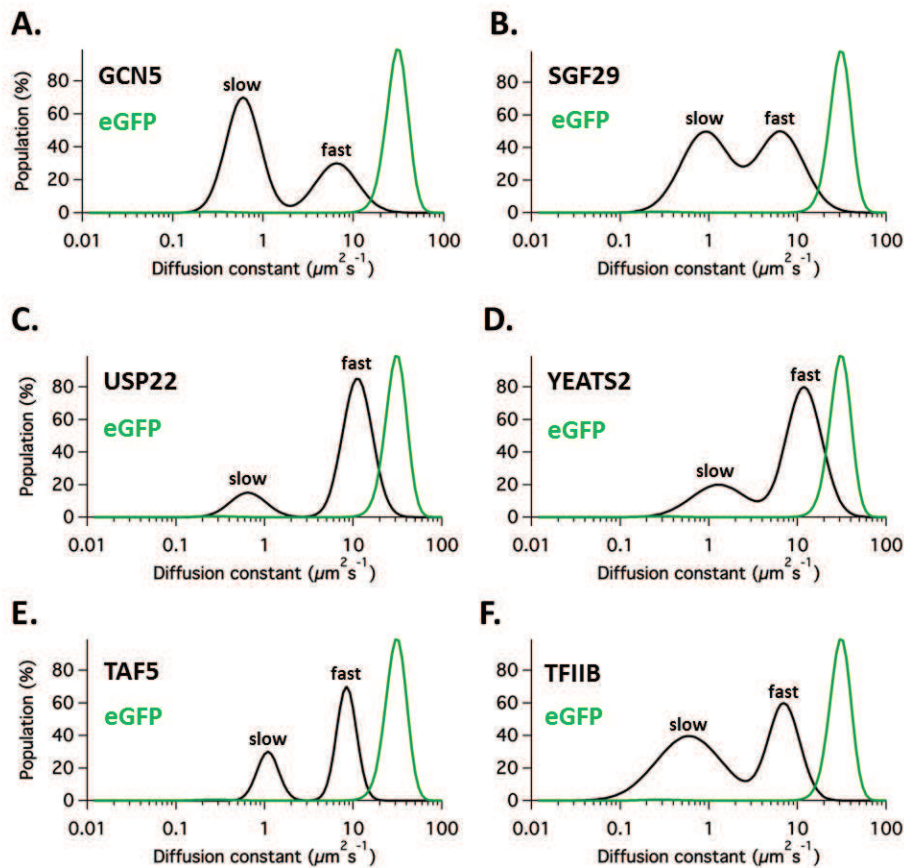
**Figure 3.12. Distribution of diffusion constants of eGFP-tagged human transcription factors**

Fitting of FCS autocorrelation curves of eGFP–tagged proteins with the Maximum Entropy approach results in bi-modal distributions of diffusion constants (*D*s). In every graph the black line corresponds to the distribution of *D*s of the indicated factor. The peaks corresponding to *fast* and *slow* species are indicated. The green line always corresponds to the distribution of *D*s of unconjugated eGFP ($D_{eGFP}$) that forms a single peak and is included in all graphs for direct comparison. **A., B.**: Distribution of *D*s of GCN5 and SGF29, common subunits of SAGA and ATAC. **C.**: *D*s of USP22, SAGA specific subunit. **D.**: Distribution of *D*s of YEATS2 an ATAC specific component. **E., F.**: Distribution of *D*s of two SAGA, ATAC independent factors TAF5 (TFIID component) and TFIIB. Note that in all cases the overlap between the distribution of $D_{eGFP}$ and the distribution of *D*s of any other factor is very small, suggesting that the tested factors are not freely diffusing in our experimental conditions.

Autocorrelation curves of all eGFP-tagged proteins included in this work, when analyzed by Maximum Entropy approach, gave bi-modal distributions (two-peaks) of diffusion constants (Figure 3.12, panels A-F, black line). As expected, free eGFP diffusing within the nucleus gives a well-defined single peak (Figure 3.12, panels A-F, green line) with a higher diffusion constant than every other factor. This type of distribution demonstrates that for our factors of interest the diffusion process is dominated by the presence of two distinct diffusing species.

### 3.1.7 The highly dynamic behaviour of SAGA and ATAC complex is an inherent property of the complexes

Our first goal was to estimate the diffusion constant ($D$) of selected SAGA and ATAC subunits and compare it with that other factors, that were also used in our FRAP experiment described in 3.1.3. This mobility parameter, is characteristic for a diffusing protein in a given environment and provides an objective reference value to compare mobility properties between different factors. As a control of a freely diffusing protein we used eGFP and we determined its mean diffusion constant. $D_{eGFP}$ was found 32 $\mu m^2.s^{-1}$, a value that is in very good agreement with what has been found in previous studies(37 $\mu m^2.s^{-1}$, (Kloster-Landsberg *et al*, 2012), 29.1 $\mu m^2.s^{-1}$, (Bancaud *et al*, 2009)). In addition, the distribution of $D$s from individual measurements forms a single, well-defined peak (Figure 3.12, green line). This shows that our FCS protocol can be used to estimate with high precision the mobility parameters of our proteins of interest, in the nucleus of living cells. An overall comparison of the distribution of diffusion constants between unconjugated eGFP and all the other proteins tested so far is shown in Figure 3.12. As expected, eGFP has the highest $D$ value compared to the $D$ of the *fast* component ($D_{fast}$) of any other protein. More importantly, as illustrated in the graphs of Figure 3.12, the overlap between the distribution of $D_{eGFP}$ and the $D$s of the *fast* species of any other factor is very small. This strongly suggests that, under our experimental conditions, the eGFP-tagged proteins are overexpressed in very low levels and are not freely diffusing but rather incorporated into protein complexes. A comparison of the mean values between $D_{eGFP}$ and $D_{fast}$ of every factor is shown in Figure 3.13, panel A. The graph in the latter figure shows clearly that $D_{fast}$ values of all factors are much smaller than that of $D_{eGFP}$. As expected, a much bigger difference was found between $D_{eGFP}$ and $D_{slow}$ of the tested factors (Figure 3.13, panel B).

Apart from the calculation of diffusion constant per se, another parameter of the diffusing proteins that can be estimated by FCS is their apparent mass. The estimation of this value was particularly important for our study as it provides an approximation of the average size of the diffusing particle (protein complex in our case) in which the eGFP-fused factor is incorporated. The mean apparent molecular mass (MW) of the molecules of each component of the diffusing eGFP–tagged factors can be calculated by comparing their $D_{fast}$ and/or $D_{slow}$ with $D_{eGFP}$ (see chapter 2.10.2). In our case, the estimation of apparent mass for the *fast* component of the tested proteins is in the MDa range (Table 3.1). This is strong evidence that the factors we observe in the femtoliter volume are not diffusing as free proteins but they are part of diffusing objects with much higher mass. As the range of ATAC and SAGA complexes is 1-2 MDa, those bigger objects probably represent the complexes (or subcomplexes) in which the eGFP tagged proteins

incorporate. This is in agreement with the graphs in which the distribution of $Ds$ of the two species is plotted and compared to $D_{eGFP}$ (Figure 3.12). Those graphs show that there is only a minor overlap between the distribution of $D_{eGFP}$ (green line) and the distribution of $D_{fast}$ of any other protein. Regarding $D_{slow}$, it corresponds to a much higher apparent MW that is in the order of $10^3$ MDa (Table 3.1, p.100). This MW is far larger than the size of any of the known complexes that the eGFP tagged proteins has been found to participate. We consider that this fraction corresponds to the molecules that are interacting (more or less transiently) with chromatin or other nuclear structures.
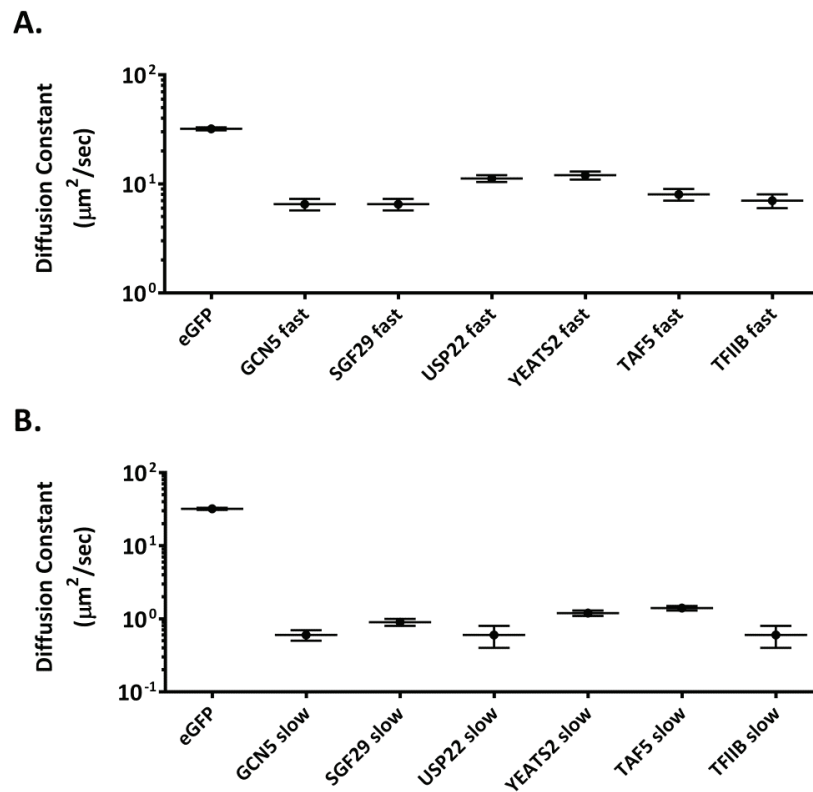


**Figure 3.13. Comparison of $D_{eGFP}$ with $D_{fast}$ and $D_{slow}$ of the eGFP-tagged factors**
Graphs summarising the comparison between the calculated diffusion constant values of the *fast* and *slow* component of each tested factor and the diffusion constant of eGFP. **A.**: $D_{eGFP}$ and $D_{fast}$ of all tested factors are plotted in the same graph. Note that similar $D_{fast}$ values are obtained for the eGFP-tagged proteins and are all lower than $D_{eGFP}$. **B.**: $D_{eGFP}$ and $D_{slow}$ of all tested factors are plotted in the same graph. Note that the y axis scale is different as $D_{slow}$ values for all tested factors are much lower than $D_{eGFP}$. Vertical error bars indicate the relative uncertainty of the fitting parameters.

As shown in Figure 3.13, in graphs A and B, $D_{fast}$ and $D_{slow}$ of SAGA and ATAC components were found similar to that of TAF5 and TFIIB. This result is in good agreement with our FRAP analyses which indicated that ATAC and SAGA are very dynamic complexes. However, in our FRAP experiments we were not able to estimate the mass of the diffusing particles. On the contrary,

with the FCS approach we have calculated the apparent masses of the *fast* species, which are in the range of ATAC and SAGA size. Overall, these observations support the idea that the dynamic behaviour we observe with both techniques is an inherent property of the complexes and does not reflect the behaviour of freely diffusing subunits. Another observation which is also in agreement with our FRAP results, is that we do not find significant differences between the *D*s of common or specific subunits of the complexes (Table 3.1). Nevertheless, more subunits of each complex should be included in the FCS measurements and more measurements should be performed to test if the small differences we observe are significant.

The comparison of $D_{fast}$ and $D_{slow}$ among the studied factors does not reveal any striking differences. These values seem to be very similar both between subunits of ATAC and SAGA, but also between ATAC/SAGA components and other factors, like TFIIB and TAF5 (Figure 3.13), (Table 3.1). On the other hand, differences between factors seem to appear if we observe the distribution of diffusion constants *D*s (Figure 3.12 ) and the estimation of the percentage of *fast* and *slow* ($r_2$) components for each factor (Table 3.1). GCN5 and SGF29, subunits shared between ATAC and SAGA, seem to have the highest fractions of slow component ($r_2$). For GCN5, $r_2$ was found 70% with the same parameter for SGF29 being 50%. Surprisingly, $r_2$ for USP22 (SAGA specific) and YEATS2 (ATAC specific) were more than two times lower, estimated at 15% and 20% respectively. Thus, these two factors seem to have the lowest fractions of *slow* component. For TFIIB and TAF5 the percentage of *slow* component is 40% and 30% respectively. Since the two complex specific subunits (USP22, YEATS2) have relatively low $r_2$ values, it would be expected that the common subunits (GCN5, SGF29) have similarly low $r_2$. This discrepancy may be related to biological reasons that have to be further investigated.

### 3.1.8 Effects of transcription inhibition on SAGA and ATAC subunits analysed by FCS

Next, we wanted to decipher the effect of transcription inhibitors on the behaviour of the complexes in living cells. As described in chapter 3.1.4, FRAP analysis of the effects of two transcription inhibitors indicated that global SAGA and probably ATAC dynamics are not significantly altered after transcription inhibition. Initially, we observed that actinomycin D is reducing the rate of fluorescence recovery of most SAGA and ATAC subunits and in some cases seems to create an immobile fraction. However, as actinomycin D also affects the mobility of unconjugated eGFP we hypothesised that this transcription inhibitor affects the overall structure of chromatin. It is not trivial to distinguish if the effect of actinomycin D on the mobility of a transcription factor is due to transcription inhibition or due to the induced changes in the structure of the nucleus. Thus, we used DRB, for which there is no indication that it directly affects chromatin structure at a global scale. Indeed we found that after treatment with this inhibitor, eGFP mobility is not affected. In addition the mobility of most ATAC and SAGA subunits remained unchanged. These experiments suggested that ATAC and SAGA dynamics are independent of active transcription.

To further investigate the effects of transcription inhibition on the mobility of ATAC and SAGA components, we performed FCS measurements on cells treated with DRB. We wanted to compare the results obtained by the two techniques, but also to exploit FCS as it can potentially detect smaller changes in the mobility parameters of the chosen factors. At present, we have collected FCS data after DRB treatment for GCN5, SGF29 (ATAC and SAGA common subunits) and USP22 (SAGA specific subunit).
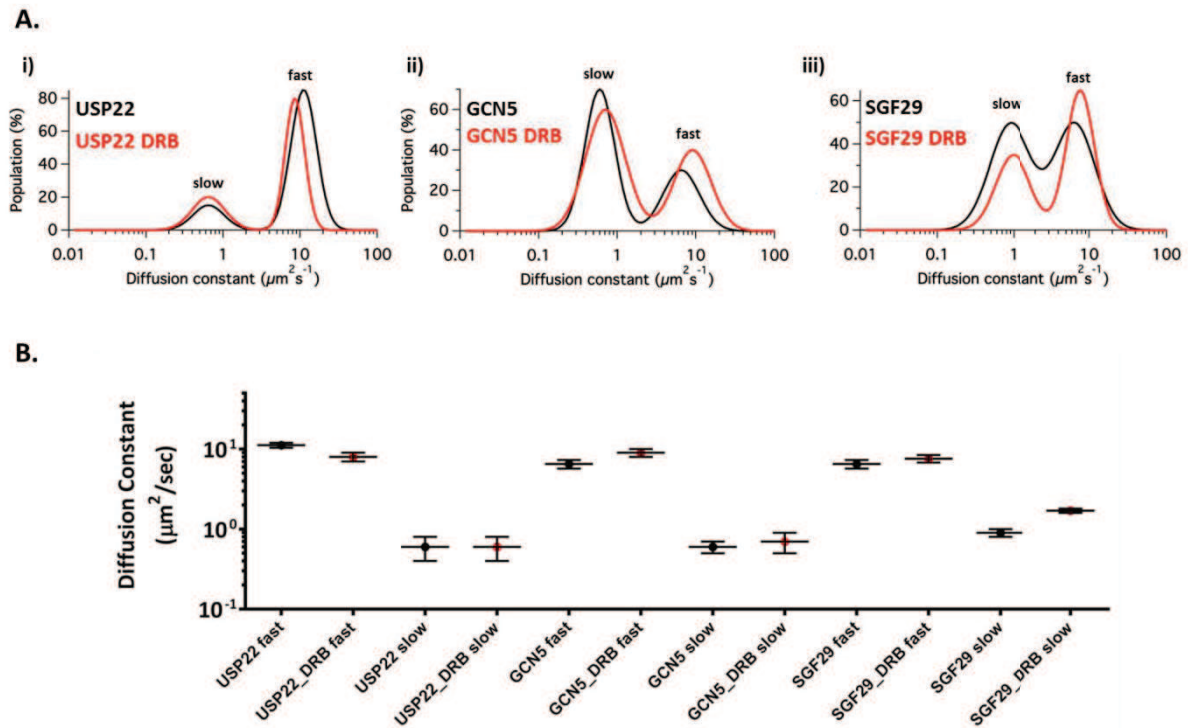
**Figure 3.14. Comparison of $D$s between eGFP and the eGFP-tagged factors.**
Graphs summarising the comparison between the calculated diffusion constant values of the *fast* and *slow* component of each tested factor and the diffusion constant of eGFP. **A.:** $D_{eGFP}$ and $D_{fast}$ of all tested factors are plotted in the same graph. Note that similar $D_{fast}$ values are obtained for the eGFP-tagged proteins and are all lower than $D_{eGFP}$. **B.:** $D_{eGFP}$ and $D_{slow}$ of all tested factors are plotted in the same graph. Note that the y axis scale is different as $D_{slow}$ values for all tested factors are much lower than $D_{eGFP}$. Vertical error bars indicate the relative uncertainty of the fitting parameters.

Like for the FCS data from experiments performed in control conditions, autocorrelation curves collected after DRB treatment were fitted by the Maximun Entropy approach. The distribution of $D$s for USP22, GCN5 and SGF29 in control and DRB treated cells is shown in the graphs of Figure 3.14, A. It is clear that for all three proteins, the distribution of $D$s after DRB treatment still forms two peaks. Graph B in Figure 3.14 shows a comparison of the values of the estimated mean $D$s for the *fast* and *slow* components of USP22, GCN5 and SGF29 in control and transcription inhibition conditions. These values, together with the estimated $r_2$ are summarised in Table 3.1 below.

**Table 3.1. Summary of FCS derived mobility parameters of all eGFP-tagged proteins**

| | Control | | | | | | DRB | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diffusion constant ($\mu m^2.s^{-1}$) | | $r_2$ (%) | App. MW (MDa) | | $n$ | Diffusion constant ($\mu m^2.s^{-1}$) | | $r_2$ (%) | $n$ |
| | *fast* | *slow* | | *fast* | *slow* | | *fast* | *slow* | | |
| **eGFP** | 32±1 | - | - | - | - | 1000 | | | | |
| **GCN5** | 6.5±0.8 | 0.6±0.1 | 70±3 | ~3,2 | ~4,1×10³ | 192 | 9±1 | 0.7±0.2 | 60±6 | 156 |
| **SGF29** | 6.5±0.8 | 0.9±0.1 | 50±4 | ~3,2 | ~1,2×10³ | 209 | 7.6±0.8 | 1.7±0.1 | 35±2 | 202 |
| **USP22** | 11.2±0.8 | 0.6±0.2 | 15±2 | ~0,6 | ~4,1×10³ | 456 | 8±1 | 0.6±0.2 | 20±3 | 176 |
| **YEATS2** | 12±1 | 1.2±0.1 | 20±2 | ~0,5 | ~0,5×10³ | 207 | | | | |
| **TAF5** | 8±1 | 1.4±0.1 | 30±5 | ~1,7 | ~0,3×10³ | 122 | | | | |
| **TFIIB** | 7±1 | 0.6±0.2 | 40±2 | ~2,5 | ~4,1×10³ | 225 | | | | |

Mobility parameters of eGFP–tagged factors derived from the analysis of all the measurements with the Maximum Entropy approach. The error margin (±) refers to the relative uncertainty of the fitting parameters. In each cell, FCS measurements were sequentially repeated, 15 × 5 s in one to three points per nucleus. Thus, $n$ values represent ~15 × the number of total selected points.

Overall there is no dramatic change in the mobility parameters of any of the three tested proteins. We observed only small changes particularly in the $r_2$ of SGF29 and GCN5 which are both increased. To determine if these changes are significant we will have to increase the number of measurements in these conditions and determine whether the observed shifts appear systematically. With the data available so far, we can rather concluded that our preliminary FCS results show that transcription inhibition does not induce significant changes in the mobility parameters of GCN5, SGF29 and USP22.

The FCS findings summarised above confirm the outcome of our FRAP measurements which showed that ATAC and SAGA subunits are highly dynamic, with their mobility being unaffected by transcription inhibition.

## 3.2 DISCUSSION (I)

SAGA and ATAC have been studied mainly with by genetic, biochemical and structural approaches. Numerous studies have investigated gene specific functions of ATAC and SAGA but also genome-wide approaches have been applied to study the distribution of binding sites of the complexes in a global scale. However, the systematic evaluation of the dynamics of the complexes in living cells has not been performed. It has been shown that nuclear, and particularly transcription associated proteins exhibit a broad range of mobility in the nuclei of living cells (van Royen *et al*, 2011). Moreover, it is commonly accepted that the stability of binding interactions of a transcription factor with chromatin is related to its function. To obtain information on the regulation of the function of these complexes from living nuclei of human cells we followed an experimental approach based on live-cell fluorescence microscopy techniques, namely FRAP and FCS.

### 3.2.1 Human ATAC and SAGA function involves transient interactions with chromatin

The first goal of our FRAP experiments was to determine ATAC and SAGA dynamics and compare it with that of other transcription factors. Our FRAP analyses showed that for every tested ATAC and SAGA subunit fluorescence recovery is complete in the order of seconds. Particularly, full recovery for all components is seems to occur after ~20-25 sec (or 0,4 min, to compare with other factors see Figure 1.14). This finding classifies ATAC and SAGA in a category of highly mobile factors. In addition we did not detect an immobile fraction for any of the subunits indicating that the interactions of the two complexes with chromatin are very transient. Moreover, the comparison with the FRAP curve of freely diffusing eGFP clearly shows that the tested subunits are not freely diffusing but exhibit much slower recoveries that we could attribute to interactions with chromatin as components of larger complexes. Thus, we obtained the first indications that the two complexes do not stably associate with chromatin but only transiently interact with their substrates.

We confirmed our initial observations by comparing SAGA and ATAC FRAP behaviour with that of transcription factors that have been investigated in previous studies. In our experiments, we included both highly mobile but also less dynamic factors. Indeed, we showed that FRAP kinetics of SAGA and ATAC components are very similar to TAF5 and TFIIB, two factors that in

previous studies were found highly mobile as well (Chen *et al*, 2002; Sprouse *et al*, 2008; de Graaf *et al*, 2010; Ihalainen *et al*, 2012). Moreover, we compared FRAP curves of ATAC and SAGA components with that of TBP and RPB1, two factors for which it is known that they fully recover in the order of minutes (Kimura *et al*, 2002a; Hieda *et al*, 2005; de Graaf *et al*, 2010). We found that TBP and RPB1 recovery rates were dramatically slower than any ATAC and SAGA subunit. These findings seem to differentiate ATAC and SAGA components from factors like TBP and RPB1, for which a significant immobile fraction and clear biphasic recovery is observed. Such types of recoveries are attributed to two distinct populations of molecules: the first representing fast diffusing molecules transiently interacting with chromatin in a scanning motion and the second one including the fraction of the molecules that are engaged in functions requiring more stable association with chromatin.

Thus fast recovery rates observed in our first set of FRAP experiments showed that, in a global scale, SAGA and ATAC only transiently interact with chromatin.

Having found that ATAC and SAGA components are highly mobile we complemented our FRAP analysis with FCS experiments. This technique is particularly suitable to resolve diffusion properties of highly dynamic components within a small volume of measurement. We took advantage of the non-averaging principle of FCS, to better distinguish between signal and background noise and determine bound and unbound fractions of the eGFP-tagged ATAC and SAGA subunits. For all tested complex components we found bi-modal distributions of diffusion constants corresponding to two distinct diffusing states of each factor. Our FCS results confirmed that the tested ATAC and SAGA subunits have similar diffusion properties with TFIIB and TAF5. Particularly, in our FCS set up, the estimated diffusion constant of TAF5 was found $8\pm1$ $\mu m^2.sec^{-1}$ in good agreement with a value of $7.5\pm0.8$ $\mu m^2.sec^{-1}$ estimated by quantitative FRAP in the study of de Graaf *et al*, (2010).

Moreover, the estimated apparent mass for the fast component of ATAC and SAGA subunits was found in the same range with the size of the complexes. Hence, we have strong evidence that in our experimental conditions the eGFP-fused subunits do not diffuse freely but in the context of the respective complexes. This conclusion is further supported form the clear separation between the distribution of the estimated diffusion constants for eGFP and any other tested protein. In addition the estimated diffusion constant of the slow component (ranging from 0,6 to 1,2 $\mu m^2.s^{-1}$) is higher than what has been observed for factors that interact with high affinity with specific DNA regulatory sites. For example, the diffusion constant of the slow component of eGFP-RAR (retinoic acid receptor) was recently estimated by FCS between 0,05 and 0.10

µm$^2$.sec$^{-1}$ (Brazda *et al*, 2011). This comparison supports a model where SAGA and ATAC only transiently interact with chromatin.

Thus, the combination of FRAP and FCS measurements we carried out, provided strong evidence that ATAC and SAGA are highly dynamic complexes that exhibit only transient interactions with chromatin. This high mobility of the complexes can also explain why ChIP-seq analysis of specific ATAC and SAGA subunits reveals only a limited number of high affinity chromatin binding sites (Krebs *et al*, 2011). From this set of experiments we also obtain strong evidence that SAGA DUB activity along the coding regions of expressed genes cannot be attributed to interactions with Pol II. If active elongation was the driving force of this process an increased immobile fraction of SAGA should have been observed. This conclusion is further supported by the FRAP and FCS experiments we performed after Pol II elongation inhibition.

It should be noted that TBP and RPB1 were not included in this set of FCS measurements because according to our preliminary observations, the imaging parameters (e.g. laser power, acquisition time etc.) we used to measure fast diffusing proteins like ATAC/SAGA subunits, TAF5 and TFIIB cannot be applied on proteins stably interacting with chromatin. To perform FCS with these factors a different FCS protocol has to be established.


### 3.2.2   Dynamics of SAGA and ATAC are not linked to active transcription

To investigate if the mobility of SAGA or ATAC is related to active transcription, we analysed the dynamics of subunits of the two complexes after transcription inhibition with two transcription inhibitors. We showed that actinomycin D alters the mobility properties of eGFP, thus global changes in chromatin may be induced from this drug. In addition, this effect of actinomycin-D in the global structure of chromatin may become more apparent when the diffusion of chromatin interacting proteins is examined. Even more significant effect could be expected if the diffusing proteins are part of complexes with higher mass and multiple chromatin interaction domains. Thus, in certain cases, transcription associated effects of this drug may not be easily distinguished from mobility changes related to chromatin conformation. Since DRB has no known effect on global nuclear architecture, we used this drug to inhibit Pol II elongation. FRAP analysis after transcription inhibition did not reveal significant changes for SAGA components or SAGA/ATAC common subunits. Our FCS analysis in the same conditions does not show significant changes for USP22. Small changes, that have to be further supported with an increased number of measurements, are observed for GCN5 and SGF29. Hence, our live cell experiments provided evidence that the dynamics of SAGA association with chromatin is not

affected by Pol II elongation inhibition. These findings are in agreement with observations made in the recent study of Bonnet *et al*, (2014b). The authors tested if the global DUB activity of SAGA is dependent on Pol II transcription by assessing the efficiency of the removal of the mark in conditions of transcription inhibition (actinomycin D treatment). It was found, that after drug treatment conditions, ubiquitin is removed from H2Bub in a SAGA dependent manner in a few minutes. This finding provided evidence that the interaction of SAGA with its substrate must be very dynamic, since in a very short time its activity seems to dramatically affect the genome-wide abundance of the mark (Bonnet *et al*, 2014b). The highly mobility of the complex in living cells, revealed form our FRAP and FCS experiments, explains how SAGA can exhibit its catalytic activity genome-wide in a very short time. In addition, our live imaging experiments seen to also confirm that the mobility properties of the complex are not dependent on Pol II active transcription. Consequently, the recruitment of SAGA complex is directly linked to its fast diffusion in the nucleus.

Regarding ATAC, the only complex specific subunit (ZZZ3) examined by FRAP under DRB treatment shows an increased half time of recovery indicating reduced association with chromatin. This observation could mean that, in contrast to the SAGA complex transcription-independent mode of action, ATAC mobility (and potentially function) is related to active transcription. Since it has been shown that ATAC can be stably associated with the active form of the Mediator complex (Krebs *et al*, 2010), we could speculate that this kind of interaction creates a more direct link between the dynamics of this complex and transcription. However, the observed effect should definitely be supported both by an increased number of examined cells but also from including more ATAC specific components in our measurements. Finally, expanding the range of FCS tested subunits after DRB treatment could decipher any differences between the two complexes.

 Since DRB is inhibiting Pol II elongation but not the recruitment of the enzyme at the PIC, it could be interesting to test the effect of a third inhibitor that completely impairs the function of polymerase and examine the effects on the dynamics of the two complexes. For this purpose, α-amanitin could be used, as it is known that it leads to RPB1 degradation and Pol II disassembly (Nguyen *et al*, 1996). FRAP and FCS analysis of ATAC and SAGA subunits after treatment with this drug could decipher if there is any link between Pol II recruitment and the dynamics of the complexes.

### 3.2.3 Approaches to clarify the interpretation of FCS data

We showed that the estimated apparent mass of the fast component of ATAC/SAGA subunits is in the range of the size of the complexes thus we assume that the MDa size particles we observe with FCS result from the incorporation of the eGFP subunits in the complexes. However, this can be further investigated with some additional experiments that will link directly the diffusion constant and the estimated apparent mass of this component with the each complex. A possible approach would be to use siRNA mediated downregulation of subunits that are necessary for the integrity of each complex (e.g. siSPT20 for SAGA and siATAC2 for ATAC). In such conditions the integrity of the complexes will be disrupted (Guelman *et al*, 2009; Nagy *et al*, 2009). FCS measurements of an eGFP-fused complex-specific subunit (other than the siRNA target) should result in a distribution of diffusion constants that, in comparison to the control condition, should be shifted towards the direction of free eGFP. Similar experiment should also be performed with a FRAP readout to further support the recovery rates we observe correspond to eGFP tagged subunits incorporated into complexes. In addition, the completion of biochemical experiments testing the efficiency of eGFP-tagged subunits incorporation into the respective complexes is necessary to verify the above conclusions.

### 3.2.4 FRAP experiments to study the effect of SAGA and ATAC on Pol II recruitment in living human cells

An important finding of the study of Bonnet *et al*, (2014b) was that in yeast strains that SAGA complex integrity was impaired, the levels of nascent mRNA were significantly reduced for all tested genes and Pol II occupancy, as indicated by ChIP-seq analysis, was reduced. This result suggests that SAGA is involved in the regulation of RNA pol II recruitment in a genome wide scale, although the mechanism of this interaction has to be investigated. To test, in living human cells, if the global efficiency of Pol II recruitment is also related to SAGA function we could perform a series of FRAP experiments. Particularly, we could perform FRAP on RPB1-eGFP transfected cells in shATX7L3 or shADA3 conditions that have been shown to efficiently affect the DUB and HAT activity of the complex respectively (Bonnet *et al*, 2014b). Next we would compare the dynamics of the protein with FRAP measurements perform in shContol cells. In parallel, we could perform FRAP using the same eGFP-fused Pol II subunit but this time in cells transfected with siRNA against SAGA and/or ATAC subunits that affect the integrity of the complex(es). siRNA against SPT20 could be used to disrupt SAGA and siRNA against ATAC2 or ZZZ3 to disrupt ATAC. These are examples of subunits have been shown to disrupt integrity and

impair recruitment of the respective complexes (Guelman *et al*, 2009; Nagy *et al*, 2009; Nagy *et al*, 2010). Since it has been shown that depletion of this SAGA DUB activity and H3K9ac (SAGA/ATAC HAT activity), affects RNA pol II occupancy in a genome wide level(Bonnet *et al*, 2014b), it is possible that the mobile fraction of RPB1 in such conditions would be increased. This set of experiments could provide further information on the effects of SAGA function in the regulation of Pol II recruitment in a global scale in living human cells.

### 3.2.5 FRAP/FCS experiments to investigate SAGA and ATAC recruitment in living cells

An additional question that could be approached with FRAP and FCS experiments regards the regulation of SAGA and ATAC recruitment in living cells. Several SAGA and ATAC subunits have domains that interact directly with chromatin and recognise specific histone modifications. For example the bromodomain of GCN5 is known to bind acetylated lysines and the tandem tudor domains of SGF29 bind H3K4me3. Other chromatin interacting domains are also found in ADA2a and ADA2b (SANT, ZnF, SWIRM domains), ATXN7 (SCA7 domain). We could design experiments with a FRAP and/or FCS readout that could provide information on how these domains affect SAGA and/or ATAC dynamics in living cells.

For example, chromatin binding properties of the tudor domains of SGF29 are well characterised and have been shown to play important role in SAGA recruitment by anchoring the complex to H3K4me3 and affect the levels of H3 acetylation by the complex (Vermeulen *et al*, 2010; Bian *et al*, 2011; Schram *et al*, 2013). Although SGF29 is a common subunit of SAGA and ATAC, its specific chromatin binding property is exhibited only in the context of SAGA (Vermeulen *et al*, 2010). Thus, it would be interesting to investigate further the role of this subunit in SAGA recruitment in living cells. To this end, we could generate eGFP tagged SGF29 deletion mutants either for one or both tudor domains and test if they exhibit altered nuclear dynamics compared to wild type eGFP tagged SGF29. Alternatively we could mutate conserved residues that have been shown to be particularly important for H3K4me3 binding and test the dynamics of the eGFP-tagged mutants. In addition, to investigate the differential role of this subunit in the regulation of recruitment of SAGA and ATAC we could use siRNA to knock it down and test if we observe different effects in the dynamics of transfected complex-specific subunits. An alternative approach to examine the role of H3K4m3 in SAGA and ATAC recruitment, would be the use siRNA to downregulate the enzymes that deposit this modification. To this end, ASHL2 core subunit of the SET1/MLL/COMPASS-like methyltransferase complexes could be targeted. Since

it has been shown that siRNA ASHL2 dramatically reduces H3K4m3 (Steward *et al*, 2006), dynamics of SAGA and ATAC subunits could be examined in this context. In this way we could observe and study directly the effect of this modification in the mobility of the complexes.

### 3.2.6 Alternative systems to investigate site specific recruitment of ATAC and SAGA in living cells

It has been shown that components of multisubunit, transcription associated complexes may have very different mobility characteristics that can be resolved either in the nucleoplasm or on the promoters that these complexes function. Characteristic examples are the different recovery rates observed between TFIID and Pol I subunits (Dundr *et al*, 2002; de Graaf *et al*, 2010) and Pol II (Sprouse *et al*, 2008) but also from observations regarding the inherent inefficiency of Pol I and II transcription (Darzacq *et al*, 2007; Gorski *et al*, 2008). This type of findings suggest that the association of individual subunits into a functional holocomplex may be a much more dynamic process than it was initially thought and that it can also occur on the sites of action of the complex (e.g. promoters). Thus, there is evidence that during the formation of a functional complex, individual subunits may exchange with very different rates. With our FRAP experiments, we did not detect any significant differences in the recovery rates among subunits of SAGA or among subunits of ATAC. In addition comparison of FRAP rates between ATAC and SAGA specific components showed that subunits of the two complexes behave in a very similar way. Moreover, the FCS experiments we performed do not show major differences among the subunits of the complexes that can be considered significant, at least not with current number of obtained measurements. Thus, with the data available so far, we have not indication that SAGA or ATAC subunits have significantly different mobility characteristics that would suggest multiple exchange rates during their recruitment at their sites of function.

However, in our experimental setup, FRAP measurements are performed in the nucleoplasm of U2OS cells where our complexes of interest are highly mobile exhibiting only very transient interaction with chromatin. Consequently, within this large population of fast diffusing molecules only a very small fraction is expected to participate in promoter specific interactions. Thus, the sensitivity of FRAP is probably not sufficient to detect possible differences in exchange rates that may take place in the very limited number of loci within the examined pool of molecules. On the other hand, although our FCS measurements could detect more subtle diffusion changes between individual subunits, they are performed in randomly chosen

positions within the nucleoplasm that is very unlikely to correspond to genomic loci that SAGA and ATAC are recruited to activate transcription in a gene specific manner.

An alternative approach that could allow us to resolve any differences in the recruitment of ATAC and SAGA subunits would be to quantify their enrichment on an inducible transgene array in real time. Various types of this system have been used in pioneer studies that investigated the regulation of the different steps of gene expression in real time in living cells (Bertrand *et al*, 1998; Tsukamoto *et al*, 2000; Janicki *et al*, 2004; Shav-Tal *et al*, 2004). In contrast to what is observed in the nucleoplasm, where freely diffusing molecules interact with non-amplified loci, the observation of recruitment on the multiple copies of a transgene arrays provides the required signal intensity to detect the recruitment of a factor on an amplified locus during transcription activation in living cells. Such an array was used in a previous study (Rafalska-Metcalf *et al*, 2010), to dissect the temporal recruitment of transcription activation associated proteins after VP16 mediated activation of the array. Interestingly, GCN5 and PCAF were found highly enriched on the activated array locus. Thus, it is highly probable that other SAGA and/or ATAC subunits are recruited on the array in a similar manner. In future experiments we could use the same system to track the recruitment of eGFP-tagged SAGA and ATAC subunits and test if differences in their order of recruitment exist. We could expand our analysis by using eGFP-tagged point or deletion mutants of certain subunits and test how the recruitment is affected. Moreover, taking advantage of the amplified signal we could perform FRAP measurement on subunits already recruited on the array and compare their kinetics.


## CONCLUSIONS

With the combination of FRAP and FCS we demonstrated for the first time directly in living cells, the dynamic nature of human SAGA and ATAC complexes. Our analyses showed that both complexes only transiently interact with chromatin. All tested subunits have similar mobility characteristics indicating that the different rates of association in the holocomplexes either does not occur or could not be detected in our conditions. The dynamic behaviour of the complexes is not affected by transcription inhibition supporting a transcription independent mode of function. The mechanisms regulating the recruitment of the complexes but also the effects of their function on Pol II dynamics can be further investigated with additional live-cell experiments based on the approach followed in the present study.

# 4 Manuscript in preparation II: ''Dynamics of intracellular distribution of ATAC and SAGA components.''

## 4.1 RESULTS (II)

### 4.1.1 Exogenously overexpressed ADA2a and ADA2b localize to different cellular compartments in fixed human U2OS cells

In chapter 3.1, the results obtained by the investigation of ATAC and SAGA dynamics in the nucleus of living cells were described. For the purpose of these experiments several ATAC and SAGA subunits were tagged with eGFP. The localisation of these fusion proteins was examined in fixed U2OS cells 16-24 hrs after transfection of very low quantities of expression vectors (see chapter 2.3). Following exogenous overexpression of higher amounts of plasmids all tested subunits appeared mainly in the nucleus except ADA2a, which above a certain level of overexpression accumulates in the cytoplasm in the tested cells (Figure 4.1, panel A). ADA2a is one of the four subunits composing the HAT core module of ATAC complex together with GCN5, ADA3 and SGF29. The latter three subunits are also present in the SAGA HAT module together with ADA2b (see chapter 1.5.7). In contrary to ADA2a, ADA2b was always found nuclear when exogenously overexpressed (Figure 4.1).
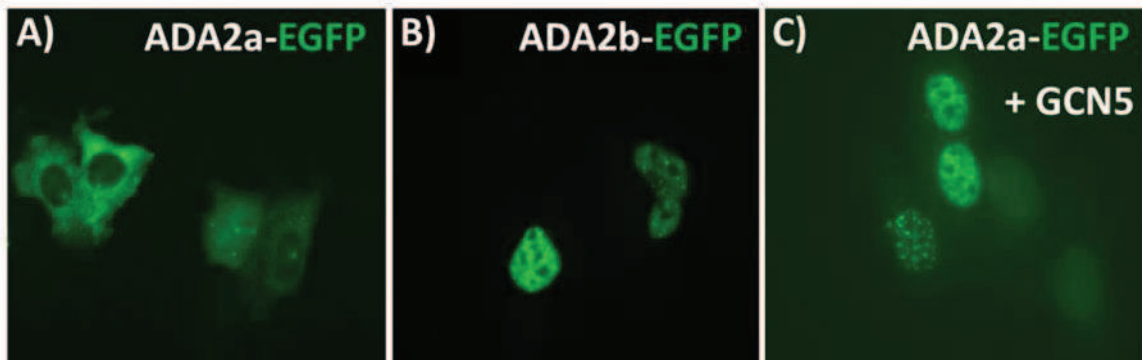
**Figure 4.1. Overexpressed ADA2a and ADA2b have different localisation patterns in U2OS cells.**
ADA2a (ATAC-specific) and ADA2b (SAGA-specific) were cloned in frame with eGFP and overexpressed in U2OS cells that were subsequently fixed and observed by epifluorescence microscopy. **A**: When ADA2a-eGFP is overexpressed alone it appears accumulated in the cytoplasm of most fixed cells. **B**: Overexpressed ADA2b-eGFP appears always nuclear. **C**: The localisation pattern of overexpressed ADA2a-eGFP changes from mainly cytoplasmic to mainly nuclear when GCN5 is simultaneously overexpressed. Fixed cells were observed in an epifluorescence microscope (DM4000, Leica) with same exposure time (1 sec).

Previous work from our lab, has shown that exogenous TAF10, a subunit found in both TFIID and SAGA complexes requires one of its three interaction partners (TAF3, TAF8, SPT7L) to localise in the nucleus (Soutoglou *et al*, 2005). This study indicated that certain components of multisubunit complexes may localise differentially in cells, depending on the simultaneous presence/availability or absence of other subunits of the complex they participate in. We hypothesised that the differential localisation patterns of exogenously overexpressed ADA2a and ADA2b may reflect the different requirements of these two proteins in the type and/or the amount of available factors in a given cellular compartment that affect their intracellular localisation.

Thus, we decided to test if the observed differences in the localisation of overexpressed ADA2a and ADA2b are related to the assembly of the HAT module of the corresponding complexes (possibly in the cytoplasm) and potentially with the regulation of the assembly of ATAC and SAGA holocomplexes. To test this hypothesis, the three other components of the HAT core module (GCN5, ADA3 and SGF29) were individually co-transfected with ADA2a-eGFP. To estimate objectively the changes in the localisation of ADA2a, an automated microscope was used as descried in section 2.10.3. As shown in Figure 4.1, (panel C) and Figure 4.2, (panel A), among the three core HAT components it is only GCN5 that is significantly changing the localisation pattern of overexpressed ADA2a from cytoplasmic to nuclear. Thus, we showed that the amount of neo-synthetized GCN5 present in the cells is directly related to the localisation pattern of overexpressed ADA2a. This requirement may reflect the interaction of the respective

endogenous proteins. In support of our observation, an *in vitro* interaction between ADA2a and GCN5 has already been demonstrated (Gamper *et al*, 2009).
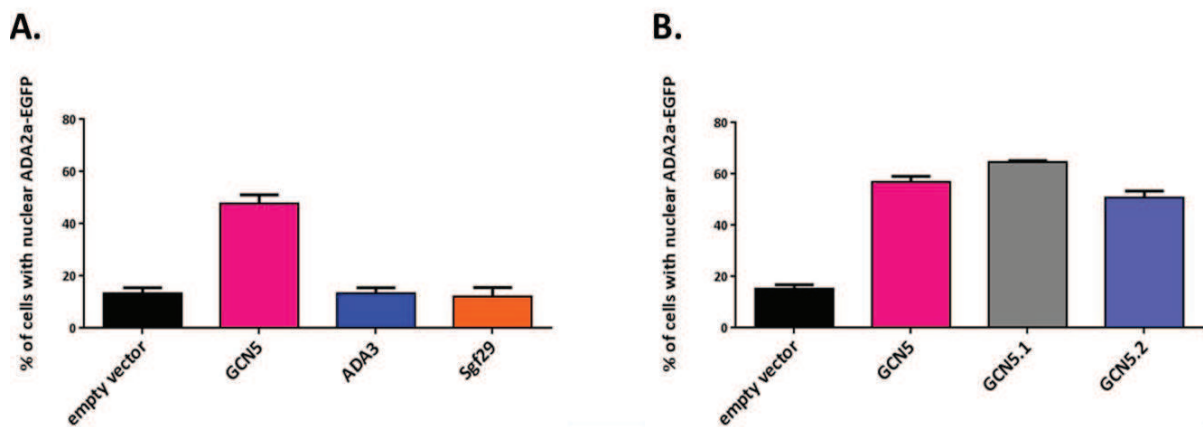


**Figure 4.2. GCN5 abundance but not the enzymatic activity alters ADA2a localisation.**
**A.** Each of the three other subunits of the ATAC HAT core module was co-trasnfected with ADA2a-eGFP. eGFP signal localisation was assessed using an automated microscope following cell segmentation. More than 100 cells were measured per condition. GCN5, the catalytic subunit of the module seems to dramatically affect the localisation pattern of overexpressed ADA2a-eGFP from mainly cytoplasmic to mainly nuclear. **B.** GCN5, a GCN5 single (GCN5.1) and a GCN5 double (GCN5.2) catalytic mutant were co-transfected with ADA2a-eGFP to test whether its localisation depends on the acetyltransferase activity of GCN5. No significant difference is observed between the translocation effect of catalytically active GCN5 and its mutants suggesting that the interaction between the two proteins is probably structural rather than biochemical.

Previous studies have shown that protein acetylation is a post-translational modification which can be involved in the regulation of nuclear import-export of different factors (Gay *et al*, 2003; Thevenet *et al*, 2004; di Bari *et al*, 2006; Liu *et al*, 2012). Thus, we decided to test if the acetyltransferase activity of GCN5 could have a direct role on the regulation ADA2a cellular localisation. To this end, we co-transfected wild type, a single and a double catalytic mutant of GCN5 with ADA2a-eGFP (Bu *et al*, 2007). The localisation pattern of ADA2a-eGFP was assessed with the same automated method as in the previous experiments. As shown in Figure 4.2, (panel B.), there is no significant difference between the effect of wild type GCN5 and its single and double catalytic mutants on the localisation pattern of ADA2a. Thus the shift from a mainly cytoplasmic to a mainly nuclear localisation is independent of the enzymatic activity of GCN5.

This set of experiments showed that ADA2a and ADA2b, the single different subunits between the ATAC and SAGA HAT modules, have distinct intracellular localisation pattern when exogenously overexpressed. Moreover we found that the localisation of ADA2a depends on the available amount, but not on the enzymatic activity, of GCN5.

### 4.1.2 ADA2a and ADA2b have different intracellular dynamics as revealed by live- cell imaging

Our data from fixed cells show that overexpressed ADA2a accumulates in the cytoplasm of U2OS cells and that this localisation pattern can change to nuclear, depending on the presence of exogenous GCN5 in the cells. This means that elevated amounts of available GCN5 are necessary for the nuclear localisation of ADA2a, when overexpressed. Thus, it could be assumed that GCN5 is involved in the import of overexpressed ADA2a into the nucleus. However, the localisation data gathered so far provides information only on the steady state localisation of ADA2a in the different conditions (presence or absence of GCN5). It has been shown that endogenous ADA2a is nuclear (Orpinell *et al*, 2010), which is in agreement with the observation that ADA2a-eGFP cytoplasmic accumulation occurs only above a certain level of overexpression. Thus, the cytoplasmic accumulation pattern of exogenously overexpressed ADA2a that we observe could mean either that the protein cannot be imported into the nucleus due to insufficient amount of import partners or that it is imported but also rapidly exported from the nucleus. To investigate whether overexpressed ADA2a-eGFP is indeed constrained in the cytoplasm (when overexpressed alone) or in the nucleus (when overexpressed together with GCN5) we decided to use FLIP (see chapter 1.6.1.1).

With this live imaging technique the dynamic exchange of a protein between two cellular compartments (e.g. the nucleus and the cytoplasm) can be investigated in living cells. In FLIP, when a cytoplasmic region is repeatedly bleached, mobile molecules localised the cytoplasm should cause a loss of fluorescence in this compartment. Importantly, in the case that the eGFP fused protein is not excluded from the nucleus but shuttles between the two compartments, quantification of average fluorescence intensity of the nucleus will be also show a drop over time. On the contrary, if nucleocytoplasmic shuttling is negligible (e.g. no nuclear export occurs during in the timeframe of the experiment) then the pool of fluorescent nuclear molecules will not be affected. In the reverse FLIP experiment, the selected bleached region can be in the nucleus and average fluorescence intensity quantified in the cytoplasm. This approach is usually applied when the protein is mainly cytoplasmic due to much higher nuclear export rate than the import rate.
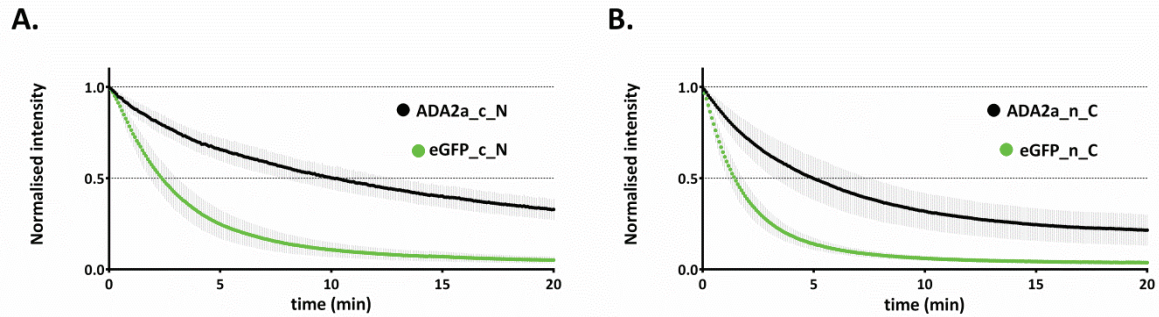
**Figure 4.3. Exogenously overexpressed ADA2a-eGFP is shuttling between the nucleus and the cytoplasm.**
Cells overexpressing either ADA2a-eGFP or free eGFP were subjected to two types of FLIP (nuclear or cytoplasmic) to demonstrate nucleocytoplasmic shuttling of the proteins. **A.**: Graph shows average normalised values of fluorescence intensity over time, in the whole nucleus of ADA2a-eGFP (black line) or eGFP (green line) expressing cells, when a region in the cytoplasm of the cells is repetitively bleached. Fluorescence loss in the nucleus occurs in both cases but is faster for the freely diffusing eGFP. **B.**: Average normalised values of fluorescence intensity in the whole cytoplasm of ADA2a-eGFP or eGFP overexpressing cells, when a region in the nucleus of the cells is repetitively bleached are plotted over time. Similar to (A) fluorescence loss in the cytoplasm was observed for both protein but in a faster rate for eGFP. Black line: **c_N**: bleached ROI in the cytoplasm (c) normalised intensity measured in the nucleus (N). **n_C**: bleached ROI in the nucleus (n) normalised intensity measured in the cytoplasm (C). Error bars show ±S.D. from the mean. n>9 for every tested protein.

Since we had no information about the nucleocytoplasmic shuttling of ADA2a, in our first set of FLIP experiments we applied both approaches described in the previous paragraph in cells overexpressing ADA2a-eGFP at a level that a strong cytoplasmic signal is observed. For FLIP graphs, when the selected bleached region was in the cytoplasm and monitoring of fluorescence was performed in the nucleus, we use the following suffix next to the protein name:'c_N'. On the other hand, when the selected bleached region was in the nucleus and monitoring of fluorescence was performed in the cytoplasm the 'n_C' suffix is used. Raw FLIP data were corrected for observational photobleaching and background signal and normalised to 1, with 1corresponding to the average nuclear fluorescence value of the first post-bleach frame. The total post-bleach time, for which we monitored fluorescence changes, is 20 min for all FLIP experiments. This relatively long duration was necessary and sufficient to obtain reliable information on the intracellular dynamics of the examined proteins.

The result of this first set of FLIP experiments is shown in Figure 4.3. In graph (A), nuclear normalised fluorescence intensity is plotted over time when the bleached region is in the cytoplasm. We performed such measurements not only in ADA2a (black line) expressing cells but also in cells overexpressing eGFP (green line). The decay of fluorescence over time shows that both in the case of ADA2a-eGFP but also for eGFP, nuclear fluorescence intensity is

decreased over time. In graph (B) of the same figure, normalised cytoplasmic fluorescence intensity is plotted over time, with the bleached region in this case being in the nucleus. Average cytoplasmic fluorescence intensity is reduced over time both for overexpressed eGFP but also for ADA2a-eGFP. As expected for a freely diffusing protein, these results show that unconjugated eGFP is shuttling between the nucleus and the cytoplasm. However, as shown in Figure 4.3 the nuclear (graph A) and cytoplasmic (graph B) fluorescence of ADA2a-eGFP expressing cells is also significantly reduced over time when the opposite compartment is repetitively bleached. However, this reduction occurs in a slower rate than for eGFP.

From the qualitative comparison of the FLIP curves of the two proteins, it can be concluded that overexpressed ADA2a is not stably localised in the cytoplasm of live cells. Although a stable cytoplasmic localisation was suggested from the previous set of experiments on fixed cells, FLIP reveals that ADA2a is shuttling between the two compartments. Nevertheless, FLIP curves show that exogenous ADA2a is shuttling with a rate slower than that of eGFP. That could be either due to the bigger size of the ADA2a-eGFP fusion protein, which is slowing down the transfer through the nuclear pores and/or because ADA2a-eGFP may be bound to other proteins as well (i.e. subunits of ATAC complex, import factors in the cytoplasm, or export factors and transient chromatin interactions in the nucleus), which could slow down the shuttling between the two compartments.

Thus, in addition to the above conclusions obtained from fixed cells, these live cell imaging experiments demonstrate that exogenously overexpressed ADA2a is able to enter the nucleus alone, without the presence of exogenous GCN5. Moreover, the above experiments indicate that in spite of the fact that overexpressed ADA2a can enter in the nucleus, it is not retained but rather rapidly exported thus the majority of molecules accumulates in the cytoplasm. This finding is particularly important to understand the interaction between ADA2a and GCN5 in living cells. Furthermore, these experiments suggest that an elevated amount of GCN5 is not related to the import of ADA2a into the nucleus.

Having demonstrated that exogenous ADA2a is shuttling between the nucleus and the cytoplasm, we wanted to test whether ADA2b has a similar behaviour but also to examine what is the effect of the presence of exogenously overexpressed GCN5 in the dynamics of overexpressed ADA2a. Thus, for the series of FLIP experiments described hereafter, the repeatedly bleached region of interest (ROI) was always in the cytoplasm of the cells, while the measurement of normalised fluorescence intensity was always performed in the nucleus. Since the signal in the cytoplasm of cells overexpressing ADA2b-eGFP is low and the protein is mainly

nuclear the reverse FLIP measurements (bleaching in the nucleus, quantification in the cytoplasm) were not performed.

When the selected FLIP protocol (cytoplasmic bleach, nuclear fluorescence monitoring) was applied on cells overexpressing ADA2b-eGFP, only a minor loss of nuclear fluorescence was observed throughout the total time frame of the experiment (Figure 4.4, panel B, blue curve). This result shows that, even when overexpressed, ADA2b is stably localised and retained in the nucleus, and it does not shuttle back to the cytoplasm. As shown by the comparison of the FLIP curves for ADA2a and ADA2b, in the same graph (Figure 4.4, panel B), there is a striking difference in the intracellular dynamics the two proteins.
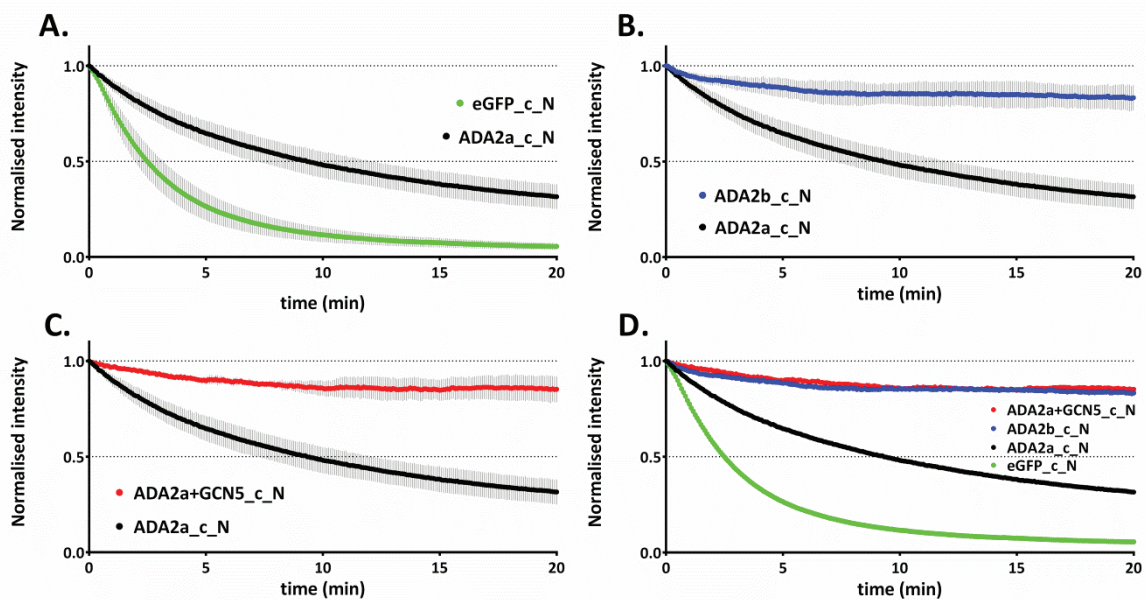


**Figure 4.4. Differential intracellular dynamics of ADA2a and ADA2b revealed by FLIP.**
Cells were subjected to FLIP (bleached ROI in the cytoplasm) and normalised nuclear fluorescence intensity measured in individual cell nuclei over time. **A.**: Graph showing loss of intensity values of nuclear ADA2a-eGFP and eGFP expressing cells. **B.**: A comparison between nuclear ADA2a-eGFP and ADA2b-eGFP FLIP values reveals a striking difference in the behaviour of the two proteins. **C.**: FLIP curves show that when GCN5 is overexpressed together with ADA2a-eGFP the latter is stabilised in the nucleus and does not shuttle to cytoplasm. **D.**: FLIP curves for all four proteins plotted in the same graph. The intracellular dynamics of overexpressed ADA2a in the presence of exogenous GCN5 is similar to that of ADA2b. Green line: eGFP, black line: ADA2a-eGFP, blue line: ADA2b-eGFP, red line: ADA2a-eGFP+GCN5. **c_N**: bleached ROI in the cytoplasm (c) normalised intensity measured in the nucleus (N). n>7 for all tested proteins.

We have demonstrated in fixed cells that overexpressed ADA2a-eGFP localises predominantly in the nucleus when co-transfected with GCN5 (Figure 4.1 panel C, Figure 4.2 panel A). To better understand the intracellular dynamics of ADA2a in the two conditions (presence/absence of exogenous GCN5) in live cells, we also carried out FLIP in cells overexpressing both ADA2a-eGFP and GCN5. The result of this experiment and the comparison with the FLIP curve of ADA2a-eGFP is shown in Figure 4.4, panel C. The red line corresponding to nuclear fluorescence intensity of ADA2a-eGFP in the presence of exogenous GCN5, does not show a significant decrease after cytoplasmic bleaching compared to the extensive loss of fluorescence intensity that is observed for ADA2a-eGFP alone. The FLIP curve of ADA2a-eGFP co-transfected with GCN5 resembles more to the curve of ADA2b (see the graph with all FLIP curves in Figure 4.4, panel D.). This FLIP curve indicates that when the amount of GCN5 is increased by overexpression, exogenous ADA2a is no longer exported back to the cytoplasm, but stays in the nucleus. Together, our fixed and live cell FLIP experiments demonstrate that the presence of an increased amount of GCN5 is necessary to stabilise exogenously expressed ADA2a in the nucleus, but not for importing it the nucleus.

### 4.1.3 Intranuclear mobility of exogenous ADA2a is significantly affected by the co-expression of GCN5

The results of the FLIP experiments described in chapter 4.1.2, showed that GCN5 can stabilise overexpressed ADA2a in the nucleus of U2OS cells. To obtain additional information on the ADA2a-GCN5 interaction we set out to test whether the GCN5-induced change in the shuttling of ADA2a between compartments could also be related to changes in the mobility of the ADA2a-GCN5 complex within the nuclear compartment itself. To answer this question we performed FRAP in the nucleus of cells overexpressing either only ADA2a-eGFP, or ADA2a-eGFP together with GCN5. As detailed in chapter 1.6.2, FRAP kinetics reveal information about the mobility of a protein in the examined region of the cell. Fast FRAP recoveries can be related to highly mobile proteins that interact transiently with their environment, whereas slower FRAP recoveries could reflect more stable interactions of the examined protein with its environment.
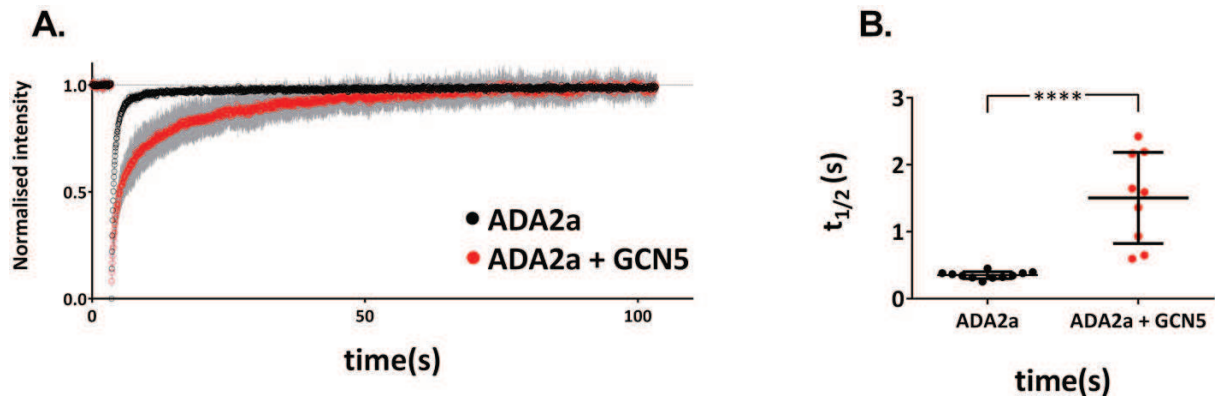
**Figure 4.5. FRAP reveals the effect of GCN5 on nuclear ADA2a mobility.**
**A:** FRAP curves showing a reduction in the mobility of nuclear ADA2a-eGFP when GCN5 is also overexpressed. Mean FRAP values of overexpressed ADA2a-eGFP alone (black circles) and ADA2a-eGFP in the presence of exogenous GCN5 (red circles) plotted over time. **B:** Scatter plot showing the differences between the graphically estimated $t_{1/2}$ of ADA2a-eGFP when only the later protein was overexpressed (black dots) and when exogenous GCN5 was also present (red dots). Unpaired t-test was performed between groups to estimate the statistical significance of the changes (*p*-value< 0,0001).

The comparison of the FRAP curves obtained in nuclei overexpressing ADA2a-eGFP alone or ADA2a-eGFP together with GCN5 is shown in Figure 4.5, panel A. The recoveries of fluorescence suggest that the mobility of ADA2a-eGFP (black curve) is reduced when GCN5 is simultaneously overexpressed (red curve). Particularly, nuclear ADA2a-eGFP alone fully recovers in less than 20 seconds, whereas the addition of GCN5 results in a much slower recovery which takes at least 70 seconds to occur. The half time of fluorescence recovery ($t_{1/2}$) was graphically calculated in both conditions. When ADA2a-eGFP is overexpressed alone it is highly mobile in the nucleus and recovers quickly, with an average $t_{1/2}$ value of 0,35 seconds. A significant difference was observed in the presence of exogenous GCN5, where the $t_{1/2}$ of ADA2a was increased from 0,35 to 1,50 seconds (Figure 4.5, panel B).

FRAP shows that the presence of GCN5 not only affects the intracellular dynamics of ADA2a by retaining it in the nucleus, but it also affects the mobility of ADA2a in the nuclear compartment. An explanation for the significant reduction of mobility of ADA2a that we observe by FRAP could be that ADA2a via its interaction with GCN5 is probably targeted to relatively immobile nuclear structures. It can be assumed that these structures represent chromatin sites that high affinity binding occurs and that GCN5 targets ADA2a to ATAC-specific regulated loci. Such an interaction could attribute the observed reduction of mobility to more stable binding of the complex/sub-complex. However, there is no indication that ATAC has so many high affinity sites on chromatin, the binding on which could cause such a dramatic decrease (at least not in the scale that we observe) in ADA2a mobility when exogenous GCN5 is present. Thus, the most probable

explanation is that ADA2a via its interaction with GCN5 is deposited in the nucleus in sites of unknown identity and function, in the form of non-functional sub-complexes.

The above observations in living cells combined with our fixed cells experiments lead to certain conclusions regarding the effect of GCN5 on the localisation and intracellular dynamics of ADA2a. Although initially we hypothesised that GCN5 is necessary for the import of ADA2a in the nucleus we can now claim that: i) overexpressed ADA2a is able to enter the nucleus without requiring exogenous amounts of GCN5 and ii) that GCN5 is not required for the import of overexpressed ADA2a in the nucleus, but the presence of exogenous GCN5 is necessary to retain ADA2a in the nuclear compartment. The combination of FLIP and FRAP experiments shows clearly that the available amount of GCN5 is directly related to the localisation and nuclear mobility of ADA2a. Moreover, as we observe by FLIP, exogenously overexpressed ADA2b is stably localised in the nucleus without the requirement for additional amounts of GCN5. This striking difference of intracellular dynamics between the two proteins could reflect the differences in the regulatory pathways of ATAC- and SAGA- HAT module assembly.

### 4.1.4 The replacement of the ADA2a SWIRM domain with the corresponding domain of ADA2b retains ADA2a in the nucleus

With a combination of experiments in fixed and live cells we showed that exogenous ADA2a and ADA2b differ not only in their intracellular localisation pattern, but also in their dynamic properties within the cell. In an attempt to identify amino-acid sequences that could play a role in the characteristic localisation of each protein we used publicly available bioinformatics tools to predict nuclear localisation signals (NLSs) and nuclear export signals (NESs). However, no prediction of high confidence NLSs or NESs was found that would justify experimental verification. Thus we decided to focus on larger regions of the proteins, and particularly examine the potential role of the different domains of the two ADA2 proteins in their localisation. Both ADA2a and ADA2b have a very similar domain organization; they contain a zinc finger (Zn), a SANT and a SWIRM domain, with 52%, 57% and 52% homology, respectively (see chapter 1.5.8). Thus we decided to use a domain swapping approach in order to test if individual structural domains of the individual proteins could have an effect on their localisation.
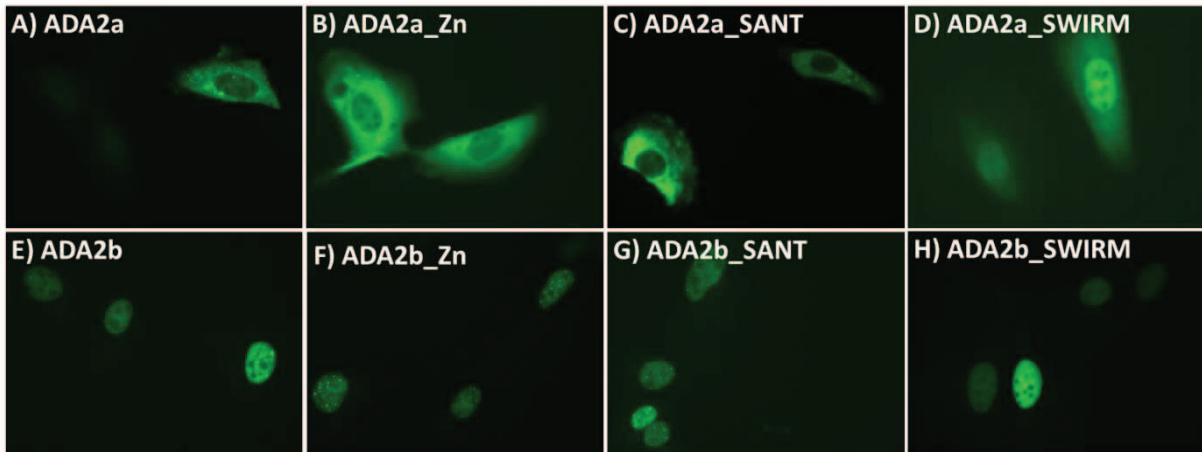
**Figure 4.6. Swapping domains between ADA2a and ADA2b affects the localisation of ADA2a.**
Overexpressed eGFP-tagged ADA2a and ADA2b swap mutants observed by epifluorescence microscopy in fixed U2OS cells. The localisation of the overexpressed proteins in representative cells is shown in this figure. **A.**: Overexpressed ADA2a-eGFP appears accumulated in the cytoplasm of fixed cells. **B, C**: Overexpressed ADA2a_Zn-eGFP and ADA2a_SANT-eGFP mutants accumulate mainly in the cytoplasm. Nuclear localisation of ADA2a_Zn-eGFP was found mildly increased in most cells. **D.**: ADA2a_SWIRM-eGFP mutant appeared in the nucleus of U2OS cells more frequently compared to wt-ADA2a or to the other two swap mutants. **E-H.**: The localisation of all ADA2b swap mutants did not change in comparison to wt-ADA2b indicating that its localisation is not directly related to the swapped regions of the protein. Fixed cells were observed in an epifluorescence microscope (DM4000, Leica) with same exposure time applied for all eGFP fused proteins (1 sec).

To this end mutants of ADA2a and ADA2b were used in which the Zn, SANT and SWIRM domains are swapped between the two proteins. Those chimera coding cDNAs were subcloned in vectors for eukaryotic expression and tagged with eGFP.

Next the nuclear localization of these chimeras was tested by epifluorescence microscopy in fixed cells. First, we transfected the 6 swapped mutants in U2OS cell (see Figure 4.6; when a domain name appears after ADA2a, or ADA2b, it indicates that the labelled domain is originating from the other protein) and compared the localisation of each one with the pattern observed for the respective wild type protein (ADA2a or ADA2b). As shown in Figure 4.6 (panels E-H), all ADA2b domain swapped mutants were nuclear without any detectable difference in their localisation pattern compared to wt-ADA2b. On the other hand, differences were observed between wt-ADA2a and the respective domain-swapped mutants. Particularly, slightly decreased cytoplasmic accumulation was observed for ADA2a_Zn, whereas ADA2a_SWIRM appears clearly more nuclear than wt-ADA2a. The third swapped mutant, ADA2a_SANT, exhibits the same localisation as wt-ADA2a in fixed cells (Figure 4.6, panel C). Since we found that some ADA2a mutants containing ADA2b domains show different cellular localisation compared to wt-

ADA2a, we decided to carry out a second set of FLIP experiments to investigate the effects of the swapped domains on the intracellular localisation of ADA2a.
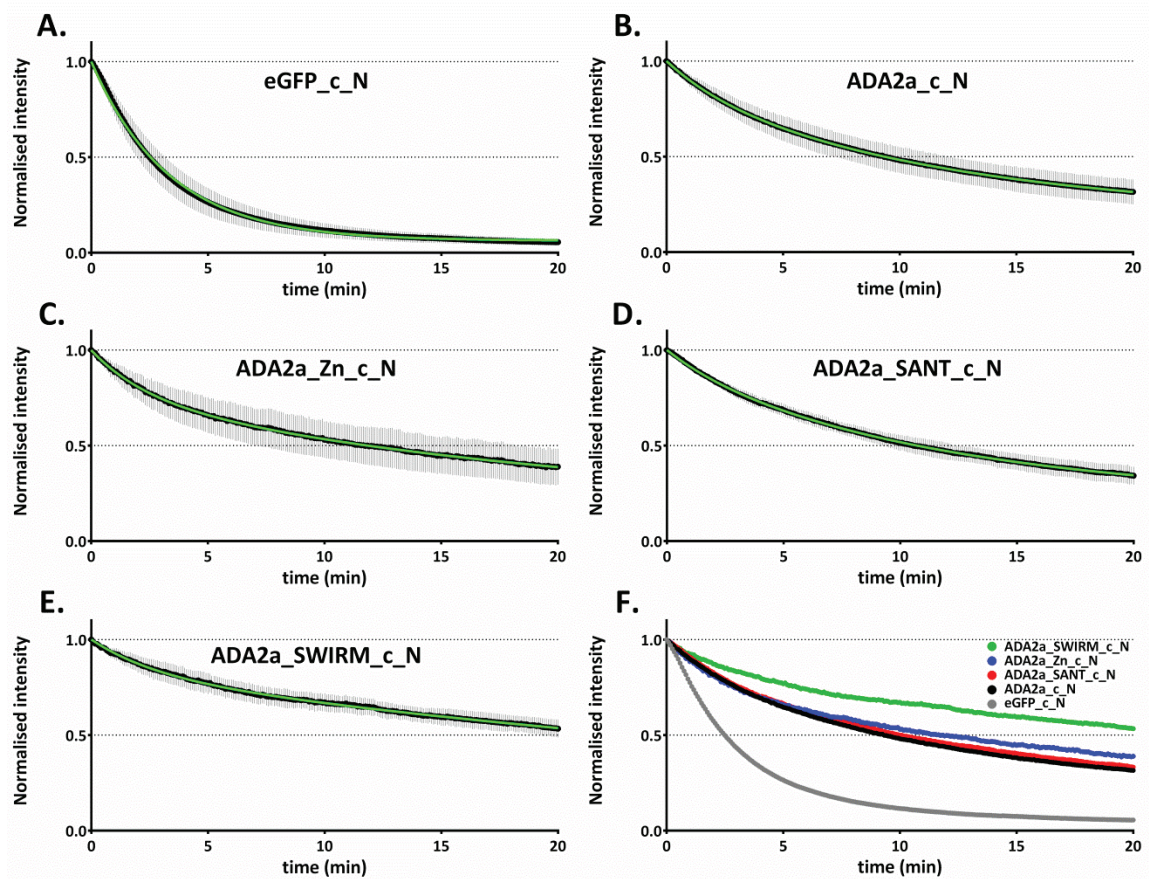


**Figure 4.7. ADA2a SWAP domain mutants have different intracellular dynamics.**
**A-E**: Individual plots of normalised FLIP curves of eGFP, ADA2a (wt), and the three ADA2a SWAP mutants plotted over time (m). Black circles represent the average values and grey bars show ±S.D. within each sample. Green line represents the curve that was fit in each case to calculate the $t_{1/2}$ of fluorescence loss. **F:** For qualitative comparisons between the different proteins an overlay of the average curve of each sample is plotted in the same graph. ADA2a (black curve) and ADA2a_SANT (red curve) have almost identical dynamics. The rate of fluorescence loss is slightly reduced for ADA2a_Zn (blue curve) compared to wt-ADA2a. A big difference is observed between wt-ADA2a and ADA2a_SWIRM (green curve). Note that in at t=20 m the normalised nuclear fluorescence intensity is still above 50% (0,5) of the initial value. n>5 for all tested proteins.

FLIP measurements were obtained using the same protocol as before. For all three ADA2b-eGFP SWAP mutants no detectable change was observed when subjected to FLIP (data not shown). On the contrary, the analysis of FLIP curves of ADA2a-eGFP swap mutants revealed differences both between the mutants but also between the mutants and wt-ADA2a (Figure 4.7).

We found that there was no significant effect on the nucleocytoplasmic shuttling of wt-ADA2a (black curve) and the ADA2a_SANT swap mutant (red curve) Figure 4.7, panel F. This could mean that the SANT domain does not contain any amino acid sequences related to the nuclear localisation of ADA2b or the shuttling of ADA2a between the nucleus and the cytoplasm.

However, when we analysed the FLIP curve of ADA2a_ZN mutant (Figure 4.7, panel F, red curve) a small decrease in the rate of nuclear fluorescence loss in comparison to wt-ADA2a (black curve) was found. The latter observation could mean that the Zn domain of ADA2b may contain sequences that act as a weak NLS or that the sequence of ADA2a Zn finger domain contains a weak nuclear export signal. Alternatively, increased chromatin interaction affinity of ADA2b Zn finger domain could increase the stability of ADA2a_ZN mutant in the nucleus. The fact that we could not identify bioinformatically a canonical NLS or NES export signal in ADA2a or ADA2b sequences makes it less probable that the effects of any swapped mutant is due this NLS or NES type of signals. Since we are comparing proteins amongst which relatively large domains have been exchanged, it is possible that the interaction with factors that play a role in nucleo-cytoplasmic shuttling or interactions with other complex subunits is altered.

Finally, the analysis of ADA2a_SWIRM mutant (Figure 4.7, panel F, green curve) showed that this domain of ADA2b is significantly reducing the rate of loss of nuclear fluorescence intensity when exchanged with the equivalent of ADA2a. As mentioned above the reciprocal ADA2b swapped mutant (ADA2b-SWIRM) had no detectable effect in ADA2b localisation dynamics. It is possible that the SWIRM domain contains sequences can that act as NLS in the context of ADA2a but are not essential for the import of ADA2b in the nucleus. Moreover, like in the case of ADA2a_Zn mutant, increased chromatin interactions mediated by this domain may also be contributing to changes of the nucleocytoplasmic shuttling rate of ADA2a_SWIRM mutant.
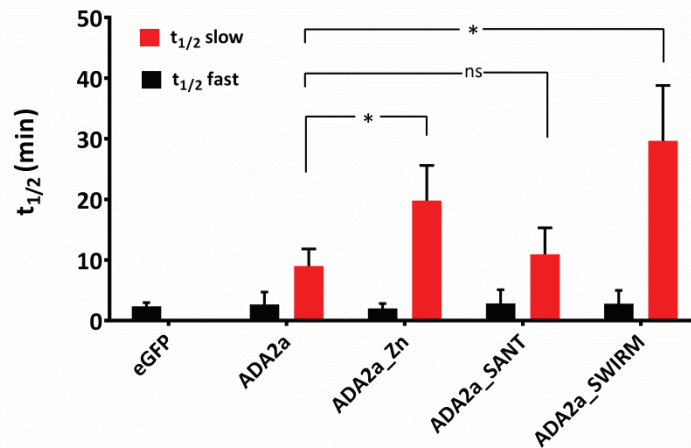
**Figure 4.8. ADA2s SWAP mutants have different half time of loss than ADA2a.**
Bar chart showing the half times of loss estimated from the fit of the averaged FLIP curves of eGFP, wt ADA2a and the three ADA2a domain swapped mutants. The half time of loss of eGFP derived from one phase exponential fit. For every other protein a two phase exponential fit was used to estimate the $t_{1/2}$ of loss of the slow and fast components are plotted on the graph. All proteins have very similar $t_{1/2}$ of loss of the fast component. Significant differences for the $t_{1/2}$ of loss of the slow component were found between wt-ADA2a and ADA2a_Zn but the most striking increase of $t_{1/2}$ of loss was observed in the ADA2a_SWIRM mutant.

In order to obtain information about the effects of ADA2b domains in the dynamics of ADA2a translocation, the average FLIP curves were fitted to a single and double exponential decay equation. Our preliminary analysis indicates that for ADA2a and ADA2b domain swapped mutants, the curve fitting was only possible with a double exponential decay equation. On the contrary eGFP could only be fitted with a single exponential decay equation. This shows that although for eGFP there is only one population of diffusing molecules, there are two populations of ADA2a molecules shuttling between the nucleus and the cytoplasm. The half times of loss obtained from the fits are illustrated in Figure 4.8. The estimated $t_{1/2}$s of loss are in accordance with the conclusions made by the qualitative comparison of the FLIP curves in Figure 4.7. The faster among the two ADA2a populations, shows similar half time of loss with that of eGFP, indicating that this pool of ADA2a molecules is probably freely diffusing between the compartments. It must be noted that the effect of the mutants (ADA2a_Zn and ADA2a_SWIRM) only affects the $t_{1/2}$ of loss of the slow component. The latter component could represent ADA2a molecules that are incorporated into holocomplexes or subcomplexes that are either no longer exported from the nucleus or are exported but with a reduced export rate.

The investigation of intracellular mobility of ADA2a and ADA2b swapped mutants showed that ADA2b localisation is not exclusively dependent on any of its three domains. On the other hand, it seems that the presence of ADA2b domains in ADA2a sequence can affect the

nucleocytoplasmic shuttling of the latter protein. However, the potential link between the observed differences in intracellular dynamics of the two proteins and the regulation of assembly of the respective complexes has to be further investigated.

### 4.1.5 The cytoplasmic and nuclear assembly pathways of endogenous ATAC and SAGA complexes are different

The analysis of exogenously expressed ADA2a and ADA2b proteins by different imaging techniques revealed distinct features in the intracellular dynamics of the two proteins. We also demonstrated that the presence of exogenous GCN5 is required for the retention of overexpressed ADA2a in the nucleus and is enhancing the association of ADA2a with nuclear structures. Thus, we hypothesised that the abundance of GCN5 could regulate the formation of ATAC HAT module and consequently ATAC holocomplex formation. In addition, the fact that exogenously expressed ADA2b always appears nuclear may suggest that the assembly of the SAGA HAT module, or SAGA holocomplex, is regulated in a different way.
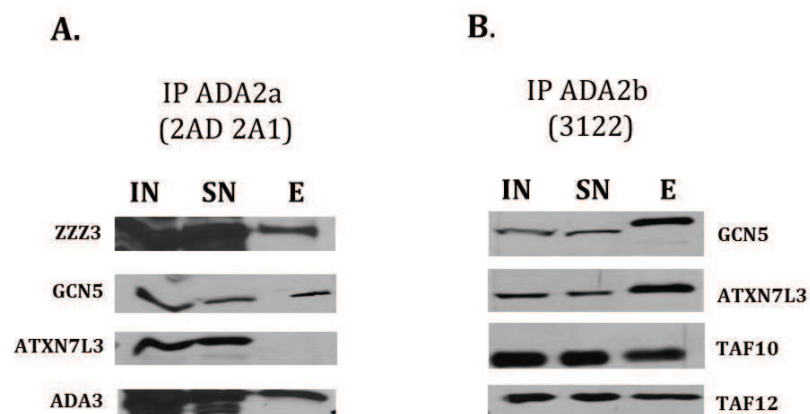


**Figure 4.9. anti-ADA2a and –ADA2b antibodies efficiently IP ATAC and SAGA complexes respectively.**
Proteins from HeLa NEs were immunoprecipitated either with the anti-ADA2a (2AD2A1) mouse monoclonal antibody or with the anti-ADA2b (3122) rabbit polyclonal antibody and eluted by glycine. Protein separation was performed by SDS-polyacrylamide gel electrophoresis and analyzed by Western blotting (WB) using antibodies against the indicated subunits. **A.**: anti-ADA2a antibody immunoprecipitates ATAC specific subunit ZZZ3 together with GCN5 and ADA3 (common subunits between ATAC and SAGA). ATXN7L3 (SAGA specific), although present in the input (IN) and supernatant (SN) of the IP, is not enriched in the eluate (E), indicating the specificity of the antibody. **B.**: SAGA specific subunits ATXN7L3, TAF10 and TAF12 but also GCN5 (shared component of the two complexes) are detected in IP performed with the anti-ADA2b antibody. 5µl of input and supernatant (IP flow-through) and 15µl of the eluted proteins were loaded for the WB analysis of both IPs.

Nevertheless, the information obtained so far was exclusively based on the behaviour of overexpressed proteins. Therefore, we set out to investigate if SAGA and ATAC subunits interact with endogenous ADA2a or ADA2b in the cytoplasm and the nucleus. These endogenous interactions could provide insights into the different assembly mechanisms of SAGA/ATAC HAT modules or holocomplexes. To this end, we decided to follow an approach based on quantitative

proteomics. First, we have generated and/or tested available antibodies that would work efficiently when used in immunoprecipitation experiments showing that the purified anti-ADA2a monoclonal mouse antibody and the anti-ADA2b polyclonal rabbit antibody immunoprecipitate (IP) the corresponding ATAC and SAGA complexes from human HeLa cell extracts, as detected by western blot analyses (Figure 4.9).

Next, we prepared nuclear (NE) and cytoplasmic (CE) protein extracts from HeLa cells. HeLa cells can be grown in cell suspension cultures allowing us to produce highly concentrated protein nuclear and cytoplasmic extracts from a large number of cells. This is important to make sure that both the nuclear and the cytoplasmic fractions have sufficient amounts of total protein and can be used for the following steps but also that the cross contamination between the extracts is minimised (Figure 4.10). Since SAGA and ATAC are mainly nuclear, what is really important for our analysis is that during the fractionation the nuclear proteins do not contaminate the CE. Note that, it is well known that although γ-tubulin is mainly cytoplasmic, significant amounts also exist in the nucleus (Lesca *et al*, 2005; Höög *et al*, 2011; Eklund *et al*, 2014). Thus, the pattern observed in Figure 4.10 corresponds to very efficient fractionation.
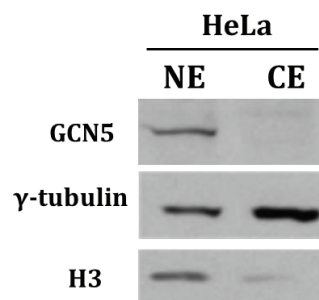


**Figure 4.10. Purity of nuclear and cytoplasmic extracts.**
Equal amounts of protein (8 µg) from NE and CE were separated by SDS-polyacrylamide gel electrophoresis and analysed by Western blotting. The indicated antibodies were used to estimate the abundance of different factors in the nuclear (NE) and cytoplasmic (CE) extracts generated from HeLa cells. As expected, GNC5 and H3 are highly enriched in the NE and almost absent in the CE. On the contrary γ-tubulin is enriched in the cytoplasmic fraction.

Subsequently, the above-tested specific antibodies recognising endogenous ADA2a or ADA2b were used to immunoprecipitate the corresponding endogenous proteins from the cytoplasmic and the nuclear extracts. To identify proteins associated with ADA2a or ADA2b in each compartment, the immunoprecipitated proteins were eluted from the immunoaffinity columns, digested with trypsin and analysed by shotgun proteomics using Multidimensional Protein Identification Technology (MudPIT) (see chapter 2.9) (Washburn *et al*, 2001).

**Figure 4.11. Endogenous ADA2a associates with ATAC components only in the nucleus.**
NSAF quantification of SAGA and ATAC subunits in IP samples against endogenous ADA2a (ATAC specific) and ADA2b (SAGA specific). The two proteins were immunoprecipitated from cytoplasmic and nuclear extracts prepared from HeLa cells. **A**: Quantification of ATAC subunits present in anti-ADA2a IP. ATAC subunits are enriched only in the nucleus but not in the cytoplasm of HeLa cells. **B**: Quantification of SAGA subunits present in anti-ADA2b IP. SAGA subunits are significantly enriched both in the nucleus and the cytoplasm. Orange bars correspond to the IP from the nuclear extract and grey bars to the IP from the cytoplasmic extract.

The abundance of each SAGA and ATAC subunit in the MudPIT datasets was calculated by Normalized Spectral Abundance Factor (NSAF) value (Zybailov *et al*, 2006). The NSAF values were calculated from the spectral counts of each identified protein. To account for the fact that larger proteins tend to contribute more peptides/spectra, spectral counts were divided by protein length to provide a spectral abundance factor (SAF). SAF values were then normalized against the sum of all SAF values in the corresponding run, allowing the comparison of protein levels across different runs. With this method the abundance of proteins between cytoplasmic and nuclear IPs could be compared quantitatively. Importantly, in the anti-ADA2a and anti-ADA2b IPs each known ATAC (for ADA2a IP) and SAGA (for ADA2b IP) subunit was identified in

HeLa cell NE extracts, respectively, with high confidence, validating our approach (see yellow bars in Figure 4.11 panels A and B). Next we compared the presence of ATAC subunits in the nuclear and the cytoplasmic anti-ADA2a immunoprecipitations (IPs). As it is shown in Figure 4.11 (panel A), ADA2a could be immunoprecipitated only from the NE, where we found ADA2a associated with all known ATAC subunits. On the contrary, ADA2a was not significantly enriched from cytoplasmic anti-ADA2a IP. Consequently, no other ATAC subunits could be identified in this IP. Next we analysed the nuclear and cytoplasmic anti-ADA2b IPs (Figure 4.11, panel B), and found that ADA2b associates with all SAGA subunits in the nucleus, but also associates with almost all SAGA subunits in the cytoplasmic extracts (Figure 4.11, panel B, compare yellow and grey columns). Thus, the results of the analysis of the cytoplasmic anti-ADA2b IP revealed a significant difference compared to what we found for ADA2a interactions in the same compartment. ADA2b can clearly be IP-ed together with almost all SAGA subunits from the cytoplasm, while ADA2a cannot. This difference cannot be explained by the different efficiencies of the used antibodies as they both worked perfectly well in IP-ing the corresponding complex from the NEs. This first set of analysed IPs may suggest that there is a difference in the compartment in which SAGA and ATAC assembly takes place. It seems that ADA2b is abundant in the cytoplasm, can be IP-ed, and SAGA subunits already interact to form the holocomplex in the cytoplasm. Thus, SAGA may be imported to nucleus as a fully assembled complex. In contrast, neo synthetized endogenous ADA2a is not enough abundant in the cytoplasm, and ADA2a-containing ATAC can only be IP-ed from the nucleus. At this point, it has to be noted that IPs against ADA2a and ADA2b were performed on the same batch of cytoplasmic and nuclear extracts, thus it is unlikely that the observed distribution of subunits could be attributed to inconsistencies in the preparation of extracts between experiments.
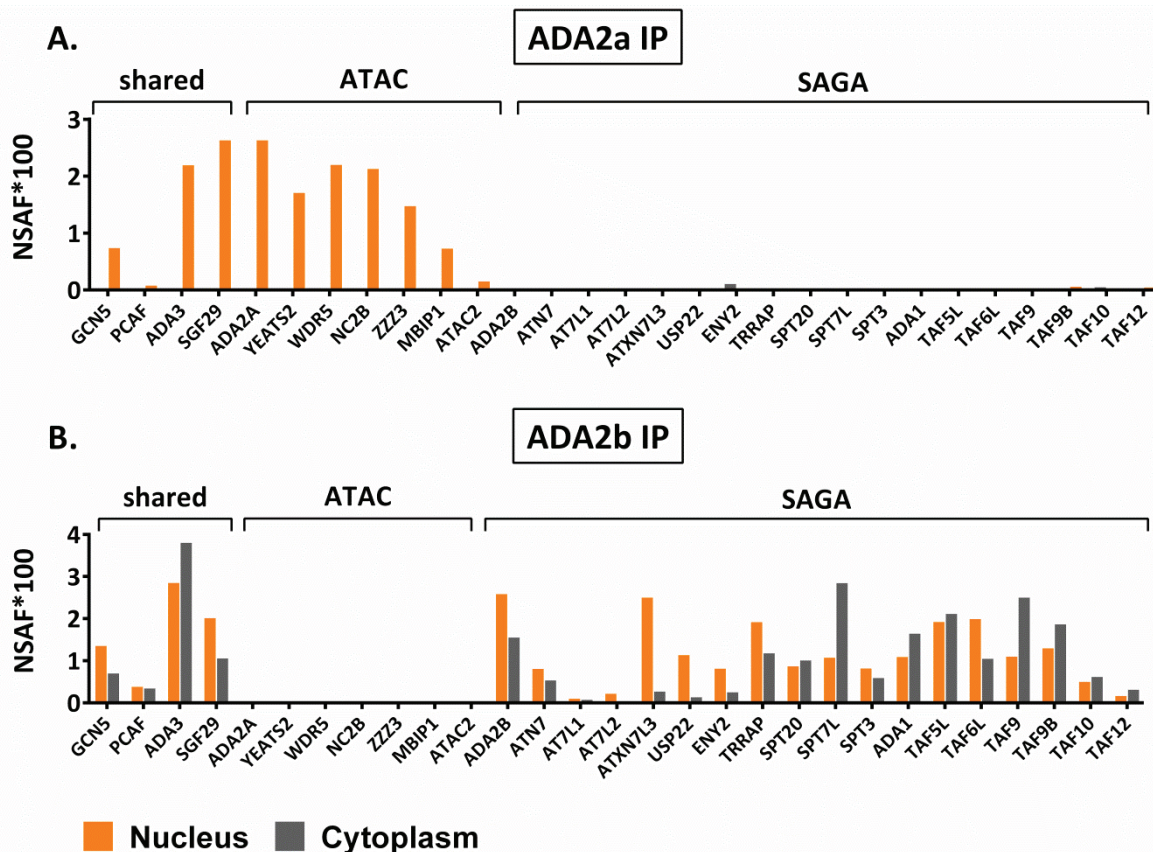
**Figure 4.12. ZZZ3 interacts with ATAC components only the nucleus.**
NSAF quantification of SAGA and ATAC subunits in IP samples against endogenous ZZZ3 (ATAC specific) and SPT20 (SAGA specific). The two proteins were immunoprecipitated from cytoplasmic and nuclear extracts prepared from HeLa cells. **A**: Quantification of ATAC subunits present in anti-ZZZ3 IP. ATAC subunits are enriched only in the nucleus but not in the cytoplasm of HeLa cells. **B**: Quantification of SAGA subunits present in anti-SPT20 IP. SAGA subunits are significantly enriched both in the nucleus and the cytoplasm. Orange bars correspond to the IP from the nuclear extract and grey bars to the IP from the cytoplasmic extract.

To further test the above-described hypothesis about the differential assembly of the Ada2a- and ADA2b-containing endogenous complexes, we decided to perform additional IPs against another SAGA- or ATAC-specific subunit and analyse them by MudPIT and NSAF, as described above. We selected SPT20 and ZZZ3 that have been shown to be specific SAGA or ATAC components, respectively, and against which we had previously generated IP grade antibodies (Nagy *et al*, 2010; Krebs *et al*, 2011). As shown in Figure 4.12 (panel A), the results of the analysis of the IPs against the ATAC-specific subunit ZZZ3 are similar to what we found in the case of anti-ADA2a IPs. ZZZ3 was found associated with all ATAC subunits in the nucleus, whereas no significant identification of ZZZ3 or other ATAC subunit was possible in the IPs performed from the cytoplasmic extract. Moreover, the results we obtained for the SPT20 IPs seem to be in agreement with what we observed for ADA2b. Indeed, SPT20 associates with all SAGA subunits in the nucleus but also interacts with several SAGA subunits in the cytoplasm. These results

verify our initial observations made from the quantitative analysis of the anti-ADA2a and anti-ADA2b IPs

To further support our findings we performed a third set of IPs from nuclear and cytoplasmic extracts, this time targeting a subunit that is shared between the two complexes. We chose ADA3, a common subunit of ATAC and SAGA present in the HAT modules of the complexes.
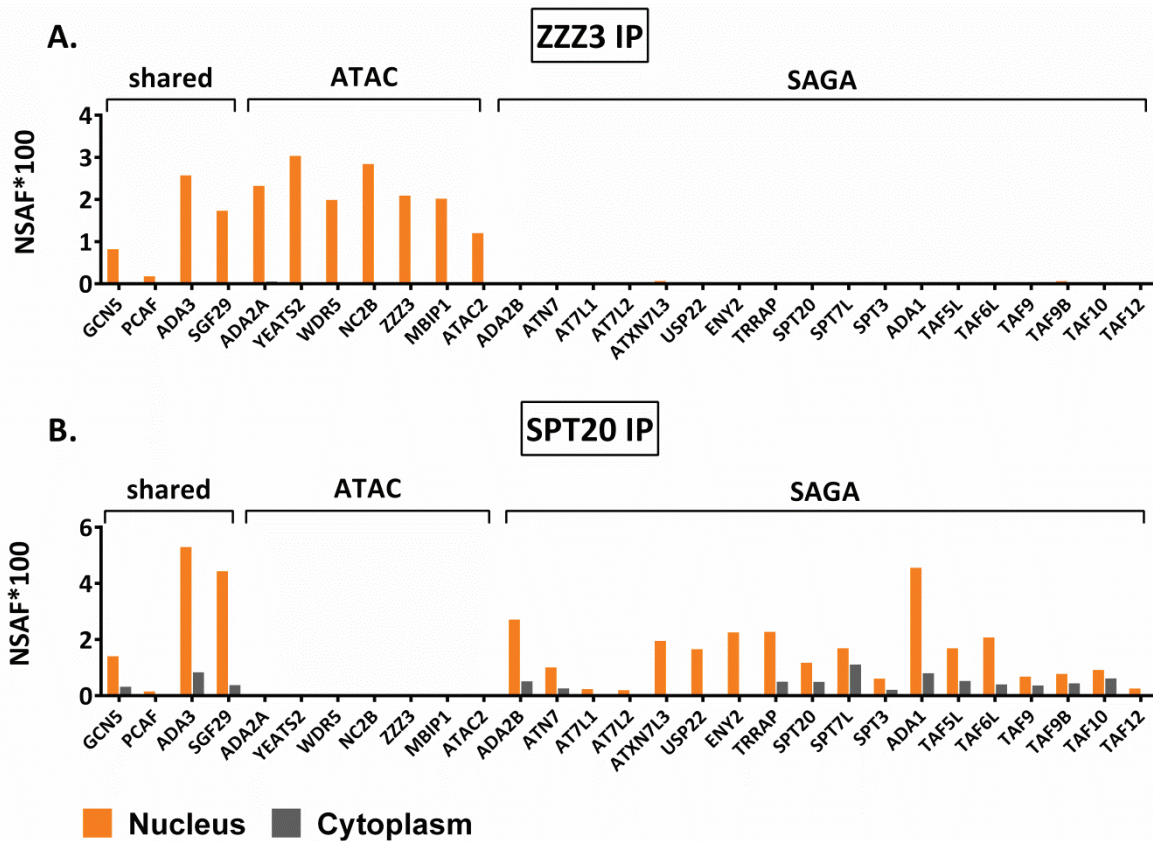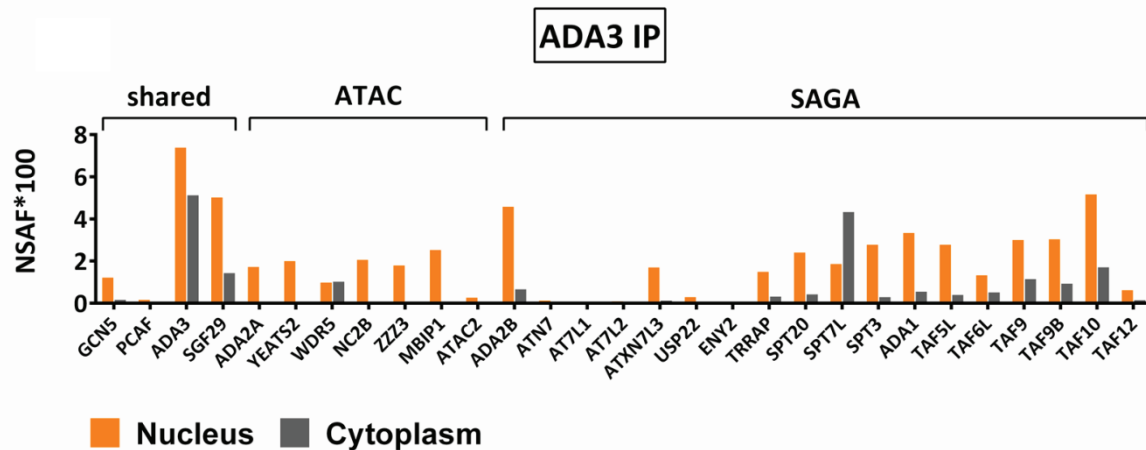


**Figure 4.13. ADA3 interacts with ATAC components only in the nucleus.**
NSAF quantification of SAGA and ATAC subunits in IP samples against endogenous ADA3, common subunit of ATAC and SAGA. ADA3 was immunoprecipitated from cytoplasmic and nuclear extracts prepared from HeLa cells. Quantification of ATAC and SAGA subunits present in anti-ADA3 IP. Only SAGA subunits associate with ADA3 in the cytoplasm of HeLa cells On the contrary, ATAC subunits are enriched only in the nucleus. Orange bars correspond to the IP from the nuclear extract and grey bars to the IP from the cytoplasmic extract.

The obtained NSAF values from the MudPIT analysis of this last set of nuclear and cytoplasmic IPs are illustrated in Figure 4.13. The nuclear IPs further indicated that, as expected, ADA3 interacts with components of both ATAC and SAGA complexes in this compartment. On the contrary, the quantitation of the identified proteins in the cytoplasmic IPs against ADA3 indicates different types of subunits interaction. Although shared between ATAC and SAGA, ADA3 associates almost exclusively with SAGA components in the cytoplasm.

Altogether, the quantitative proteomic analysis of this set of IPs revealed a clear difference in the association of SAGA and ATAC components in the two cellular compartments. We obtained evidence indicating that SAGA subunits assemble in the cytoplasm to form the SAGA holocomplex that should be actively imported into the nucleus. On the contrary our data lead us to the conclusion that the presence of the neo synthetized ATAC subunits in the cytoplasm is extremely dynamic and/or temporary, and the full assembly of ATAC complex probably takes place in the nucleus.

# 4.2 Discussion (II)

Many transcription-associated complexes share structural and functional subunits. The regulatory mechanisms that determine the allocation of these proteins between complexes, that in many cases have very different functions, are not clear. Since SAGA and ATAC have three common subunits (GCN5/PCAF, ADA3, SGF29), questions are raised on the way the incorporation of these components in each complex is regulated. So far it is not known which cellular mechanisms are responsible for the allocation of these proteins between the two complexes. Moreover, these three common subunits together with the complex specific ADA2a (for ATAC) and ADA2b (for SAGA) are part of the HAT containing module of the respective complexes. Since ATAC- and SAGA- specific HAT modules exhibit different catalytic properties, it is possible that the regulation of the assembly of their HAT modules is related to the different requirements of the cells under different conditions.

## 4.2.1 Evidence for differential subcellular compartmentalization of human ATAC and SAGA complex assembly

Although it has been shown that endogenous ADA2a (ATAC HAT-module specific) and other subunits of the same complex are mainly cytoplasmic (Orpinell *et al*, 2010), we observed that exogenously overexpressed ADA2a-eGFP was predominantly localised in the cytoplasm of human U2OS cells. This pattern indicated that either the overexpressed protein is not transferred into the nucleus due to saturation of the required import-associated factors or that it can still be imported but since it is present at levels much higher than the physiological, it is actively transported back to the cytoplasm. On the contrary, the SAGA-HAT module specific subunit, ADA2b, always appeared nuclear when overexpressed, showing that distinct mechanisms regulate the localization of the two proteins. Indeed, our live-cell imaging experiments, based on FLIP, showed that the reason of cytoplasmic accumulation of ADA2a-eGFP is not a blocked/saturated nuclear import pathway. Hence, we demonstrated that overexpressed ADA2aeGFP is able to enter into the nucleus and that the predominant steady state pattern of cytoplasmic accumulation, observed in fixed cells, is attributed to the rapid export of excess ADA2a-eGFP from the nucleus. This shuttling behaviour seems to be distinct from what has been described for TAF10 (common subunit of TFIID and SAGA complexes). It was shown that exogenously overexpressed TAF10 is predominantly cytoplasmic and is not

imported into the nucleus (Soutoglou *et al*, 2005). On the other hand what we observe for ADA2a is more similar to RPB1. Evidence has been provided that even when overexpressed, RPB1 can be imported into the nucleus despite the fact that it does not seem to incorporate into functional complexes (Sugaya *et al*, 2000; Custódio *et al*, 2006).

It has been shown that the relative abundance between individual subunits of transcription associated complexes can affect their intracellular distribution and this distribution can be related to the regulation of the assembly of the complexes (Soutoglou *et al*, 2005; Boulon *et al*, 2010). With a series of fixed and live cell imaging experiments, we found that among the three components of the HAT core module (GCN5, ADA3 and SGF29) it is only the presence of exogenously overexpressed GCN5 that can stabilise the overexpressed ADA2a-eGFP in the nucleus. However, as mentioned above, elevated levels of GCN5 are not required for the nuclear import of ADA2a-eGFP.

These findings suggest that i) ADA2a can be stably localised in the nucleus only in association with other factors and ii) that there is a given amount of ADA2a protein that the cells can incorporate into an ATAC complex or subassembly (potentially HAT-core) and that this amount seems to be defined from the availability of GCN5. Our findings provide strong indications that the interaction between GCN5 and ADA2a may be a key step in the assembly of ATAC specific HAT module. Moreover, the absence of a similar translocation effect by the co-expression of ADA3 or SGF29 (the two other subunits of the HAT module) underlines the specificity of GCN5-ADA2a interaction. The fact that we did not detect similar requirements for the stable localisation of ADA2b in the nucleus, could mean that either ADA2b can be imported and stabilised in the nucleus in the absence of interacting partners or that its nuclear import and retention is regulated via a different mechanism than that of ADA2a. Thus, it is possible that apart from their differential effect on GCN5 HAT specificity, ADA2a and ADA2b may also be involved in the regulation of ATAC and SAGA complex assembly.

To investigate further the above observations we wanted to identify the interacting partners of ADA2a and ADA2b in nucleus and cytoplasm. To this end, we performed IPs (followed by MudPIT analysis) targeting ATAC and/or SAGA in the two cellular compartments in three different ways. Initially, we IP-ed ADA2a and ADA2b, since those are the only different subunits between the HAT modules of the complexes and when exogenously overexpressed are found to have very different intracellular distribution and dynamics. Next, apart from ADA2a and ADA2b, we targeted an additional ATAC specific (ZZZ3) and a SAGA specific (SPT20) subunit. Finally, we chose to IP ADA3, a subunit that is found in both complexes and particularly in their HAT modules. In the cases of quantified anti-ADA2a and -ZZZ3 IPs, we found that all ATAC specific

subunits were detected only in samples deriving from nuclear extracts and never from cytoplasmic. On the contrary, IPs against ADA2b and SPT20 enriched our analysed samples for the total of SAGA subunits both in the nuclear and the cytoplasmic IPs. Importantly, the same pattern was confirmed in the analysed anti-ADA3 IPs. All ATAC and SAGA subunits were identified in anti-ADA3 IPs from nuclear extracts but only SAGA components were detected in the same IPs from cytoplasmic extracts. Therefore, in contrast to what we observe at high levels of ADA2a overexpression, our MudPIT data suggest that when the relative abundance of ATAC subunits is at the physiological levels, ADA2a is in the nucleus incorporated in the ATAC holocomplex.

We have considered the possibility that the observed differences in our IP-MudPIT analysis are due to different efficiency of antibodies in extracts deriving from the two compartments. Particularly, the affinity of the antibodies we used to IP endogenous ADA2a and ZZZ3 could be affected by potential PTMs that could occur on the antibody epitope in the nucleus and the cytoplasm. Alternatively, the antibody epitope may not be equally accessible in the two compartments due to binding of a protein that only occurs in the cytoplasm. However, we consider the above explanations highly unlikely, since the quantitation of samples from the cytoplasmic IP against the ATAC/SAGA shared subunit ADA3, resulted in the exclusive enrichment of SAGA subunits from the cytoplasm. Moreover, although our MudPIT results from different subunits of the same complex support the same conclusions, the generation of biological replicates for each analysed IP will be a priority for future experiments.

Thus, the combination of our imaging (obtained from overexpressed proteins) and quantitative proteomics data (obtained from endogenous proteins) revealed different aspects of SAGA and ATAC subunits interactions in the nucleus and the cytoplasm, that are probably related to the assembly pathways of the complexes. Importantly, as mentioned above, similar types of interactions between subunits of other transcription associated complexes have been described before.

Based on our observations, we can propose that the high enrichment for all SAGA components in the cytoplasmic IPs indicates that SAGA subunits are synthesised and associated into a full complex in the cytoplasm. It is possible that the events required for the cytoplasmic assembly of the complex lead to increased residence time and consequently higher availability of SAGA than ATAC-specific subunits in this compartment. This could explain why SAGA is always IP-ed as a holo-complex from the cytoplasm. Since it has been shown that the assembly of several multisubunit complexes may occur co-translationally (Duncan & Mata, 2011), such a mechanism of SAGA assembly could also be in agreement with what we observe. We could speculate that

after the association of all subunits is completed, the complex is actively imported into the nucleus where it is stably localised. On the other hand, our data suggest that ATAC subunits, synthesised in the cytoplasm, are probably directly and rapidly transported into the nucleus where the association of the holocomplex occurs. That could explain why we could not obtain significantly enriched IPs against the selected endogenous ATAC specific proteins from the cytoplasmic extracts. Rapid nuclear import of neo-synthetized ATAC subunits is a process which is supported from our FLIP experiments on overexpressed ADA2a-eGFP; we observed a rapid drop of nuclear fluorescence intensity when a cytoplasmic ROI was repeatedly bleached. Partial assembly of ATAC subcomplexes in the cytoplasm is not excluded but, if it occurs, they must be small enough to associate quickly and be rapidly imported in to the nucleus. Our data suggest that it is in the nucleus that the assembly of ATAC holocomplex takes places. Moreover, the effect of GCN5 abundance in ADA2a localisation indicates that, to a certain extent, ATAC holo-complex assembly depends on the availability of the subunits shared between this complex and SAGA.

### 4.2.2 Evaluating changes in localisation of ADA2a relative to ADA2b and GCN5 abundance

So far in our study, the localisation of ADA2a and ADA2b was assessed by overexpressing their eGFP-tagged versions. This approach was necessary and useful to reveal dynamic aspects that have to be amplified in order to be observed. For example, it would have been rather challenging to detect ADA2a nucleocytoplasmic shuttling at expression levels closer to the endogenous.

However, in future experiments, it would be very informative to examine the localisation of the endogenous proteins as well. ADA2a has been shown to be nuclear in a previous study from our lab (Orpinell *et al*, 2010). Although ADA2b is expected to localise in the nucleus too, the respective immunofluorescence (IF) experiment using an antibody against the endogenous protein has not been performed. It would be interesting to use IF against the endogenous proteins as readout to detect potential changes in their localisation in different conditions. For example the localisation of endogenous ADA2a and ADA2b could be examined when shared or complex specific subunits have been downregulated. That would be a direct way to obtain information about the regulation of localisation and import of the endogenous proteins. The first experiment of this series could include an anti-ADA2a IF in conditions that GCN5 is downregulated. We have observed that ADA2a requires a certain amount of available GCN5 to be stably localised in the nucleus. Thus we would expect that GCN5 knockdown would lead to cytoplasmic accumulation of endogenous ADA2a. Moreover, it would be interesting to test if in

endogenous conditions there is a common set of proteins that regulates ADA2b and ADA2a assembly and if there is competition for partners that favour nuclear import and/or retention. To test that, we could overexpress ADA2b and test the effect on the localisation of endogenous of ADA2a. In such conditions, if factors that regulate localisation are shared between the two proteins then endogenous ADA2a could be found enriched in the cytoplasm. This would be an indication that ADA2b affinity for these factors is higher and thus the presence of this subunit in the nucleus is favoured.

In addition, we have recently generated cell lines in which ADA2a-eGFP is highly overexpressed and accumulates in the cytoplasm. Those cells can be used to test conditions in which the localisation of the protein is shifted from cytoplasmic to nuclear. The question of potential competition for association with GCN5 (or other factors regulating localisation) between ADA2a and ADA2b could be investigated using these clones. For example, we could reduce the available amount of ADA2b by siRNA and examine if increased nuclear localisation of ADA2a is observed. Such a result would support the idea that ADA2b has higher affinity for GCN5 and it is this interaction that prioritises the incorporation of available GCN5 in SAGA specific HAT-modules. Similarly, we could investigate a potential increase in the nuclear signal of these clones in conditions that the integrity of the whole SAGA complex has been impaired.

### 4.2.3 Implications for the role of SWIRM domain in intracellular dynamics and assembly of human SAGA and ATAC

We observed that swapping the SWIRM domain of ADA2a with that of ADA2b affects the dynamics of the intracellular localisation of the former protein and particularly increases the fraction that is localised in the nucleus. However, the causes of this change are not clear. As it has been shown that the SWIRM domains mediate interactions with chromatin (Qian *et al*, 2005; Da *et al*, 2006), yet not with the same affinity (Yoneyama *et al*, 2007), the first possible explanation for the increased nuclear localisation of ADA2a_SWIRM swap mutant is that the SWIRM domain of ADA2b interacts more stably with chromatin than the one found in wt-ADA2a. Although it has been shown *in vitro* that ADA2a binds double stranded DNA and dinucleosomes (Qian *et al*, 2005), the properties of the respective ADA2b domain have not been studied with similar approaches. However, a role for the regulation of the HAT activity of the respective dATAC and dSAGA complexes in *D. melanogaster* has been attributed in the C-terminal region of dADA2a and dADA2b. C-terminal dADA2a and dADA2b chimeras were the only hybrid proteins that could partially rescue dADA2a or dADA2b mutant flies (Vamos & Boros, 2012). On the other

hand, *in vitro* studies showed that human ADA2b SWIRM is probably not required for efficient mononucleosome and chromatin acetylation nor for the formation of a heterotrimer with GCN5 and ADA3 (Gamper *et al*, 2009).

Another explanation for the phenotype we observe could be that swapping the SWIRM domain between the two proteins affects the interactions and association of the mutants with other complex subunits. To test this possibility, ADA2a and ADA2b swapped mutants can be transfected in U2OS cell and IP-ed via the eGFP tag. Then WB and/or MudPIT analysis could be performed to test if the interacting partners are affected from the presence of swapped domains. In this way it could be deciphered if the observed changes of intracellular distribution are related to the higher chromatin affinity of ADA2b SWIRM or to the association of ADA2a_SWIRM with SAGA subunits. In the latter case the increased nuclear localisation of ADA2a_SWIRM should be attributed to the incorporation into a SAGA-specific HAT module and potentially to a SAGA complex. The combinatorial effect of these two possibilities (increased chromatin binding affinity and interaction with different subunits) cannot be excluded.

## 4.2.4 Alternative imaging approaches that allow direct monitoring of nucleocytoplasmic shuttling

So far, FLIP was selected as the live-cell imaging technique to obtain information about the intracellular dynamics of overexpressed ADA2a in presence/absence of exogenous GCN5 and compare it with ADA2b. This technique is particularly useful to evaluate the continuity of protein transport between cellular compartments separated by a barrier like the nuclear membrane. However, the loss of fluorescence is an indirect way to study nucleocytoplasmic transport. Similar, indirect information for the shuttling of the proteins could be obtained in a FRAP approach by bleaching the whole nucleus or the whole cytoplasm and observe how recovery occurs (or not) due to shuttling of unbleached molecules from the reverse compartment.

Instead of photobleaching techniques, an alternative approach that would allow to directly observe the two processes (nuclear import/export of overexpressed ADA2a) could be based on photoconversion experiments. Particularly, a photoconvertible/photoswitchable protein could be used to tag ADA2a, ADA2b, GCN5 or any other protein of interest. Examples of such tags are Dendra2 and mEos (Gurskaya *et al*, 2006; Baker *et al*, 2010), that have been used in photoconversion experiments to track the dynamics of proteins in living cells. Prior to photoconversion, these proteins emit in green and after illumination with a blue/UV laser pulse their emission is switched to red. Thus, by photoconverting a pool of the overexpressed protein

(green), a distinct population of red molecules appears. With the appropriate imaging protocol the conversion of a sufficient number of molecules from green to red would allow to track the mobility of the converted population. Specifically, in the first set of experiments, a cytoplasmic ROI would be selected, to test if the molecules (e.g. ADA2a-Dendra2) that switch to red are transported into the nucleus. The second set of experiments should involve the selection of a nuclear ROI for photoconversion, to test if the molecules that convert to red are exported from the nucleus. In contrast to FLIP, photoconversion experiments as those described above would provide direct information about the directionality and different rates of nucleocytoplasmic shuttling of the proteins. On the other hand, it has to be noted that the shuttling processes as observed by photoconversion, are expected to reach a plateau of exchange much shorter that what is observed by FLIP. This has to do with the fact that, in photoconversion, the total number of molecules that is observed is much smaller than the molecules that are bleached during a FLIP experiment. A possible drawback of this approach could be that the time-frame may prove too short to reveal the distinct (slow or fast) components that may contribute differentially to the shuttling. However, combined information, gathered from FLIP and photoconversion experiments could decipher some aspects of the intracellular dynamics of ATAC and SAGA components.

### 4.2.5 Identifying domains responsible for ADA2a-GCN5 interaction in living cells

Apart from demonstrating that GCN5 affects the localisation pattern of overexpressed ADA2a, we also showed that the presence of GCN5 significantly affects the intranuclear dynamics of ADA2a reducing its rate of fluorescence recovery after photobleaching. This effect has to be further investigated to clarify which type of interactions are responsible for this dramatic change of ADA2a mobility. It would be interesting to test which domains of GCN5 could be involved in this effect. Since the bromodomain of GCN5 is known to function as a reader of acetylated lysines, we could assume that part of the observed reduction in ADA2a mobility is mediated from interactions of this domain with chromatin in the context of a GCN5-ADA2a complex. A deletion mutant lacking that domain could be co-transfected with ADA2a-eGFP to test if the effect is significantly different. Other GCN5 and ADA2a domains could also be included in such kind of experiments to dissect in living cells the region(s) of the two proteins involved in the observed effect.

## CONCLUSIONS

In this work, we described the differential dynamics of intracellular distribution of SAGA and ATAC specific ADA2 paralogues and provided evidence that the observed patterns are related to different assembly pathways of the complexes. We showed that exogenously overexpressed ADA2a and ADA2b differ significantly in their subcellular localisation dynamics. In addition, we demonstrated that the relative abundance of GCN5 and ADA2a affects the subcellular distribution of the latter protein. The application of quantitative proteomic analysis, based on MudPIT allowed us to expand our observations on endogenous proteins and provided additional evidence that the cytoplasmic and nuclear assembly pathways of SAGA and ATAC complexes are different. Since ATAC and SAGA also share subunits with other complexes, the complexity of the mechanisms underlying complex assembly should not be underestimated and combinatorial experimental approaches should be applied to obtain a complete view of these highly regulated processes.

# 5 Supplementary Data
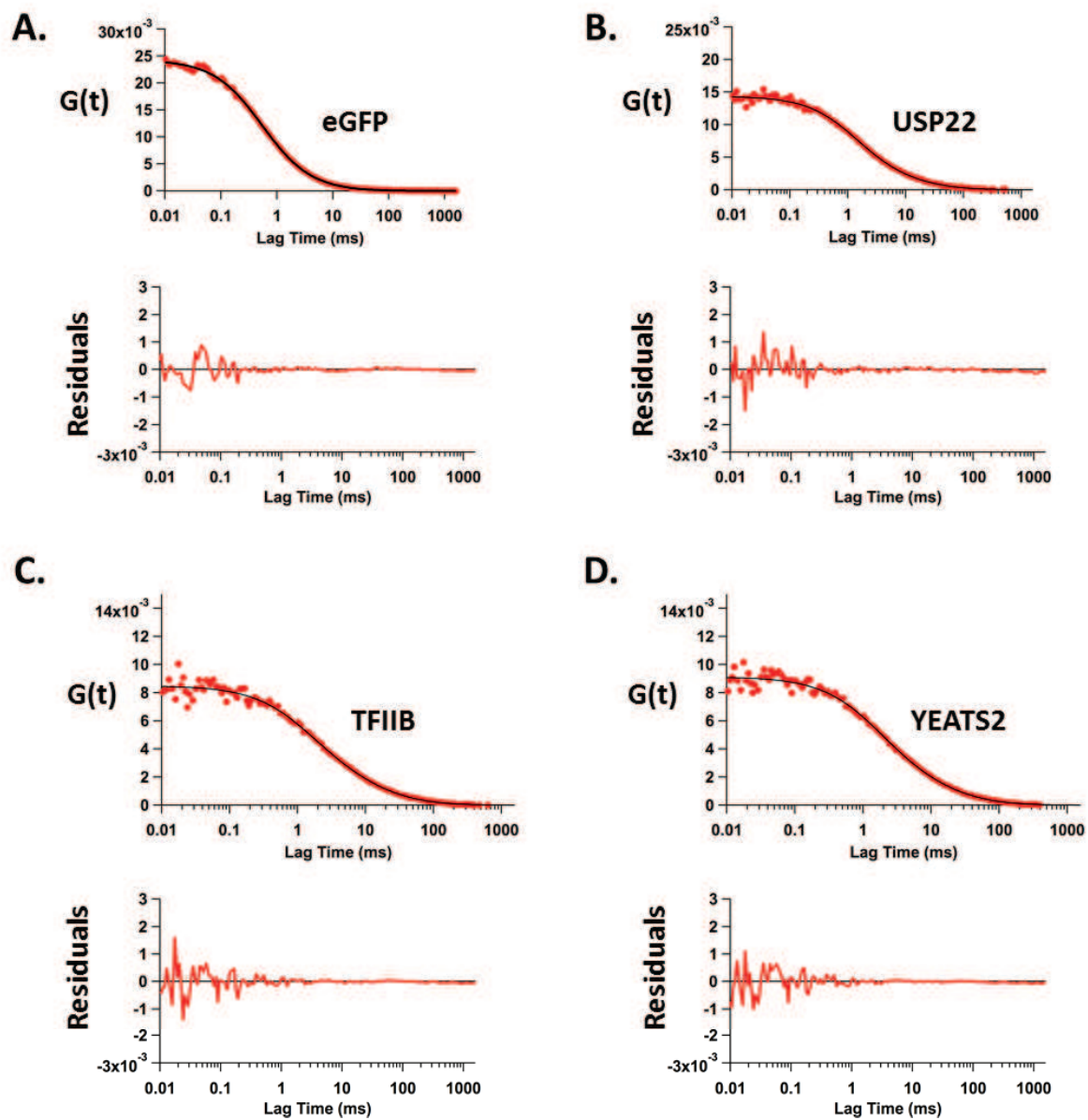


**Figure 5.1. Examples of nonlinear fitting of average autocorrelation curves of free eGFP and eGFP-tagged factors.**
**A.**: Fit of free eGFP autocorrelation curve with the one-component free diffusion model. **B, C, D.**: Fit of eGFP tagged USP22 (B), TFIIB (C) and YEATS2 (D) with the two-component, free-diffusion model. Below each graph, the residuals of the respective fit are shown.

# Résumé de thèse

## Introduction:

La transcription chez les eucaryotes est un phénomène complexe et très précisément régulé impliquant de très nombreux facteurs. Des activateurs transcriptionnels spécifiques de gènes, de nombreux co-facteurs ainsi que les facteurs généraux de transcription sont nécessaires pour permettre le recrutement de l'ARN polymérase II au promoteur et l'initiation de la transcription. Les modifications post-traductionnelles d'histones de la chromatine jouent un rôle important dans l'activation ou la répression de la transcription (Fuda *et al*, 2009).

L'acétylation des queues N-terminales des histones est une modification post-traductionnelle extensivement étudiée. Les histones acétyltransférases (HATs) induisent une ouverture de la structure de la chromatine augmentant ainsi l'accessibilité à l'ADN des facteurs nécessaires pour la transcription. Les HATs sont des complexes de modification de la chromatine qui sont fréquemment altérés dans les cancers. Des anomalies d'acétylation modifient l'expression de gènes suppresseurs de tumeurs et de proto-oncogènes. L'hyperacétylation de proto-oncogènes peut augmenter leur expression et donc leur propriétés oncogéniques alors que l'hypoacétylation de gènes suppresseurs de tumeurs va réprimer leur expression (Di Cerbo & Schneider, 2013).

Les complexes HAT humains SAGA et ATAC ont une organisation modulaire, portent différentes activités enzymatiques et semblent fonctionner comme co-activateurs de la transcription à des loci spécifiques. Les deux complexes partagent certaines sous-unités et ont des activités enzymatiques distinctes (Nagy & Tora, 2007). En plus de leur activité de modification de l'architecture chromatinienne, les complexes SAGA et ATAC agissent également sur des substrats non-histoniques. Ces deux complexes ont donc, en plus de leur rôle de co-activateurs de la transcription, d'importantes fonctions en tant que régulateurs de différents processus biologiques comme différentes voies de transduction du signal (e.g. PKC, MAPK, p53) ou la régulation du cycle cellulaire (Pankotai *et al*, 2005; Nagy *et al*, 2009; Orpinell *et al*, 2010). Différentes études permettent de lier des modification du niveau d'expression, de l'activité ou de la localisation cellulaire de plusieurs sous unités de ces complexes avec différents aspects de la carcinogénèse (Liu *et al*, 2003; Kurabe *et al*, 2007) (Mirza *et al*, 2013)

Une étude récente (Bonnet *et al*, 2014b), montre que SAGA agit sur l'ensemble des gènes transcrits et peut déubiquitiner l'histone H2B en quelques minutes. Cependant, nous n'avons que peu d'information sur les propriétés dynamiques de SAGA et ATAC dans des cellules vivantes ainsi que sur les voies de régulations contrôlant l'assemblage de ces complexes multiprotéiques. Pour mieux comprendre les mécanismes cellulaires controlant l'activité des complexes SAGA et ATAC dans des cellules vivantes humaines et pour définir les mécanismes d'assemblage de ces complexes, j'ai réalisé une approche expérimentale basée sur des techniques d'imagerie sur cellules vivantes en parallèle à des approches biochimiques couplées à des analyses protéomiques.

## Résultats:

**Partie I**

J'ai généré des constructions permettant l'expression de protéines fusionnées à eGFP pour différentes sous-unités des complexes SAGA et ATAC et réalisé des expériences de FRAP (Fluorescence Recovery After Photobleaching) après transfection de ces constructions dans les cellules U2OS. Les analyses de mes premières expériences montrent que toutes les sous-unités de SAGA et ATAC analysées interagissent de façon extrêmement dynamique avec leur environnement nucléaire (Figure 1).
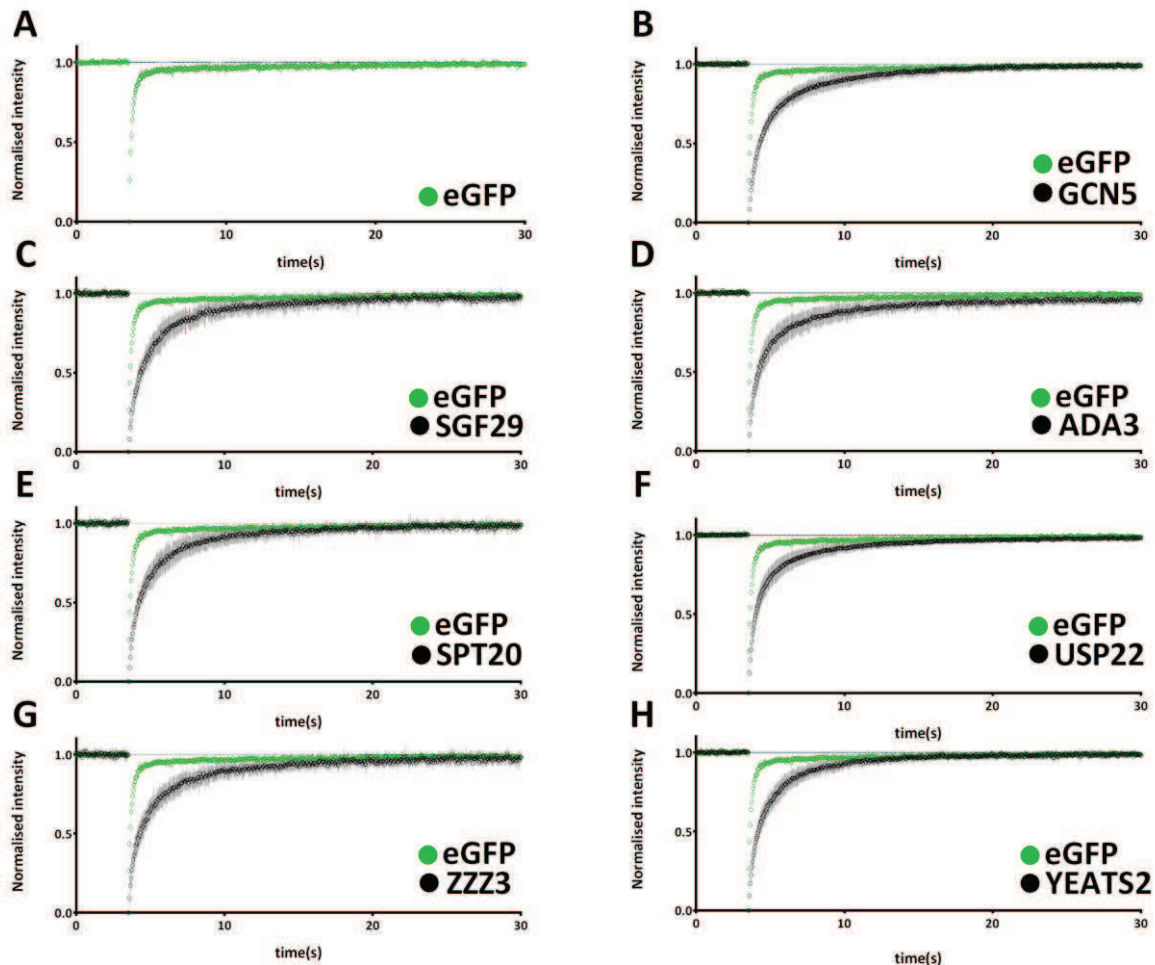
**Figure 1. FRAP courbes de sous-unités de SAGA et ATAC.**
Les courbes représentent les cinétiques de redistribution de fluorescence après photoblanchiment (FRAP) en présence de eGFP seule et de sous-unités de SAGA et ATAC taguées eGFP en cellules humaines U2OS. Pour toutes les courbes FRAP, (sur cette figure et celles qui suivent), le niveau de fluorescence préblanchie est normalisé à 1 et la profondeur de blanchiment normalisée à 0. Sur toutes les courbes, le bruit de fond a été enlevé, ainsi que le photoblanchiment observé. Pour toutes les courbes FRAP, la moyenne d'intensité du ROI photoblanchi est tracée en fonction du temps. **A.**: Recouvrement FRAP de GFP utilisé comme control de protéines diffusant librement. Le recouvrement de la GFP seule est plus rapide que celui de toutes les sous-unités SAGA et ATAC testées. **B, C**: recouvrement FRAP de GCN5 et SGF29, sous-unités partagées entre les deux complexes. **D, E**: recouvrement FRAP de SPT20 et USP22, sous-unités spécifiques de SAGA. **F**: recouvrement FRAP de ZZZ3, sous-unité spécifique d'ATAC. Dans tous les cas, les cellules ont été imagées dans 100 régions avant photoblanchiment de la région nucléaire sélectionnée.

J'ai comparé les dynamiques de ces protéines fusionnées à eGFP avec d'autres facteurs de transcription bien caractérisé et j'ai pu montrer que les complexes SAGA et ATAC humains font partie d'une classe de facteurs de transcription ayant une très forte mobilité intranucléaire (Figure 2).
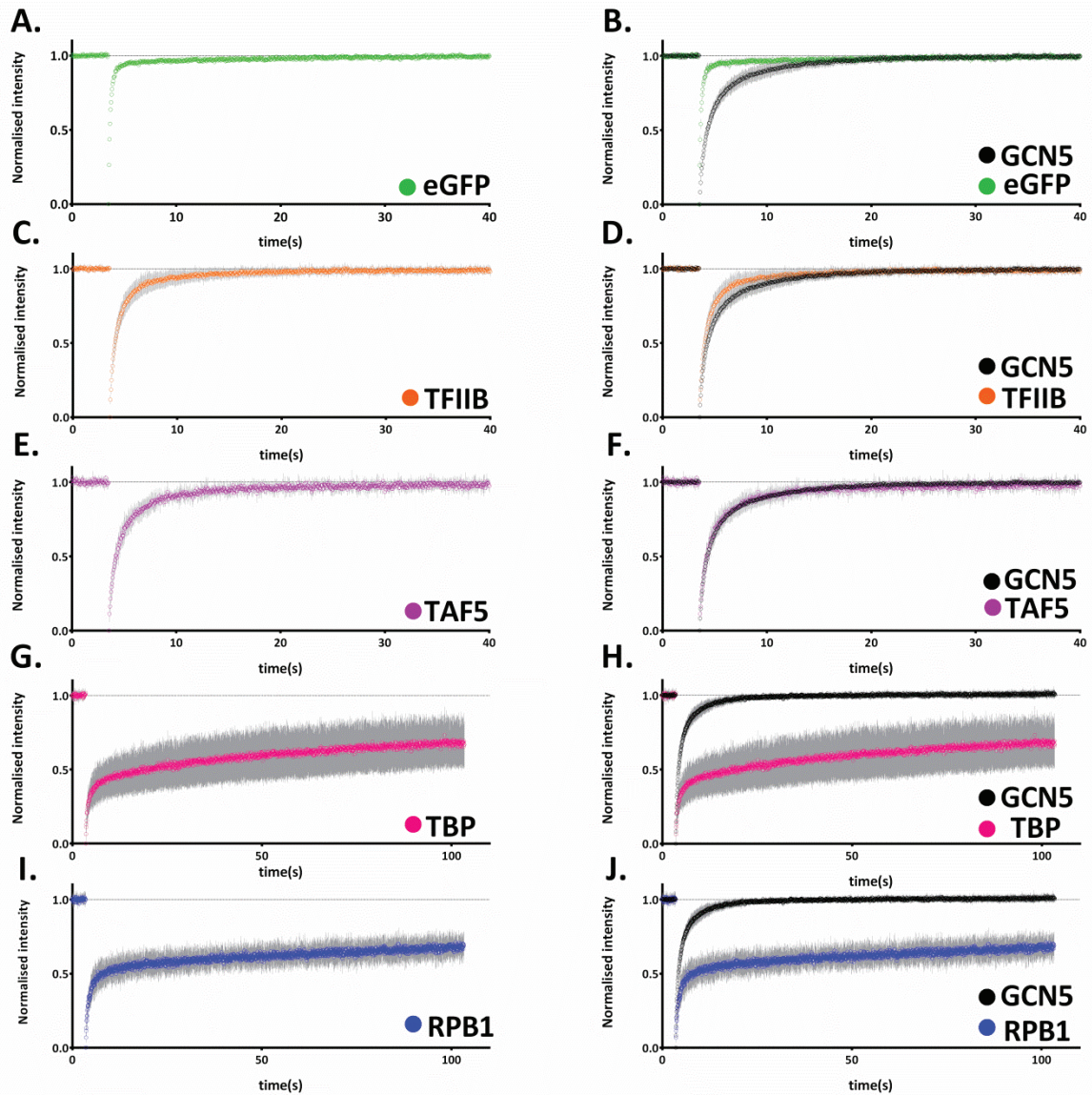
**Figure 2. Comparaison qualitative entre les courbes FRAP de GCN5 et autres facteurs de transcription.**

**A, B**: GCN5 récupère dans l'ordre de la seconde bien plus lentement que GFP diffusant librement **C** : La courbe FRAP TFIIB montre un recouvrement rapide. **C, D, E, F**: le recouvrement de GCN5 est similaire de celui de TFIIB et TAF5, indiquant que SAGA et ATAC appartiennent à une classe de complexes avec une mobilité intranucléaire importante. G, H : recouvrement FRAP de TBP, sous-unité de TFIID et comparaison avec GCN5. Par rapport à GCN5, TBP récupère dans l'ordre de la minute **I, J**: recouvrement FRAP de RPB1, sous-unité de RNA Pol II et comparaison avec GCN5. Par rapport à GCN5, RPB1 récupère dans l'ordre de la minute. Dans tous les cas, les cellules ont été photographiées sur approximativement 100 régions avant le blanchiment et le recouvrement de fluorescence a été suivi sur 2000 régions supplémentaires sur un temps total de 100secs. A noter que sur les graphes A-F, uniquement les 40 premières secondes de recouvrement sont montrées. Afin d'améliorer sur les premiers temps.

Afin de déterminer si la dynamique de ces complexes est modifiée par une inhibition de la transcription, j'ai répété ces expériences de FRAP dans des cellules contrôles et après traitement par DRB ou actinomycin-D (Figure 3). L'interprétation des résultats obtenus montre que la dynamique de ces deux complexes n'est pas altérée de façon significative après inhibition de la transcription par l'ARN polymérase II. Ces résultats sont en accord avec ceux décrits par Bonnet *et al*. (2014b) montrant que l'inhibition de la transcription par l'actinomycin-D n'affecte pas la dynamique de désubiquitination de H2B par le complexe SAGA.
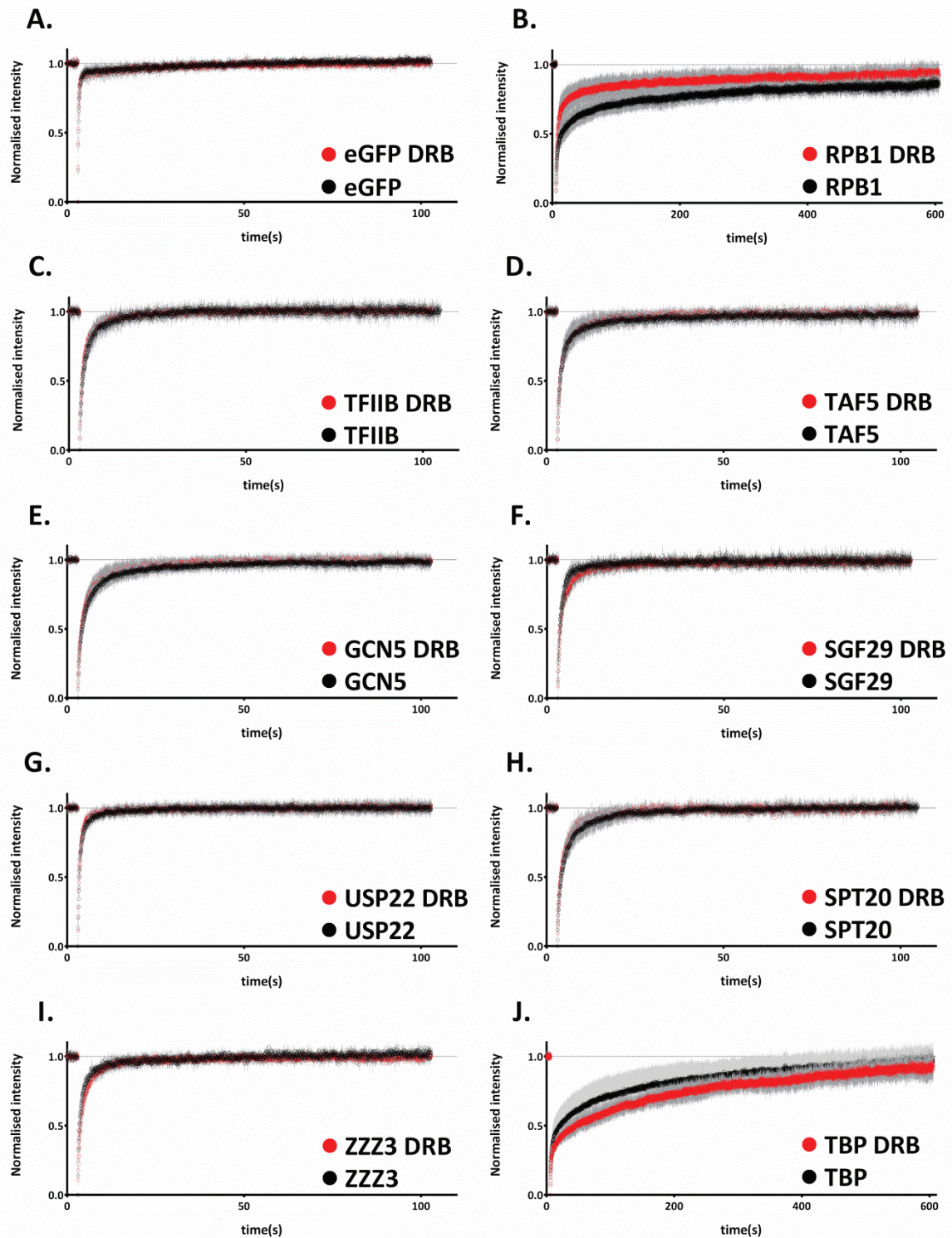
**Figure 3. DRB a un effet limité sur les courbes FRAP de la plupart des facteurs testés.** La comparaison qualitative des courbes FRAP entre les cellules non traitées et cellules traitées au DRB, montre que pour la plupart des facteurs testés, il n'y a pas de différence notable de recouvrement de fluorescence. En revanche, le traitement DRB entraine une réponse plus rapide pour RPB1 (**B**) et plus lente pour TBP (**J**). A noter qu'il n'y a pas d'effet détectable sur les courbes FRAP de GFP.

Ayant mis en évidence une très grande mobilité des sous-unités de SAGA et ATAC dans le noyau cellulaire, j'ai décidé de mettre en place dans le laboratoire une technique complémentaire afin d'obtenir des informations additionnelles sur les propriétés biophysiques de ces complexes. Pour cela j'ai utilisé la technique de FCS (Fluorescence Correlation Spectroscopy) qui permet l'étude de protéines ayant une mobilité élevée. L'utilisation de cette technique dans les cellules vivantes m'a permis de vérifier la forte mobilité des sous-unités de SAGA et ATAC dans le noyau et de confirmer que les propriétés dynamiques de ces complexes ne sont pas modifiées par l'inhibition de la transcription (Figure 4).
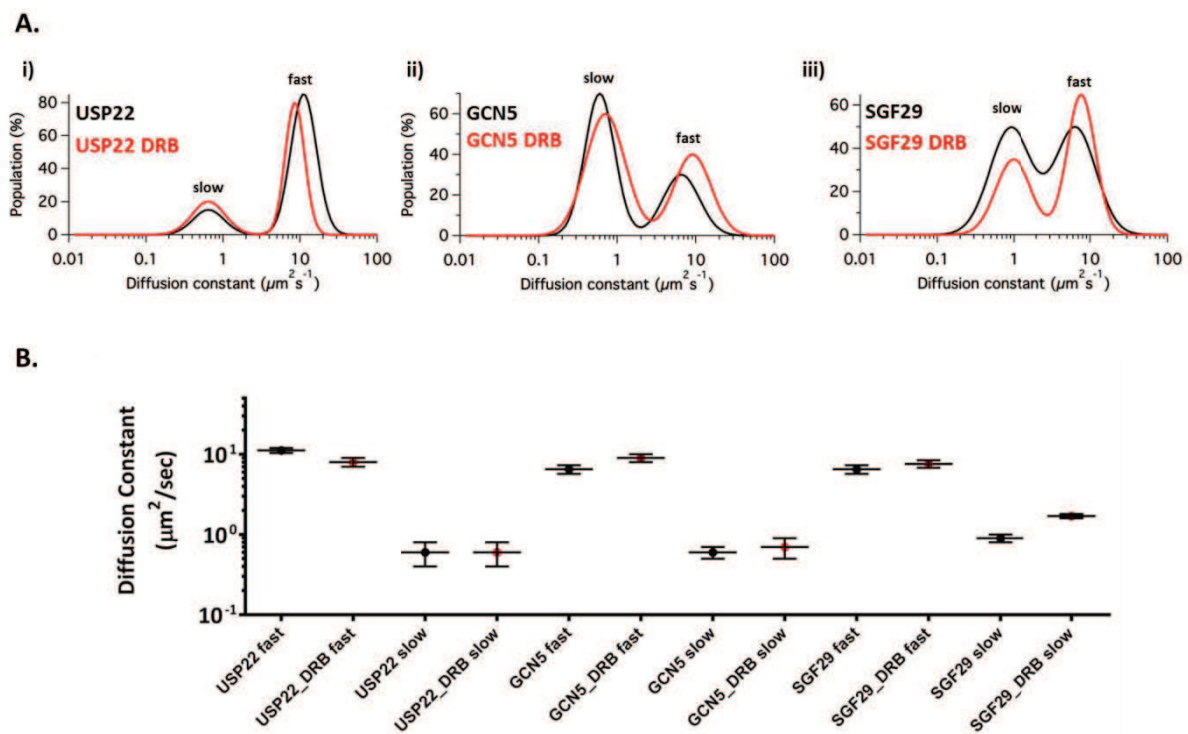


**Figure 4. Comparaison de _D_s entre GFP et les facteurs tagués GFP.**
Les graphes montrent la comparaison des constantes de diffusion des composantes rapides et lentes de chaque facteur testé et la constante de diffusion de GFP. **A.**: DGFP et $D_{fast}$ de tous les facteurs testés sont contenus dans le même graphe. A noter que des valeurs similaires $D_{fast}$ ont été obtenues pour toutes les protéines taguées GFP et sont plus faibles que $D_{GFP}$. **B.:** $D_{GFP}$ and $D_{slow}$ de tous les facteurs testés sont contenus dans le même graphe.

**Partie II**

Un criblage visant à analyser la localisation de sous-unités de SAGA ou ATAC fusionnées à eGFP m'a permis d'observer que la sous-unité ADA2a, un des quatre constituants du module HAT du complexe ATAC est essentiellement localisé dans le cytoplasme de la majorité des cellules transfectées. ADA2a interagit avec GCN5, ADA3 et SGF29 pour former le module HAT du complexe ATAC. ADA2b, paralogue d'ADA2a et sous-unité correspondante du module HAT présente toujours une localisation nucléaire. Nous supposons que les différences de localisation observées pour ces deux sous-unités correspondraient à des différences d'assemblage des modules HAT de ces deux complexes pouvant également expliquer des différences fonctionnelles.
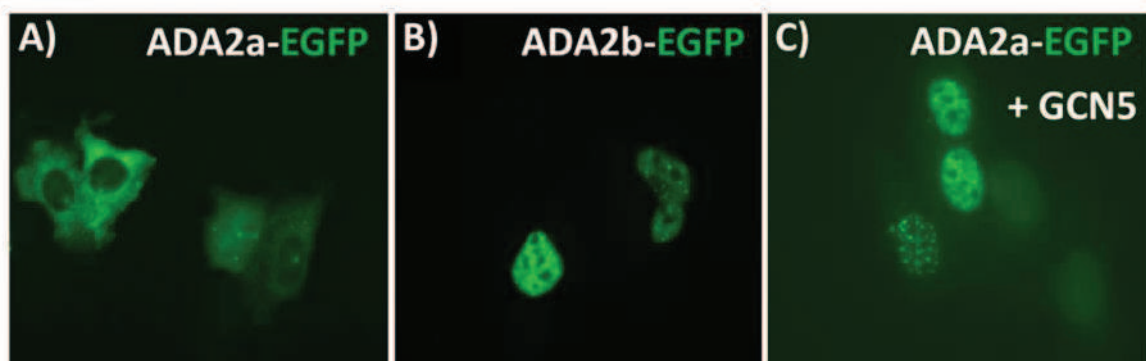


**Figure 5. Les surexpressions de ADA2a et ADA2b ont des localisations cellulaires différentes en cellules U2OS.**
ADA2a (sous-unité spécifique ATAC) et ADA2b (sous-unité spécifique SAGA) ont été clonés fusionnées à la eGFP et surexprimées en cellules U2OS ensuite fixées et analysées par microscopie à fluorescence. **A.**: Lorsque ADA2a-eGFP est surexprimée, elle s'accumule dans le cytoplasme de la plupart des cellules fixées. **B.**: Lorsque ADA2b-eGFP est surexprimée seule, elle s'accumule toujours dans le noyau. **C.**: Changement de localisation cellulaire de ADA2a-eGFP, de cytoplasmique à nucléaire, en présence de GCN5 surexprimée. Les cellules fixées, ont été analysées par microscopie à fluorescence (DM4000, Leica) avec le même temps d'exposition.

Pour mieux comprendre la localisation cellulaire d'ADA2a recombinante, j'ai caractérisé par des expériences d'imagerie différentes sous-unités de SAGA et ATAC fusionnées à eGFP. En utilisant différentes combinaisons de co-transfection, j'ai identifié GCN5 comme le facteur limitant pour la translocation nucléaire d'ADA2a (Figure 5.).

J'ai étudié par FLIP (Fluorescence Loss In Photobleaching) les protéines ADA2a et ADA2b fusionnées à eGFP afin de déterminer leur dynamique de localisation intracellulaire dans les cellules vivantes. Cette analyse m'a fourni de nouvelles informations concernant la régulation de ces deux protéines et de l'activité HAT des complexes SAGA et ATAC (Figure 6).
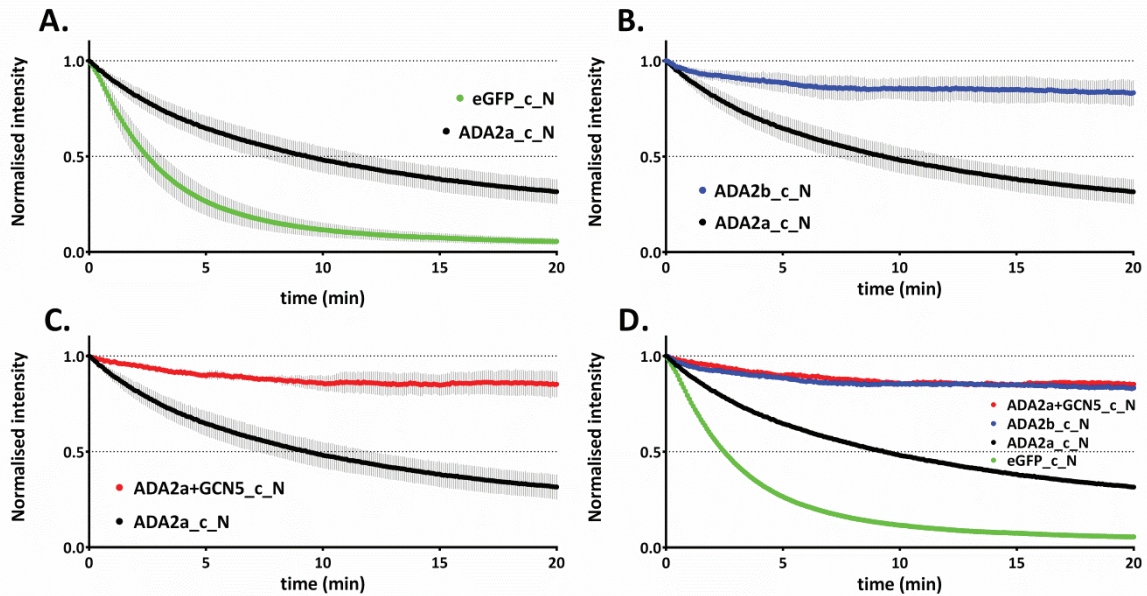
**Figure 6. Différentes dynamiques intracellulaires d'ADA2A et ADA2b sont révélées par FLIP.**
Les cellules ont été analysées par FLIP et l'intensité de fluorescence nucléaire normalisée a été mesurée dans le noyau de chaque cellule au cours du temps. **A**.: Graphe montrant la perte d'intensité de valeurs d'ADA2a-eGFP et eGFP nucléaires dans les cellules exprimant ces protéines. **B**.: La comparaison de valeurs FLIP entre ADA2a-eGFP et ADA2b-eGFP nucléaires révèle une différence importante de comportement de ces deux protéines. **C**) Les courbes FLIP montrent que lorsque GCN5 est surexprimé avec ADA2a-eGFP, ce dernier est stabilisé dans le noyau et n'est pas transporté dans le cytoplasme. **D**.: Les courbes FLIP des quatre protéines sont montrées sur le même graphe. La dynamique intracellulaire d'ADA2a surexprimé en présence de GCN5 est similaire de celle d'ADA2b. Ligne verte : eGFP, ligne noire : ADA2a-eGFP, ligne bleue : ADA2b-eGFP, ligne rouge : ADA2a-eGFP+GCN5.c_N: ROI blanchie dans le cytoplasme intensité normalisée mesurée dans le noyau.

Pour explorer les mécanismes d'assemblage des modules HAT de ces deux complexes, nous avons utilisé des approches de protéomiques basées sur la spectrométrie de masse pour identifier les partenaires des sous-unités endogènes d'ADA2a, ADA2b soit dans le noyau, soit dans le cytoplasme de cellules U2OS ou HeLa. Après immunoprécipitation de ces sous-unités à partir de chaque compartiment, les protéines associées ont été identifiées par analyse MudPIT (Multidimensional Protein Identification Technology). Cette approche nous a fourni de nouvelles informations nous aidant à mieux comprendre la distribution différentielle et la dynamique des sous-unités de ces complexes entre le noyau et le cytoplasme (Figure 3.7).
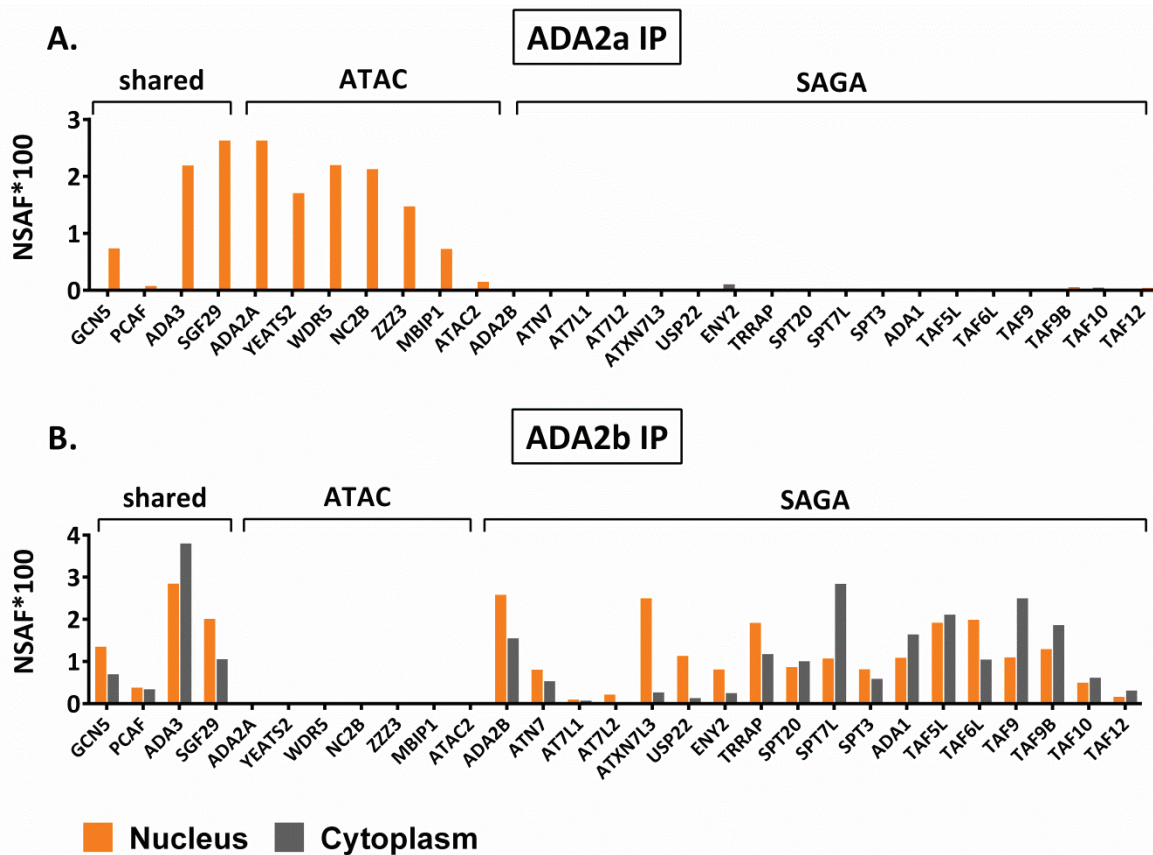
**Figure 3.7 Quantification par NSAF des sous-unités de SAGA et ATAC dans les échantillons d'IP dirigées contre ADA2a (spécifique ATAC) et ADA2b (spécifique SAGA) endogènes.**
Les deux protéines ont été immunoprécipitées à partir d'extraits cytoplasmique et nucléaires de cellules HeLa. **A.**: Quantification des sous-unités ATAC dans les données d'IP ADA2a. Les sous-unités d'ATAC sont enrichies uniquement dans le noyau et non pas dans le cytoplasme de cellules HeLa. **B.**: Quantification des sous-unités de SAGA identifiées dans l'IP ADA2b. Les sous-unités de SAGA sont enrichies à la fois dans le cytoplasme et dans le noyau.

## Conclusions:

La combinaison d'expériences de FRAP et de FCS dans les cellules vivantes, m'ont permis d'identifier pour la première fois la dynamique extrêmement importante des complexes SAGA et ATAC . Ces analyses montrent également que la mobilité de ces complexes et leur mode d'interaction avec la chromatine sont indépendants de la transcription. Les résultats obtenus en utilisant ces deux techniques donnent de nouvelles preuves aux résultats d'études biochimiques précédentes (Krebs *et al*, 2011; Bonnet *et al*, 2014b) suggérant que les complexes SAGA et ATAC interagiraient de façon transitoire avec la chromatine.

Des expériences d'imagerie sur cellules vivantes montrent, qu'après surexpression, les sous-unités ADA2a et ADA2b ont une dynamique intracellulaire différente. GCN5 a été identifiée comme la sous-unité clé du module HAT pour définir la dynamique d'ADA2a. Enfin les résultats de nos analyses MudPIT réalisées à partir d'immunoprécipitations contre les fractions cytoplasmiques ou nucléaires des sous-unités ADA2a, ADA2b, and GCN5 révèlent le réseau d'interactions entre les différents constituants des complexes ATAC et SAGA dans chaque compartiment cellulaire.

# References

Ahn SH, Kim M, Buratowski S (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Molecular Cell* **13:** 67-76

Allfrey VG, Faulkner R, Mirsky AE (1964) ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proceedings of the National Academy of Sciences of the United States of America* **51:** 786-794

Andrau JC, Van Oevelen CJ, Van Teeffelen HA, Weil PA, Holstege FC, Timmers HT (2002) Mot1p is essential for TBP recruitment to selected promoters during in vivo gene activation. *The EMBO Journal* **21:** 5173-5183

Arents G, Moudrianakis EN (1993) Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 10489-10493

Armache KJ, Mitterweger S, Meinhart A, Cramer P (2005) Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *The Journal of Biological Chemistry* **280:** 7131-7134

Asahara H, Santoso B, Guzman E, Du K, Cole PA, Davidson I, Montminy M (2001) Chromatin-dependent cooperativity between constitutive and inducible activation domains in CREB. *Molecular and Cellular Biology* **21:** 7892-7900

Atanassov BS, Evrard YA, Multani AS, Zhang Z, Tora L, Devys D, Chang S, Dent SY (2009) Gcn5 and SAGA regulate shelterin protein turnover and telomere maintenance. *Molecular Cell* **35:** 352-364

Auble DT, Hahn S (1993) An ATP-dependent inhibitor of TBP binding to DNA. *Genes & Development* **7:** 844-856

Auble DT, Wang D, Post KW, Hahn S (1997) Molecular analysis of the SNF2/SWI2 protein family member MOT1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. *Molecular and Cellular Biology* **17:** 4842-4851

Ausio J (2014) The shades of gray of the chromatin fiber: Recent literature provides new insights into the structure of chromatin. *BioEssays : news and reviews in molecular, cellular and developmental biology*

Ausio J, Dong F, van Holde KE (1989) Use of selectively trypsinized nucleosome core particles to analyze the role of the histone "tails" in the stabilization of the nucleosome. *Journal of Molecular Biology* **206:** 451-463

Azoulay J, Clamme JP, Darlix JL, Roques BP, Mély Y (2003) Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *Journal of Molecular Biology* **326:** 691-700

Baarends WM, Wassenaar E, van der Laan R, Hoogerbrugge J, Sleddens-Linkels E, Hoeijmakers JH, de Boer P, Grootegoed JA (2005) Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. *Molecular and Cellular Biology* **25:** 1041-1053

Baker SM, Buckheit RW, Falk MM (2010) Green-to-red photoconvertible fluorescent proteins: tracking cell and protein dynamics on standard wide-field mercury arc-based microscopes. *BMC Cell Biology* **11:** 15

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. *The EMBO Journal* **28:** 3785-3798

Basehoar AD, Zanton SJ, Pugh BF (2004) Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116:** 699-709

Batta K, Zhang Z, Yen K, Goffman DB, Pugh BF (2011) Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes & Development* **25:** 2254-2265

Baxevanis AD, Arents G, Moudrianakis EN, Landsman D (1995) A variety of DNA-binding and multimeric proteins contain the histone fold motif. *Nucleic Acids Research* **23:** 2685-2691

Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science (New York, NY)* **301:** 1090-1093

Bensaude O (2011) Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* **2:** 103-108

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM (1998) Localization of ASH1 mRNA particles in living yeast. *Molecular Cell* **2:** 437-445

Bhaumik SR, Raha T, Aiello DP, Green MR (2004) In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes & Development* **18:** 333-343

Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Barsyte D, Li J, Wu M, Zhou BO, Fleharty BE, Paulson A, Allali-Hassani A, Zhou J-QQ, Mer G, Grant PA, Workman JL, Zang J, Min J (2011) Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *The EMBO Journal* **30:** 2829-2842

Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P, Berger I (2013) The architecture of human general transcription factor TFIID core complex. *Nature* **493:** 699-702

Black JC, Van Rechem C, Whetstine JR (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Molecular Cell* **48:** 491-507

Bonnet J, Devys D, Tora L (2014a) Histone H2B ubiquitination: signaling not scrapping. *Drug discovery today Technologies* **12:** 27

Bonnet J, Wang C-YY, Baptista T, Vincent SD, Hsiao W-CC, Stierle M, Kao C-FF, Tora L, Devys D (2014b) The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. *Genes & Development* **28:** 1999-2012

Bonnet J, Wang YH, Spedale G, Atkinson RA, Romier C, Hamiche A, Pijnappel WW, Timmers HT, Tora L, Devys D, Kieffer B (2010) The structural plasticity of SCA7 domains defines their differential nucleosome-binding properties. *EMBO Reports* **11:** 612-618

Boulon S, Pradet-Balade B, Verheggen C, Molle D, Boireau S, Georgieva M, Azzag K, Robert M-CC, Ahmad Y, Neel H, Lamond AI, Bertrand E (2010) HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Molecular Cell* **39:** 912-924

Brand M, Leurent C, Mallouh V, Tora L, Schultz P (1999a) Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. *Science* **286:** 2151-2153

Brand M, Moggs JG, Oulad-Abdelghani M, Lejeune F, Dilworth FJ, Stevenin J, Almouzni G, Tora L (2001) UV-damaged DNA-binding protein in the TFTC complex links DNA damage recognition to nucleosome acetylation. *The EMBO Journal* **20:** 3187-3196

Brand M, Yamamoto K, Staub A, Tora L (1999b) Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. *The Journal of Biological Chemistry* **274:** 18285-18289

Brazda P, Szekeres T, Bravics B, Tóth K, Vámosi G, Nagy L (2011) Live-cell fluorescence correlation spectroscopy dissects the role of coregulator exchange and chromatin binding in retinoic acid receptor mobility. *Journal of Cell Science* **124:** 3631-3642

Bregman DB, Halaban R, van Gool AJ, Henning KA, Friedberg EC, Warren SL (1996) UV-induced ubiquitination of RNA polymerase II: a novel modification deficient in Cockayne syndrome cells. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 11586-11590

Briggs SD, Xiao T, Sun Z-WW, Caldwell JA, Shabanowitz J, Hunt DF, Allis CD, Strahl BD (2002) Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* **418:** 498

Brown CE, Howe L, Sousa K, Alley SC, Carrozza MJ, Tan S, Workman JL (2001) Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science* **292:** 2333-2337

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84:** 843-851

Bu P, Evrard YA, Lozano G, Dent SY (2007) Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. *Molecular and Cellular Biology* **27:** 3405-3416

Bucheli M, Sweder K (2004) In UV-irradiated Saccharomyces cerevisiae, overexpression of Swi2/Snf2 family member Rad26 increases transcription-coupled repair and repair of the non-transcribed strand. *Molecular Microbiology* **52:** 1653-1663

Bulseco DA, Wolf DE (2013) Fluorescence correlation spectroscopy: molecular complexing in solution and in living cells. *Methods in Cell Biology* **114:** 489-524

Buratowski S, Hahn S, Guarente L, Sharp PA (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56:** 549-561

Burke TL, Miller JL, Grant PA (2013) Direct Inhibition of Gcn5 Protein Catalytic Activity by Polyglutamine-expanded Ataxin-7. *The Journal of Biological Chemistry* **288:** 34266-34275

Cai Y, Jin J, Swanson SK, Cole MD, Choi SH, Florens L, Washburn MP, Conaway JW, Conaway RC (2010) Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. *The Journal of Biological Chemistry* **285:** 4268-4272

Cao R, Tsukada Y, Zhang Y (2005) Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Molecular Cell* **20:** 845-854

Carisey A, Stroud M, Tsang R, Ballestrem C (2011) Fluorescence recovery after photobleaching. *Methods in Molecular Biology* **769:** 387-402

Carles C, Treich I, Bouet F, Riva M, Sentenac A (1991) Two additional common subunits, ABC10 alpha and ABC10 beta, are shared by yeast RNA polymerases. *The Journal of Biological Chemistry* **266:** 24092-24096

Carre C, Shiekhattar R (2011) Human GTPases associate with RNA polymerase II to mediate its nuclear import. *Molecular and Cellular Biology* **31:** 3953-3962

Carrozza MJ, Utley RT, Workman JL, Côté J (2003) The diverse functions of histone acetyltransferase complexes. *Trends in Genetics* **19:** 321-329

Celic I, Masumoto H, Griffith WP, Meluh P, Cotter RJ, Boeke JD, Verreault A (2006) The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Current Biology* **16:** 1280-1289

Chambers RS, Wang BQ, Burton ZF, Dahmus ME (1995) The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. *The Journal of Biological Chemistry* **270:** 14962-14969

Chandrasekharan MB, Huang F, Sun Z-WW (2009) Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 16686-16691

Chen D, Dundr M, Wang C, Leung A, Lamond A, Misteli T, Huang S (2005) Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *The Journal of Cell Biology* **168:** 41-54

Chen D, Hinkley CS, Henry RW, Huang S (2002) TBP dynamics in living human cells: constitutive association of TBP with mitotic chromosomes. *Molecular Biology of the Cell* **13:** 276-284

Cheng B, Price DH (2007) Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. *The Journal of Biological Chemistry* **282:** 21901-21912

Chiang CM, Roeder RG (1995) Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. *Science* **267:** 531-536

Chicca JJ, 2nd, Auble DT, Pugh BF (1998) Cloning and biochemical characterization of TAF-172, a human homolog of yeast Mot1. *Molecular and Cellular Biology* **18:** 1701-1710

Ciurciu A, Komonyi O, Pankotai T, Boros IM (2006) The *Drosophila* histone acetyltransferase Gcn5 and transcriptional adaptor Ada2a are involved in nucleosomal histone H4 acetylation. *Molecular and Cellular Biology* **26:** 9413-9423

Clamme JP, Azoulay J, Mély Y (2003) Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophysical Journal* **84:** 1960-1968

Clapier CR, Cairns BR (2009) The biology of chromatin remodeling complexes. *Annual Review of Biochemistry* **78:** 273-304

Cojocaru M, Jeronimo C, Forget D, Bouchard A, Bergeron D, Cote P, Poirier GG, Greenblatt J, Coulombe B (2008) Genomic location of the human RNA polymerase II general machinery: evidence for a role of TFIIF and Rpb7 at both early and late stages of transcription. *The Biochemical Journal* **409:** 139-147

Coleman RA, Pugh BF (1995) Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. *The Journal of Biological Chemistry* **270:** 13850-13859

Coleman RA, Taggart AK, Benjamin LR, Pugh BF (1995) Dimerization of the TATA binding protein. *The Journal of Biological Chemistry* **270:** 13842-13849

Coleman RA, Taggart AK, Burma S, Chicca JJ, 2nd, Pugh BF (1999) TFIIA regulates TBP and TFIID dimers. *Molecular Cell* **4:** 451-457

Collart MA (1996) The NOT, SPT3, and MOT1 genes functionally interact to regulate transcription at core promoters. *Molecular and Cellular Biology* **16:** 6668-6676

Comer FI, Hart GW (2001) Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry* **40:** 7845-7852

Conaway RC, Conaway JW (1989) An RNA polymerase II transcription factor has an associated DNA-dependent ATPase (dATPase) activity strongly stimulated by the TATA region of promoters. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 7356-7360

Conaway RC, Sato S, Tomomori-Sato C, Yao T, Conaway JW (2005) The mammalian Mediator complex and its role in transcriptional regulation. *Trends in Biochemical Sciences* **30:** 250-255

Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322:** 1845-1848

Cramer P (2000) Architecture of RNA Polymerase II and Implications for the Transcription Mechanism. *Science* **288**

Cramer P, Armache KJJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn CDD, Lehmann E, Leike K, Sydow JF, Vannini A (2008) Structure of eukaryotic RNA polymerases. *Annual Review of Biophysics* **37:** 337-352

Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292:** 1863-1876

Crosio C, Fimia GM, Loury R, Kimura M, Okano Y, Zhou H, Sen S, Allis CD, Sassone-Corsi P (2002) Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Molecular and Cellular Biology* **22:** 874-885

Custódio N, Antoniou M, Carmo-Fonseca M (2006) Abundance of the largest subunit of RNA polymerase II in the nucleus is regulated by nucleo-cytoplasmic shuttling. *Experimental Cell Research* **312:** 2557-2567

Czeko E, Seizl M, Augsberger C, Mielke T, Cramer P (2011) Iwr1 directs RNA polymerase II nuclear import. *Molecular Cell* **42:** 261-266

Da G, Lenkart J, Zhao K, Shiekhattar R, Cairns BR, Marmorstein R (2006) Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 2057-2062

Daniel JA, Torok MS, Sun Z-WW, Schieltz D, Allis CD, Yates JR, Grant PA (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *The Journal of Biological Chemistry* **279:** 1867-1871

Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, Singer RH (2007) In vivo dynamics of RNA polymerase II transcription. *Nature Structural & Molecular Biology* **14:** 796-806

Davidson I (2003) The genetics of TBP and TBP-related factors. *Trends in Biochemical Sciences* **28:** 391-398

de Graaf P, Mousson F, Geverts B, Scheer E, Tora L, Houtsmuller AB, Timmers HT (2010) Chromatin interaction of TATA-binding protein is dynamically regulated in human cells. *Journal of Cell Science* **123:** 2663-2671

Demény MA, Soutoglou E, Nagy Z, Scheer E, Jànoshàzi A, Richardot M, Argentini M, Kessler P, Tora L (2007) Identification of a small TAF complex and its role in the assembly of TAF-containing complexes. *PloS One* **2**

Deng W, Roberts SG (2007) TFIIB and the regulation of transcription by RNA polymerase II. *Chromosoma* **116:** 417-429

Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399:** 491-496

di Bari MG, Ciuffini L, Mingardi M, Testi R, Soddu S, Barila D (2006) c-Abl acetylation by histone acetyltransferases regulates its nuclear-cytoplasmic localization. *EMBO Reports* **7:** 727-733

Di Cerbo V, Schneider R (2013) Cancers with wrong HATs: the impact of acetylation. *Briefings in Functional Genomics* **12:** 231-243

Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biology* **6:** 227

Dion V, Coulombe B (2003) Interactions of a DNA-bound transcriptional activator with the TBP-TFIIA-TFIIB-promoter quaternary complex. *The Journal of Biological Chemistry* **278:** 11495-11501

Dori-Bachash M, Shalem O, Manor YS, Pilpel Y, Tirosh I (2012) Widespread promoter-mediated coordination of transcription and mRNA degradation. *Genome Biology* **13**

Drapkin R, Reardon JT, Ansari A, Huang JC, Zawel L, Ahn K, Sancar A, Reinberg D (1994) Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* **368:** 769-772

Dubois MF, Bellier S, Seo SJ, Bensaude O (1994) Phosphorylation of the RNA polymerase II largest subunit during heat shock and inhibition of transcription in HeLa cells. *Journal of Cellular Physiology* **158:** 417-426

Duncan CD, Mata J (2011) Widespread cotranslational formation of protein complexes. *PLoS Genetics* **7**

Dundr M, Hoffmann-Rohrer U, Hu Q, Grummt I, Rothblum LI, Phair RD, Misteli T (2002) A kinetic framework for a mammalian RNA polymerase in vivo. *Science* **298:** 1623-1626

Eberharter A, Sterner DE, Schieltz D, Hassan A, Yates JR, Berger SL, Workman JL (1999) The ADA complex is a distinct histone acetyltransferase complex in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **19:** 6621-6631

Edwards AM, Kane CM, Young RA, Kornberg RD (1991) Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *The Journal of Biological Chemistry* **266:** 71-75

Egloff S, Zaborowska J, Laitem C, Kiss T, Murphy S (2012) Ser7 phosphorylation of the CTD recruits the RPAP2 Ser5 phosphatase to snRNA genes. *Molecular Cell* **45:** 111-122

Eisenmann DM, Arndt KM, Ricupero SL, Rooney JW, Winston F (1992) SPT3 interacts with TFIID to allow normal transcription in Saccharomyces cerevisiae. *Genes & Development* **6:** 1319-1331

Eklund G, Lang S, Glindre J, Ehlén Å, Alvarado-Kristensson M (2014) The nuclear localization of γ-tubulin is regulated by SadB-mediated phosphorylation. *The Journal of Biological Chemistry* **289:** 21360-21373

Ellenberg J, Siggia ED, Moreira JE, Smith CL, Presley JF, Worman HJ, Lippincott-Schwartz J (1997) Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. *The Journal of Cell Biology* **138:** 1193-1206

Fang SM, Burton ZF (1996) RNA polymerase II-associated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. *The Journal of Biological Chemistry* **271:** 11703-11709

Fire A, Samuels M, Sharp PA (1984) Interactions between RNA polymerase II, factors, and template leading to accurate transcription. *The Journal of Biological Chemistry* **259:** 2509-2516

Fishburn J, Mohibullah N, Hahn S (2005) Function of a eukaryotic transcription activator during the transcription cycle. *Molecular Cell* **18:** 369-378

Fleming AB, Kao CF, Hillyer C, Pikaart M, Osley MA (2008) H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Molecular Cell* **31:** 57-66

Fletcher TM, Hansen JC (1995) Core histone tail domains mediate oligonucleosome folding and nucleosomal DNA organization through distinct molecular mechanisms. *The Journal of Biological Chemistry* **270:** 25359-25362

Flores O, Lu H, Killeen M, Greenblatt J, Burton ZF, Reinberg D (1991) The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 9999-10003

Forget D, Lacombe A-AA, Cloutier P, Al-Khoury R, Bouchard A, Lavallée-Adam M, Faubert D, Jeronimo C, Blanchette M, Coulombe B (2010) The protein interaction network of the human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Molecular & Cellular Proteomics : MCP* **9:** 2827-2839

Forget D, Lacombe A-AA, Cloutier P, Lavallée-Adam M, Blanchette M, Coulombe B (2013) Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2. *Nucleic Acids Research* **41:** 6881-6891

Fromaget M, Cook PR (2007) Photobleaching reveals complex effects of inhibitors on transcribing RNA polymerase II in living cells. *Experimental Cell Research* **313:** 3026-3033

Fuda NJ, Ardehali MB, Lis JT (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461:** 186-192

Gamper AM, Kim J, Roeder RG (2009) The STAGA subunit ADA2b is an important regulator of human GCN5 catalysis. *Molecular and Cellular Biology* **29:** 266-280

Gangloff YG, Sanders SL, Romier C, Kirschner D, Weil PA, Tora L, Davidson I (2001) Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. *Molecular and Cellular Biology* **21:** 1841-1853

Gavin A-CC, Bösche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon A-MM, Cruciat C-MM, Remor M, Höfert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier M-AA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415:** 141-147

Gay F, Calvo D, Lo MC, Ceron J, Maduro M, Lin R, Shi Y (2003) Acetylation regulates subcellular localization of the Wnt signaling nuclear effector POP-1. *Genes & Development* **17:** 717-722

Gazdag E, Rajkovic A, Torres-Padilla ME, Tora L (2007) Analysis of TATA-binding protein 2 (TBP2) and TBP expression suggests different roles for the two proteins in regulation of gene expression during oogenesis and early mouse development. *Reproduction* **134:** 51-62

Geisberg JV, Moqtaderi Z, Kuras L, Struhl K (2002) Mot1 associates with transcriptionally active promoters and inhibits association of NC2 in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **22:** 8122-8134

Gelman L, Feige JN, Tudor C, Engelborghs Y, Wahli W, Desvergne B (2006) Integrating nuclear receptor mobility in models of gene regulation. *Nuclear Receptor Signaling* **4:** e010

Georgakopoulos T, Thireos G (1992) Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *The EMBO Journal* **11:** 4145-4152

Giglia-Mari G, Theil AF, Mari PO, Mourgues S, Nonnekens J, Andrieux LO, de Wit J, Miquel C, Wijgers N, Maas A, Fousteri M, Hoeijmakers JH, Vermeulen W (2009) Differentiation driven changes in the dynamic organization of Basal transcription initiation. *PLoS Biology* **7:** e1000220

Gill G, Pascal E, Tseng ZH, Tjian R (1994) A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the *Drosophila* TFIID complex and mediates transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 192-196

Gillette TG, Gonzalez F, Delahodde A, Johnston SA, Kodadek T (2004) Physical and functional association of RNA polymerase II and the proteasome. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 5904-5909

Goodrich JA, Hoey T, Thut CJ, Admon A, Tjian R (1993) *Drosophila* TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. *Cell* **75:** 519-530

Gorski SA, Snyder SK, John S, Grummt I, Misteli T (2008) Modulation of RNA polymerase assembly dynamics in transcriptional regulation. *Molecular Cell* **30:** 486-497

Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG (2007) Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Molecular Cell* **25:** 31-42

Grant PA, Duggan L, Côté J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL, Workman JL (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & Development* **11:** 1640-1650

Grant PA, Eberharter A, John S, Cook RG, Turner BM, Workman JL (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *The Journal of Biological Chemistry* **274:** 5895-5900

Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR, 3rd, Workman JL (1998) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94:** 45-53

Guelman S, Kozuka K, Mao Y, Pham V, Solloway MJ, Wang J, Wu J, Lill JR, Zha J (2009) The double-histone-acetyltransferase complex ATAC is essential for mammalian development. *Molecular and Cellular Biology* **29:** 1176-1188

Guelman S, Suganuma T, Florens L, Swanson SK, Kiesecker CL, Kusch T, Anderson S, Yates JR, 3rd, Washburn MP, Abmayr SM, Workman JL (2006) Host cell factor and an uncharacterized SANT domain protein are stable components of ATAC, a novel dAda2A/dGcn5-containing histone acetyltransferase complex in *Drosophila*. *Molecular and Cellular Biology* **26:** 871-882

Gurskaya NG, Verkhusha VV, Shcheglov AS, Staroverov DB, Chepurnykh TV, Fradkov AF, Lukyanov S, Lukyanov KA (2006) Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nature Biotechnology* **24:** 461-465

Ha I, Roberts S, Maldonado E, Sun X, Kim LU, Green M, Reinberg D (1993) Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. *Genes & Development* **7:** 1021-1032

Haldar D, Kamakaka RT (2008) Schizosaccharomyces pombe Hst4 functions in DNA damage response by regulating histone H3 K56 acetylation. *Eukaryotic Cell* **7:** 800-813

Hall DB, Struhl K (2002) The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo. *The Journal of Biological Chemistry* **277:** 46043-46050

Hampsey M (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews* **62:** 465-503

Han Y, Luo J, Ranish J, Hahn S (2014) Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex. *The EMBO Journal* **33:** 2534-2546

Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, Workman JL (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111:** 369-379

Haustein E, Schwille P (2007) Fluorescence correlation spectroscopy: novel variations of an established technique. *Annual Review of Biophysics and Biomolecular Structure* **36:** 151-169

Heidemann M, Hintermair C, Voß K, Eick D (2013) Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochimica et Biophysica Acta* **1829:** 55-62

Helmlinger D, Hardy S, Abou-Sleymane G, Eberlin A, Bowman AB, Gansmuller A, Picaud S, Zoghbi HY, Trottier Y, Tora L, Devys D (2006) Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biology* **4:** e67

Helmlinger D, Marguerat S, Villén J, Swaney DL, Gygi SP, Bähler J, Winston F (2011) Tra1 has specific regulatory roles, rather than global functions, within the SAGA co-activator complex. *The EMBO Journal* **30:** 2843-2852

Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NC, Kao CF, Pillus L, Shilatifard A, Osley MA, Berger SL (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes & Development* **17:** 2648-2663

Hernandez N (1993) TBP, a universal eukaryotic transcription factor? *Genes & Development* **7:** 1291-1308

Hershko A, Ciechanover A (1998) The ubiquitin system. *Annual Review of Biochemistry* **67:** 425-479

Hieda M, Winstanley H, Maini P, Iborra FJ, Cook PR (2005) Different populations of RNA polymerase II in living mammalian cells. *Chromosome Research* **13:** 135-144

Höög G, Zarrizi R, Stedingk K, Jonsson K, Alvarado-Kristensson M (2011) Nuclear localization of -tubulin affects E2F transcriptional activity and S-phase progression. *The FASEB Journal* **25:** 3815-3827

Hoogstraten D, Nigg AL, Heath H, Mullenders LH, van Driel R, Hoeijmakers JH, Vermeulen W, Houtsmuller AB (2002) Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair in vivo. *Molecular Cell* **10:** 1163-1174

Horiuchi J, Silverman N, Marcus GA, Guarente L (1995) ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Molecular and Cellular Biology* **15:** 1203-1209

Horiuchi J, Silverman N, Pina B, Marcus GA, Guarente L (1997) ADA1, a novel component of the ADA/GCN5 complex, has broader effects than GCN5, ADA2, or ADA3. *Molecular and Cellular Biology* **17:** 3220-3228

Houtsmuller AB, Rademakers S, Nigg AL, Hoogstraten D, Hoeijmakers JH, Vermeulen W (1999) Action of DNA repair endonuclease ERCC1/XPF in living cells. *Science* **284:** 958-961

Huisinga KL, Pugh BF (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. *Molecular Cell* **13:** 573-585

Ihalainen TO, Willman SF, Niskanen EA, Paloheimo O, Smolander H, Laurila JP, Kaikkonen MU, Vihinen-Ranta M (2012) Distribution and dynamics of transcription-associated proteins during parvovirus infection. *Journal of Virology* **86:** 13779-13784

Imbalzano AN, Kwon H, Green MR, Kingston RE (1994a) Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature* **370:** 481-485

Imbalzano AN, Zaret KS, Kingston RE (1994b) Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *The Journal of Biological Chemistry* **269:** 8280-8286

Ingvarsdottir K, Krogan NJ, Emre NC, Wyce A, Thompson NJ, Emili A, Hughes TR, Greenblatt JF, Berger SL (2005) H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the Saccharomyces cerevisiae SAGA complex. *Molecular and Cellular Biology* **25:** 1162-1172

Ishikawa-Ankerhold HC, Ankerhold R, Drummen GP (2012) Advanced fluorescence microscopy techniques--FRAP, FLIP, FLAP, FRET and FLIM. *Molecules* **17:** 4047-4132

Iyer SP, Akimoto Y, Hart GW (2003) Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase. *The Journal of Biological Chemistry* **278:** 5399-5409

Jackson-Fisher AJ, Chitikila C, Mitra M, Pugh BF (1999) A role for TBP dimerization in preventing unregulated gene expression. *Molecular Cell* **3:** 717-727

Jacobson RH, Ladurner AG, King DS, Tjian R (2000) Structure and function of a human TAFII250 double bromodomain module. *Science* **288:** 1422-1425

Janicki SM, Tsukamoto T, Salghetti SE, Tansey WP, Sachidanandam R, Prasanth KV, Ried T, Shav-Tal Y, Bertrand E, Singer RH, Spector DL (2004) From silencing to gene expression: real-time analysis in single cells. *Cell* **116:** 683-698

Jansen A, Verstrepen KJ (2011) Nucleosome positioning in Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews* **75:** 301-320

Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Thérien C, Bergeron D, Bourassa S, Greenblatt J, Chabot B, Poirier GG, Hughes TR, Blanchette M, Price DH, Coulombe B (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Molecular Cell* **27:** 262-274

Johansen KM, Johansen J (2006) Regulation of chromatin structure by histone H3S10 phosphorylation. *Chromosome Research* **14:** 393-404

Johnsson A, Durand-Dubief M, Xue-Franzén Y, Rönnerblad M, Ekwall K, Wright A (2009) HAT-HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress. *EMBO Reports* **10:** 1009-1014

Jung I, Kim SK, Kim M, Han YM, Kim YS, Kim D, Lee D (2012) H2B monoubiquitylation is a 5'-enriched active transcription mark and correlates with exon-intron structure in human cells. *Genome Research* **22:** 1026-1035

Juven-Gershon T, Hsu J-YY, Theisen JW, Kadonaga JT (2008) The RNA polymerase II core promoter - the gateway to transcription. *Current Opinion in Cell Biology* **20:** 253-259

Kadosh D, Struhl K (1997) Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* **89:** 365-371

Kang J, Ferguson D, Song H, Bassing C, Eckersdorff M, Alt FW, Xu Y (2005) Functional interaction of H2AX, NBS1, and p53 in ATM-dependent DNA damage responses and tumor suppression. *Molecular and Cellular Biology* **25:** 661-670

Kang JJ, Auble DT, Ranish JA, Hahn S (1995) Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. *Molecular and Cellular Biology* **15:** 1234-1243

Katan-Khaykovich Y, Struhl K (2002) Dynamics of global histone acetylation and deacetylation in vivo: rapid restoration of normal histone acetylation status upon removal of activators and repressors. *Genes & Development* **16:** 743-752

Kelly WG, Dahmus ME, Hart GW (1993) RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *The Journal of Biological Chemistry* **268:** 10416-10424

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, Linder-Basso D, Plachetka A, Shanower G, Tolstorukov MY, Luquette LJ, Xi R, Jung YL, Park RW, Bishop EP, Canfield TK, Sandstrom R, Thurman RE, MacAlpine DM, Stamatoyannopoulos JA, Kellis M, Elgin SC, Kuroda MI, Pirrotta V, Karpen GH, Park PJ (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* **471:** 480-485

Khorasanizadeh S (2004) The nucleosome: from genomic organization to genomic regulation. *Cell* **116:** 259-272

Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, Shilatifard A, Muir TW, Roeder RG (2009) RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* **137:** 459-471

Kim JL, Burley SK (1994) 1.9 A resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nature Structural Biology* **1:** 638-653

Kim JL, Nikolov DB, Burley SK (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365:** 520-527

Kim YJ, Björklund S, Li Y, Sayre MH, Kornberg RD (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77:** 599-608

Kimura H, Sugaya K, Cook PR (2002a) The transcription cycle of RNA polymerase II in living cells. *The Journal of Cell Biology* **159:** 777-782

Kimura M, Suzuki H, Ishihama A (2002b) Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in Schizosaccharomyces pombe involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. *Molecular and Cellular Biology* **22:** 1577-1588

Kirschner DB, vom Baur E, Thibault C, Sanders SL, Gangloff YG, Davidson I, Weil PA, Tora L (2002) Distinct mutations in yeast TAF(II)25 differentially affect the composition of TFIID and SAGA complexes as well as global gene expression patterns. *Molecular and Cellular Biology* **22:** 3178-3193

Klein J, Nolden M, Sanders SL, Kirchner J, Weil PA, Melcher K (2003) Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA. *The Journal of Biological Chemistry* **278:** 6779-6786

Klejman MP, Zhao X, van Schaik FM, Herr W, Timmers HT (2005) Mutational analysis of BTAF1-TBP interaction: BTAF1 can rescue DNA-binding defective TBP mutants. *Nucleic Acids Research* **33:** 5426-5436

Kloster-Landsberg M, Herbomel G, Wang I, Derouard J, Vourc'h C, Usson Y, Souchier C, Delon A (2012) Cellular response to heat shock studied by multiconfocal fluorescence correlation spectroscopy. *Biophysical Journal* **103:** 1110-1119

Kobayashi N, Boyer TG, Berk AJ (1995) A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly. *Molecular and Cellular Biology* **15:** 6465-6473

Kobor MS, Simon LD, Omichinski J, Zhong G, Archambault J, Greenblatt J (2000) A motif shared by TFIIF and TFIIB mediates their interaction with the RNA polymerase II carboxy-terminal domain phosphatase Fcp1p in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **20:** 7438-7449

Köhler A, Pascual-García P, Llopis A, Zapater M, Posas F, Hurt E, Rodríguez-Navarro S (2006) The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Molecular Biology of the Cell* **17:** 4228-4236

Köhler A, Schneider M, Cabal GG, Nehrbass U, Hurt E (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nature Cell Biology* **10:** 707-715

Kokubo T, Gong DW, Yamashita S, Takada R, Roeder RG, Horikoshi M, Nakatani Y (1993) Molecular cloning, expression, and characterization of the *Drosophila* 85-kilodalton TFIID subunit. *Molecular and Cellular Biology* **13:** 7859-7863

Kokubo T, Swanson MJ, Nishikawa JI, Hinnebusch AG, Nakatani Y (1998) The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Molecular and Cellular Biology* **18:** 1003-1012

Koleske AJ, Young RA (1994) An RNA polymerase II holoenzyme responsive to activators. *Nature* **368:** 466-469

Komarnitsky P, Cho EJ, Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development* **14:** 2452-2460

Kornberg RD (1974) Chromatin structure: a repeating unit of histones and DNA. *Science* **184:** 868-871

Köster M, Frahm T, Hauser H (2005) Nucleocytoplasmic shuttling revealed by FRAP and FLIP technologies. *Current Opinion in Biotechnology* **16:** 28-34

Kostrewa D, Zeller ME, Armache K-JJ, Seizl M, Leike K, Thomm M, Cramer P (2009) RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462:** 323-330

Krebs AR, Demmers J, Karmodiya K, Chang N-CC, Chang AC, Tora L (2010) ATAC and Mediator coactivators form a stable complex and regulate a set of non-coding RNA genes. *EMBO Reports* **11:** 541-547

Krebs AR, Karmodiya K, Lindahl-Allen M, Struhl K, Tora L (2011) SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Molecular Cell* **44:** 410-423

Krieger W, Langowski J (2013): QuickFit 3.0: A data evaluation application for biophysics, *http://www.dkfz.de/Macromol/quickfit/*

Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Molecular Cell* **11:** 721-729

Kumar KP, Akoulitchev S, Reinberg D (1998) Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 9767-9772

Kurabe N, Katagiri K, Komiya Y, Ito R, Sugiyama A, Kawasaki Y, Tashiro F (2007) Deregulated expression of a novel component of TFTC/STAGA histone acetyltransferase complexes, rat SGF29, in hepatocellular carcinoma: possible implication for the oncogenic potential of c-Myc. *Oncogene* **26:** 5626-5634

Kusch T, Guelman S, Abmayr SM, Workman JL (2003) Two *Drosophila* Ada2 homologues function in different multiprotein complexes. *Molecular and Cellular Biology* **23:** 3305-3319

Lang G, Bonnet J, Umlauf D, Karmodiya K, Koffler J, Stierle M, Devys D, Tora L (2011) The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements. *Molecular and Cellular Biology* **31:** 3734-3744

Langelier MF, Forget D, Rojas A, Porlier Y, Burton ZF, Coulombe B (2001) Structural and functional interactions of transcription factor (TF) IIA with TFIIE and TFIIF in transcription initiation by RNA polymerase II. *The Journal of Biological Chemistry* **276:** 38652-38657

Laprade L, Rose D, Winston F (2007) Characterization of new Spt3 and TATA-binding protein mutants of Saccharomyces cerevisiae: Spt3 TBP allele-specific interactions and bypass of Spt8. *Genetics* **177:** 2007-2017

Larschan E, Winston F (2001) The S. cerevisiae SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. *Genes & Development* **15:** 1946-1956

Lee J-SS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, Shilatifard A (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* **131:** 1084-1096

Lee KK, Florens L, Swanson SK, Washburn MP, Workman JL (2005) The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Molecular and Cellular Biology* **25:** 1173-1182

Lee KK, Sardiu ME, Swanson SK, Gilmore JM, Torok M, Grant PA, Florens L, Workman JL, Washburn MP (2011) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Molecular Systems Biology* **7:** 503

Lee KK, Swanson SK, Florens L, Washburn MP, Workman JL (2009) Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. *Epigenetics & Chromatin* **2:** 2

Lee TI, Young RA (2000) Transcription of eukaryotic protein-coding genes. *Annual Review of Genetics* **34:** 77-137

Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NA, Margaritis T, van de Pasch LA, van Heesch SAA, Brok MO, Groot Koerkamp MJ, Ko CW, van Leenen D, Sameith K, van Hooff SR, Lijnzaad P, Kemmeren P, Hentrich T, Kobor MS, Buratowski S, Holstege FC (2011) The specificity and topology of chromatin interaction pathways in yeast. *Molecular Cell* **42:** 536-549

Lenstra TL, Holstege FC (2012) The discrepancy between chromatin factor location and effect. *Nucleus* **3:** 213-219

Lesca C, Germanier M, Raynaud-Messina B, Pichereaux C, Etievant C, Emond S, Burlet-Schiltz O, Monsarrat B, Wright M, Defais M (2005) DNA damage induce gamma-tubulin-RAD51 nuclear complexes in mammalian cells. *Oncogene* **24:** 5165-5172

Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* **128:** 707-719

Li XY, Virbasius A, Zhu X, Green MR (1999) Enhancement of TBP binding by activators and general transcription factors. *Nature* **399:** 605-609

Liu D, Ishima R, Tong KI, Bagby S, Kokubo T, Muhandiram DR, Kay LE, Nakatani Y, Ikura M (1998) Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell* **94:** 573-583

Liu X, Tesfai J, Evrard YA, Dent SY, Martinez E (2003) c-Myc transformation domain recruits the human STAGA complex and requires TRRAP and GCN5 acetylase activity for transcription activation. *The Journal of Biological Chemistry* **278:** 20405-20412

Liu Y, Peng L, Seto E, Huang S, Qiu Y (2012) Modulation of histone deacetylase 6 (HDAC6) nuclear import and tubulin deacetylase activity through acetylation. *The Journal of Biological Chemistry* **287:** 29168-29174

Lu H, Flores O, Weinmann R, Reinberg D (1991) The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 10004-10008

Lu H, Zawel L, Fisher L, Egly JM, Reinberg D (1992) Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358:** 641-645

Lusser A, Kadonaga JT (2003) Chromatin remodeling by ATP-dependent molecular machines. *BioEssays* **25:** 1192-1200

Macdonald N, Welburn JP, Noble ME, Nguyen A, Yaffe MB, Clynes D, Moggs JG, Orphanides G, Thomson S, Edmunds JW, Clayton AL, Endicott JA, Mahadevan LC (2005) Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Molecular Cell* **20:** 199-211

Magde D, Elson E, Webb WW (1972) Thermodynamic Fluctuations in a Reacting System—Measurement by Fluorescence Correlation Spectroscopy. *Physical Review Letters* **29:** 705-708

Malik S, Roeder RG (2005) Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends in Biochemical Sciences* **30:** 256-263

Marcus GA, Horiuchi J, Silverman N, Guarente L (1996) ADA5/SPT20 links the ADA and SPT genes, which are involved in yeast transcription. *Molecular and Cellular Biology* **16:** 3197-3205

Martinez E, Chiang CM, Ge H, Roeder RG (1994) TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. *The EMBO Journal* **13:** 3115-3126

Martinez E, Kundu TK, Fu J, Roeder RG (1998) A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. *The Journal of Biological Chemistry* **273:** 23781-23785

Maruvada P, Baumann CT, Hager GL, Yen PM (2003) Dynamic shuttling and intranuclear mobility of nuclear hormone receptors. *The Journal of Biological Chemistry* **278:** 12425-12432

Maxon ME, Goodrich JA, Tjian R (1994) Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. *Genes & Development* **8:** 515-524

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

McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453:** 812-816

McMahon SB, Van Buskirk HA, Dugan KA, Copeland TD, Cole MD (1998) The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell* **94:** 363-374

McMahon SJ, Pray-Grant MG, Schieltz D, Yates JR, Grant PA (2005) Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 8478-8482

McNally JG, Müller WG, Walker D, Wolford R, Hager GL (2000) The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287:** 1262-1265

Mengus G, May M, Jacq X, Staub A, Tora L, Chambon P, Davidson I (1995) Cloning and characterization of hTAFII18, hTAFII20 and hTAFII28: three subunits of the human transcription factor TFIID. *The EMBO Journal* **14:** 1520-1531

Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M (2008) Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nature Cell Biology* **10:** 483-488

Mirza S, Katafiasz BJ, Kumar R, Wang J, Mohibi S, Jain S, Gurumurthy CB, Pandita TK, Dave BJ, Band H, Band V (2012) Alteration/deficiency in activation-3 (Ada3) plays a critical role in maintaining genomic stability. *Cell Cycle* **11:** 4266-4274

Mirza S, Rakha EA, Alshareeda A, Mohibi S, Zhao X, Katafiasz BJ, Wang J, Gurumurthy CB, Bele A, Ellis IO, Green AR, Band H, Band V (2013) Cytoplasmic localization of alteration/deficiency in activation 3 (ADA3) predicts poor clinical outcome in breast cancer patients. *Breast Cancer Research and Treatment* **137:** 721-731

Mohibullah N, Hahn S (2008) Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. *Genes & Development* **22:** 2994-3006

Moncollin V, Fischer L, Cavallini B, Egly JM, Chambon P (1992) Class II (B) general transcription factor (TFIIB) that binds to the template-committed preinitiation complex is different from general transcription factor BTF3. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 397-401

Mosley AL, Pattenden SG, Carey M, Venkatesh S, Gilmore JM, Florens L, Workman JL, Washburn MP (2009) Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Molecular Cell* **34:** 168-178

Mujtaba S, Zeng L, Zhou MMM (2007) Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26:** 5521-5527

Muller F, Tora L (2014) Chromatin and DNA sequences in defining promoters for transcription initiation. *Biochimica et Biophysica Acta* **1839:** 118-128

Müller F, Zaucker A, Tora L (2010) Developmental regulation of transcription initiation: more than just changing the actors. *Current Opinion in Genetics & Development* **20:** 533-540

Muratoglu S, Georgieva S, Pápai G, Scheer E, Enünlü I, Komonyi O, Cserpán I, Lebedeva L, Nabirochkina E, Udvardy A, Tora L, Boros I (2003) Two different *Drosophila* ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Molecular and Cellular Biology* **23:** 306-321

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K (2007) RNA polymerase is poised for activation across the genome. *Nature Genetics* **39:** 1507-1511

Näär AM, Lemon BD, Tjian R (2001) Transcriptional coactivator complexes. *Annual Review of Biochemistry* **70:** 475-501

Nag A, Germaniuk-Kurowska A, Dimri M, Sassack MA, Gurumurthy CB, Gao Q, Dimri G, Band H, Band V (2007) An essential role of human Ada3 in p53 acetylation. *The Journal of Biological Chemistry* **282:** 8812-8820

Nagy Z, Riss A, Fujiyama S, Krebs A, Orpinell M, Jansen P, Cohen A, Stunnenberg HG, Kato S, Tora L (2010) The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cellular and Molecular Life Sciences* **67:** 611-628

Nagy Z, Riss A, Romier C, le Guezennec X, Dongre AR, Orpinell M, Han J, Stunnenberg H, Tora L (2009) The human SPT20-containing SAGA complex plays a direct role in the regulation of endoplasmic reticulum stress-induced genes. *Molecular and Cellular Biology* **29:** 1649-1660

Nagy Z, Tora L (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26:** 5341-5357

Nechaev S, Adelman K (2011) Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochimica et Biophysica Acta* **1809:** 34-45

Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327:** 335-338

Ng H, Dole S, Struhl K (2003a) The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *Journal of Biological Chemistry* **278:** 33625-33628

Ng HH, Robert F, Young RA, Struhl K (2003b) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular Cell* **11:** 709-719

Ng SS, Yue WW, Oppermann U, Klose RJ (2009) Dynamic protein methylation in chromatin biology. *Cellular and Molecular Life Sciences* **66:** 407-422

Nguyen VT, Giannoni F, Dubois MF, Seo SJ, Vigneron M, Kédinger C, Bensaude O (1996) In vivo degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. *Nucleic Acids Research* **24:** 2924-2929

Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK (1996) Crystal structure of a human TATA box-binding protein/TATA element complex. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 4862-4867

Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK, Roeder RG, Burley SK (1995) Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377:** 119-128

Nissim-Rafinia M, Meshorer E (2011) Photobleaching assays (FRAP & FLIP) to measure chromatin protein dynamics in living embryonic stem cells. *Journal of Visualized Experiments*

Ogryzko VV, Kotani T, Zhang X, Schiltz RL, Howard T, Yang XJ, Howard BH, Qin J, Nakatani Y (1998) Histone-like TAFs within the PCAF histone acetylase complex. *Cell* **94:** 35-44

Ohkuma Y, Roeder RG (1994) Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368:** 160-163

Ohkuma Y, Sumimoto H, Hoffmann A, Shimasaki S, Horikoshi M, Roeder RG (1991) Structural motifs and potential sigma homologies in the large subunit of human general transcription factor TFIIE. *Nature* **354:** 398-401

Oki M, Aihara H, Ito T (2007) Role of histone phosphorylation in chromatin dynamics and its implications in diseases. *Subcellular Biochemistry* **41:** 319-336

Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92:** 105-116

Orpinell M, Fournier M, Riss A, Nagy Z, Krebs AR, Frontini M, Tora L (2010) The ATAC acetyl transferase complex controls mitotic progression by targeting non-histone substrates. *The EMBO Journal* **29:** 2381-2394

Palhan VB, Chen S, Peng G-HH, Tjernberg A, Gamper AM, Fan Y, Chait BT, La Spada AR, Roeder RG (2005) Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 8472-8477

Pankotai T, Komonyi O, Bodai L, Ujfaludi Z, Muratoglu S, Ciurciu A, Tora L, Szabad J, Boros I (2005) The homologous *Drosophila* transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions. *Molecular and Cellular Biology* **25:** 8215-8227

Paoletti AC, Parmely TJ, Tomomori-Sato C, Sato S, Zhu D, Conaway RC, Conaway JW, Florens L, Washburn MP (2006) Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 18928-18933

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D (2006) Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125:** 703-717

Pereira LA, van der Knaap JA, van den Boom V, van den Heuvel FA, Timmers HT (2001) TAF(II)170 interacts with the concave surface of TATA-binding protein to inhibit its DNA binding activity. *Molecular and Cellular Biology* **21:** 7523-7534

Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Molecular Cell* **23:** 297-305

Phair RD, Misteli T (2000) High mobility of proteins in the mammalian cell nucleus. *Nature* **404:** 604-609

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122:** 517-527

Pollard KJ, Peterson CL (1997) Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression. *Molecular and Cellular Biology* **17:** 6212-6222

Potma EO, de Boeij WP, Bosgraaf L, Roelofs J, van Haastert PJ, Wiersma DA (2001) Reduced protein diffusion rate by cytoskeleton in vegetative and polarized dictyostelium cells. *Biophysical Journal* **81:** 2010-2019

Pray-Grant MG, Schieltz D, McMahon SJ, Wood JM, Kennedy EL, Cook RG, Workman JL, Yates JR, Grant PA (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Molecular and Cellular Biology* **22:** 8774-8786

Qian C, Zhang Q, Li S, Zeng L, Walsh MJ, Zhou M-MM (2005) Structure and chromosomal DNA binding of the SWIRM domain. *Nature Structural & Molecular Biology* **12:** 1078-1085

Rafalska-Metcalf IU, Powers SL, Joo LM, LeRoy G, Janicki SM (2010) Single cell analysis of transcriptional activation dynamics. *PloS One* **5**

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA (2010) c-Myc regulates transcriptional pause release. *Cell* **141:** 432-445

Rapsomaniki MA, Kotsantis P, Symeonidou I-EE, Giakoumakis N-NN, Taraviras S, Lygerou Z (2012) easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. *Bioinformatics* **28:** 1800-1801

Razin SV, Gavrilov AA (2014) Chromatin without the 30-nm fiber: constrained disorder instead of hierarchical folding. *Epigenetics* **9:** 653-657

Reeves WM, Hahn S (2005) Targets of the Gal4 transcription activator in functional transcription complexes. *Molecular and Cellular Biology* **25:** 9092-9102

Reinberg D, Roeder RG (1987) Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. *The Journal of Biological Chemistry* **262:** 3310-3321

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

Robert F, Douziech M, Forget D, Egly JM, Greenblatt J, Burton ZF, Coulombe B (1998) Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Molecular Cell* **2:** 341-351

Roberts SM, Winston F (1996) SPT20/ADA5 encodes a novel protein functionally related to the TATA-binding protein and important for transcription in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **16:** 3206-3213

Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *The Journal of Biological Chemistry* **273:** 5858-5868

Roh TY, Cuddapah S, Zhao K (2005) Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes & Development* **19:** 542-552

Rojo-Niersbach E, Furukawa T, Tanese N (1999) Genetic dissection of hTAF(II)130 defines a hydrophobic surface required for interaction with glutamine-rich activators. *The Journal of Biological Chemistry* **274:** 33778-33784

Rosado-Lugo JD, Hampsey M (2014) The Ssu72 Phosphatase Mediates the RNA Polymerase II Initiation-Elongation Transition. *The Journal of Biological Chemistry*

Rossetto D, Avvakumov N, Cote J (2012) Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics* **7:** 1098-1108

Roth SY, Denu JM, Allis CD (2001) Histone acetyltransferases. *Annual Review of Biochemistry* **70:** 81-120

Roudier F, Ahmed I, Bérard C, Sarazin A, Huard T, Cortijo S, Bouyer D, Caillieux E, Berthet E, Shikhley L (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *The EMBO Journal* **30:** 1928-1938

Roy R, Adamczewski JP, Seroz T, Vermeulen W, Tassan JP, Schaeffer L, Nigg EA, Hoeijmakers JH, Egly JM (1994) The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell* **79:** 1093-1101

Ruppert S, Tjian R (1995) Human TAFII250 interacts with RAP74: implications for RNA polymerase II initiation. *Genes & Development* **9:** 2747-2755

Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nature Reviews Genetics* **8:** 424-436

Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA (2002) Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Molecular and Cellular Biology* **22:** 4723-4738

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

Schram AW, Baas R, Jansen PW, Riss A, Tora L, Vermeulen M, Timmers HT (2013) A dual role for SAGA-associated factor 29 (SGF29) in ER stress survival by coordination of both histone H3 acetylation and histone H3 lysine-4 trimethylation. *PloS One* **8**

Selleck W, Howley R, Fang Q, Podolny V, Fried MG, Buratowski S, Tan S (2001) A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. *Nature Structural Biology* **8:** 695-700

Sengupta P, Garai K, Balaji J, Periasamy N, Maiti S (2003) Measuring size distribution in highly heterogeneous systems with fluorescence correlation spectroscopy. *Biophysical Journal* **84:** 1977-1984

Sengupta P, Garai K, Balaji J, Periasamy N, Maiti S (2003) Measuring size distribution in highly heterogeneous systems with fluorescence correlation spectroscopy. *Biophysical Journal* **84:** 1977-1984

Serizawa H, Conaway JW, Conaway RC (1993) Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. *Nature* **363:** 371-374

Sermwittayawong D, Tan S (2006) SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. *The EMBO Journal* **25:** 3791-3800

Shav-Tal Y, Darzacq X, Shenoy SM, Fusco D, Janicki SM, Spector DL, Singer RH (2004) Dynamics of single mRNPs in nuclei of living cells. *Science* **304:** 1797-1800

Shilatifard A (2008) Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current Opinion in Cell Biology* **20:** 341-348

Shogren-Knaak M, Ishii H, Sun J-MM, Pazin MJ, Davie JR, Peterson CL (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311:** 844-847

Simpson RT (1978) Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry* **17:** 5524-5531

Smith KP, Byron M, Clemson CM, Lawrence JB (2004) Ubiquitinated proteins including uH2A on the human and mouse inactive X chromosome: enrichment in gene rich bands. *Chromosoma* **113:** 324-335

Sobel RE, Cook RG, Perry CA, Annunziato AT, Allis CD (1995) Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 1237-1241

Solow S, Salunek M, Ryan R, Lieberman PM (2001) Taf(II) 250 phosphorylates human transcription factor IIA on serine residues important for TBP binding and transcription activity. *The Journal of Biological Chemistry* **276:** 15886-15892

Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu RM, Zhu P, Li G (2014) Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344:** 376-380

Soutoglou E, Demény MA, Scheer E, Fienga G, Sassone-Corsi P, Tora L (2005) The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. *Molecular and Cellular Biology* **25:** 4092-4104

Spedale G, Mischerikow N, Heck AJR, Timmers MHT, Pijnappel WWM (2010) Identification of Pep4p as the protease responsible for formation of the SAGA-related SLIK protein complex. *Journal of Biological Chemistry* **285:** 22793-22799

Spedale G, Timmers HT, Pijnappel WW (2012) ATAC-king the complexity of SAGA during evolution. *Genes & Development* **26:** 527-541

Sprague BL, McNally JG (2005) FRAP analysis of binding: proper and fitting. *Trends in Cell Biology* **15:** 84-91

Sprouse RO, Karpova TS, Mueller F, Dasgupta A, McNally JG, Auble DT (2008) Regulation of TATA-binding protein dynamics in living yeast cells. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 13304-13308

Stasevich TJ, Mueller F, Michelman-Ribeiro A, Rosales T, Knutson JR, McNally JG (2010) Cross-validating FRAP and FCS to quantify the impact of photobleaching on in vivo binding estimates. *Biophysical Journal* **99:** 3093-3101

Steffen PA, Ringrose L (2014) What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nature reviews Molecular Cell biology* **15:** 340-356

Sterner DE, Grant PA, Roberts SM, Duggan LJ, Belotserkovskaya R, Pacella LA, Winston F, Workman JL, Berger SL (1999) Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Molecular and Cellular Biology* **19:** 86-98

Sterner DE, Wang X, Bloom MH, Simon GM, Berger SL (2002) The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex. *The Journal of Biological Chemistry* **277:** 8178-8186

Steward MM, Lee J-SS, O'Donovan A, Wyatt M, Bernstein BE, Shilatifard A (2006) Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature Structural & Molecular Biology* **13:** 852-854

Suganuma T, Gutiérrez JL, Li B, Florens L, Swanson SK, Washburn MP, Abmayr SM, Workman JL (2008) ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nature Structural & Molecular Biology* **15:** 364-372

Suganuma T, Workman JL (2011) Signals and combinatorial functions of histone modifications. *Annual Review of Biochemistry* **80:** 473-499

Sugaya K, Vigneron M, Cook PR (2000) Mammalian cell lines expressing functional RNA polymerase II tagged with the green fluorescent protein. *Journal of Cell Science* **113 ( Pt 15):** 2679-2683

Sun M, Schwalb B, Pirkl N, Maier KC, Schenk A, Failmezger H, Tresch A, Cramer P (2013) Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. *Molecular Cell* **52:** 52-62

Sun ZW, Allis CD (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418:** 104-108

Taatjes DJ (2010) The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends in Biochemical Sciences* **35:** 315-322

Tan Y, Xue Y, Song C, Grunstein M (2013) Acetylated histone H3K56 interacts with Oct4 to promote mouse embryonic stem cell pluripotency. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 11493-11498

Taunton J, Hassig CA, Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272:** 408-411

Thevenet L, Mejean C, Moniot B, Bonneaud N, Galeotti N, Aldrian-Herrada G, Poulat F, Berta P, Benkirane M, Boizet-Bonhoure B (2004) Regulation of human SRY subcellular distribution by its acetylation/deacetylation. *The EMBO Journal* **23:** 3336-3345

Thoma F, Koller T, Klug A (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *The Journal of Cell Biology* **83:** 403-427

Thomas-Claudepierre A-SS, Schiavo E, Heyer V, Fournier M, Page A, Robert I, Reina-San-Martin B (2013) The cohesin complex regulates immunoglobulin class switch recombination. *The Journal of experimental medicine* **210:** 2495-2502

Thomas MC, Chiang C-MM (2006) The general transcription machinery and general cofactors. *Critical Reviews in Biochemistry and Molecular Biology* **41:** 105-178

Tjeertes JV, Miller KM, Jackson SP (2009) Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *The EMBO Journal* **28:** 1878-1889

Tochio N, Umehara T, Koshiba S, Inoue M, Yabuki T, Aoki M, Seki E, Watanabe S, Tomo Y, Hanada M, Ikari M, Sato M, Terada T, Nagase T, Ohara O, Shirouzu M, Tanaka A, Kigawa T, Yokoyama S (2006) Solution structure of the SWIRM domain of human histone demethylase LSD1. *Structure* **14:** 457-468

Tsien RY (1998) The green fluorescent protein. *Annual Review of Biochemistry* **67:** 509-544

Tsukamoto T, Hashiguchi N, Janicki SM, Tumbar T, Belmont AS, Spector DL (2000) Visualization of gene activity in living cells. *Nature Cell Biology* **2:** 871-878

Umlauf D, Bonnet J, Waharte F, Fournier M, Stierle M, Fischer B, Brino L, Devys D, Tora L (2013) The human TREX-2 complex is stably associated with the nuclear pore basket. *Journal of Cell Science* **126:** 2656-2667

Vamos EE, Boros IM (2012) The C-terminal domains of ADA2 proteins determine selective incorporation into GCN5-containing complexes that target histone H3 or H4 for acetylation. *FEBS Letters* **586:** 3279-3286

Van Dyke MW, Roeder RG, Sawadogo M (1988) Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* **241:** 1335-1338

Van Dyke MW, Sawadogo M, Roeder RG (1989) Stability of transcription complexes on class II genes. *Molecular and Cellular Biology* **9:** 342-344

Van Hooser A, Goodrich DW, Allis CD, Brinkley BR, Mancini MA (1998) Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation. *Journal of Cell Science* **111 ( Pt 23):** 3497-3506

van Royen ME, Zotter A, Ibrahim SM, Geverts B, Houtsmuller AB (2011) Nuclear proteins: finding and binding target sites in chromatin. *Chromosome Research* **19:** 83-98

Värv S, Kristjuhan K, Peil K, Lõoke M, Mahlakõiv T, Paapsi K, Kristjuhan A (2010) Acetylation of H3 K56 is required for RNA polymerase II transcript elongation through heterochromatin in yeast. *Molecular and Cellular Biology* **30:** 1467-1477

Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamon AJ, Jain P, Rolleri NS, Jiang C, Hemeryck-Walsh C, Pugh BF (2011) A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. *Molecular Cell* **41:** 480-492

Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142:** 967-980

Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131:** 58-69

Verrijzer CP, Tjian R (1996) TAFs mediate transcriptional activation and promoter selectivity. *Trends in Biochemical Sciences* **21:** 338-342

Vettese-Dadey M, Grant PA, Hebbes TR, Crane- Robinson C, Allis CD, Workman JL (1996) Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *The EMBO Journal* **15:** 2508-2518

Vitaliano-Prunier A, Menant A, Hobeika M, Geli V, Gwizdek C, Dargemont C (2008) Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. *Nature Cell Biology* **10:** 1365-1371

Vogelauer M, Wu J, Suka N, Grunstein M (2000) Global histone acetylation and deacetylation in yeast. *Nature* **408:** 495-498

vom Baur E, Harbers M, Um SJ, Benecke A, Chambon P, Losson R (1998) The yeast Ada complex mediates the ligand-dependent activation function AF-2 of retinoid X and estrogen receptors. *Genes & Development* **12:** 1278-1289

Voss TC, Hager GL (2014) Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nature Reviews Genetics* **15:** 69-81

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431:** 873-878

Wang Y-LL, Faiola F, Xu M, Pan S, Martinez E (2008a) Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. *The Journal of Biological Chemistry* **283:** 33808-33815

Wang Z, Cui B, Gorovsky MA (2009) Histone H2B ubiquitylation is not required for histone H3 methylation at lysine 4 in tetrahymena. *The Journal of Biological Chemistry* **284:** 34870-34879

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K (2008b) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics* **40:** 897-903

Warfield L, Ranish JA, Hahn S (2004) Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. *Genes & Development* **18:** 1022-1034

Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **19:** 242-247

Watanabe T, Hayashi K, Tanaka A, Furumoto T, Hanaoka F, Ohkuma Y (2003) The carboxy terminus of the small subunit of TFIIE regulates the transition from transcription initiation to elongation by RNA polymerase II. *Molecular and Cellular Biology* **23:** 2914-2926

Weake VM, Dyer JO, Seidel C, Box A, Swanson SK, Peak A, Florens L, Washburn MP, Abmayr SM, Workman JL (2011) Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation. *Genes & Development* **25:** 1499-1509

Weake VM, Workman JL (2012) SAGA function in tissue-specific gene expression. *Trends in Cell Biology* **22:** 177-184

Weil PA, Luse DS, Segall J, Roeder RG (1979) Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* **18:** 469-484

Welihinda AA, Tirasophon W, Green SR, Kaufman RJ (1997) Gene induction in response to unfolded protein in the endoplasmic reticulum is mediated through Ire1p kinase interaction with a transcriptional coactivator complex containing Ada5p. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 4289-4294

Welihinda AA, Tirasophon W, Kaufman RJ (2000) The transcriptional co-activator ADA5 is required for HAC1 mRNA processing in vivo. *The Journal of Biological Chemistry* **275:** 3377-3381

Wieczorek E, Brand M, Jacq X, Tora L (1998) Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* **393:** 187-191

Wild T, Cramer P (2012) Biogenesis of multisubunit RNA polymerases. *Trends in Biochemical Sciences* **37:** 99-105

Williams SK, Truong D, Tyler JK (2008) Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 9000-9005

Wood A, Schneider J, Dover J, Johnston M, Shilatifard A (2003) The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *The Journal of Biological Chemistry* **278:** 34739-34742

Woudstra EC, Gilbert C, Fellows J, Jansen L, Brouwer J, Erdjument-Bromage H, Tempst P, Svejstrup JQ (2002) A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature* **415:** 929-933

Woychik NA, Liao SM, Kolodziej PA, Young RA (1990) Subunits shared by eukaryotic nuclear RNA polymerases. *Genes & Development* **4:** 313-323

Wu P-YJY, Ruhlmann C, Winston F, Schultz P (2004) Molecular architecture of the S. cerevisiae SAGA complex. *Molecular Cell* **15:** 199-208

Wu PY, Winston F (2002) Analysis of Spt7 function in the Saccharomyces cerevisiae SAGA coactivator complex. *Molecular and Cellular Biology* **22:** 5367-5379

Wu SY, Chiang CM (1998) Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro. *The Journal of Biological Chemistry* **273:** 12492-12498

Wu SY, Chiang CM (2001) TATA-binding protein-associated factors enhance the recruitment of RNA polymerase II by transcriptional activators. *The Journal of Biological Chemistry* **276:** 34235-34243

Wyce A, Xiao T, Whelan KA, Kosman C, Walter W, Eick D, Hughes TR, Krogan NJ, Strahl BD, Berger SL (2007) H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Molecular Cell* **27:** 275-288

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442:** 86-90

Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, Nakatani Y (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382:** 319-324

Yang XJJ, Seto E (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* **26:** 5310-5318

Yao J, Munson KM, Webb WW, Lis JT (2006) Dynamics of heat shock factor association with native gene loci in living cells. *Nature* **442:** 1050-1053

Yokomori K, Admon A, Goodrich JA, Chen JL, Tjian R (1993) *Drosophila* TFIIA-L is processed into two subunits that are associated with the TBP/TAF complex. *Genes & Development* **7:** 2235-2245

Yoneyama M, Tochio N, Umehara T, Koshiba S, Inoue M, Yabuki T, Aoki M, Seki E, Matsuda T, Watanabe S, Tomo Y, Nishimura Y, Harada T, Terada T, Shirouzu M, Hayashizaki Y, Ohara O, Tanaka A, Kigawa T, Yokoyama S (2007) Structural and functional differences of SWIRM domain subtypes. *Journal of Molecular Biology* **369:** 222-238

Young RA (1991) RNA polymerase II. *Annual Review of Biochemistry* **60:** 689-715

Zhang XY, Varthi M, Sykes SM, Phillips C, Warzecha C, Zhu W, Wyce A, Thorne AW, Berger SL, McMahon SB (2008) The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Molecular Cell* **29:** 102-111

Zhao X, Herr W (2002) A regulated two-step mechanism of TBP binding to DNA: a solvent-exposed surface of TBP inhibits TATA box recognition. *Cell* **108:** 615-627

Zhao Y, Lang G, Ito S, Bonnet J, Metzger E, Sawatsubashi S, Suzuki E, Le Guezennec X, Stunnenberg HG, Krasnov A, Georgieva SG, Schüle R, Takeyama K-I, Kato S, Tora L, Devys D (2008) A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Molecular Cell* **29:** 92-101

Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, Rosenfeld MG (2008) Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Molecular Cell* **29:** 69-80

Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP (2006) Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *Journal of Proteome Research* **5:** 2339-2347

**Nikolaos VOSNAKIS**

# "Investigating the role of human HAT (histone acetyltransferase) containing complexes, ATAC and SAGA, in living cells"

## Résumé

Les complexes acétyltransférases (HAT), SAGA et ATAC, sont des régulateurs de la transcription des gènes. Cependant, peu d'études ont été menées sur la dynamique de ces complexes au niveau cellulaire et sur les mécanismes régulant leur assemblage. Au cours de mes travaux de thèse, j'ai utilisé des approches d'imagerie sur cellules vivantes, afin de déterminer la mobilité de ces complexes en comparaison avec celle d'autres régulateurs transcriptionnels. Les résultats ont montré que les sous-unités de SAGA et ATAC intéragissent de manière transiente avec la chromatine. En complément, nous avons montré que les sous-unités spécifiques de SAGA et ATAC (ADA2b et ADA2a) ont des propriétés dynamique intracellulaire différentes et que GCN5, affecte la distribution d'ADA2a. Des analyses protéomique menées sur le comportement de ces protéines au niveau endogène, ont permis de montrer que les voies d'assemblage de ces deux complexes étaient différentes au niveau cytoplasmique et nucléaire.

## Résumé en anglais

Human SAGA and ATAC, are histone acetyltransferase (HAT) containing complexes that share a set of subunits and facilitate RNA polymerase II (Pol II) transcription. Little is known for the dynamics of the complexes in living cells and the regulation of their assembly. In this work, we used live-cell imaging to characterise the mobility of the two complexes and compare it with other actors of Pol II transcription. All tested ATAC and SAGA subunits exhibit very transient interactions with chromatin, a property that explains certain aspects of the function of the complexes. Moreover, we showed that overexpressed ATAC- and SAGA-specific HAT-module subunits (ADA2a and ADA2b respectively) have different intracellular dynamics and that the abundance of the shared subunit GCN5, affects the distribution of ADA2a. Quantitative proteomic analysis expanded our findings on endogenous proteins and provided evidence that the cytoplasmic and nuclear assembly pathways of SAGA and ATAC are different.