



HAL
open science

What do the structures of GCN5-containing complexes teach us about their function?

Dominique Helmlinger, Gábor Papai, Didier Devys, László Tora

► **To cite this version:**

Dominique Helmlinger, Gábor Papai, Didier Devys, László Tora. What do the structures of GCN5-containing complexes teach us about their function?. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 2020, pp.194614. 10.1016/j.bbagr.2020.194614 . hal-02990243v2

HAL Id: hal-02990243

<https://hal.science/hal-02990243v2>

Submitted on 5 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Contents lists available at ScienceDirect

BBA - Gene Regulatory Mechanisms

journal homepage: www.elsevier.com/locate/bbagrm

Review

What do the structures of GCN5-containing complexes teach us about their function?☆

Dominique Helmlinger^a, Gábor Papai^{b,c,d,e}, Didier Devys^{b,c,d,e,*}, László Tora^{b,c,d,e,*}^a CRBM, University of Montpellier, CNRS, Montpellier, France^b Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67404 Illkirch, France^c Centre National de la Recherche Scientifique (CNRS), UMR7104, 67404 Illkirch, France^d Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, 67404 Illkirch, France^e Université de Strasbourg, 67404 Illkirch, France

ARTICLE INFO

Keywords:

Transcription
Chromatin
KAT2A
KAT2B
ADA complex
SAGA complex
ADA two A containing (ATAC) complex
Modular organization
Histone acetyl transferase
Histone deubiquitylase

ABSTRACT

Transcription initiation is a major regulatory step in eukaryotic gene expression. It involves the assembly of general transcription factors and RNA polymerase II into a functional pre-initiation complex at core promoters. The degree of chromatin compaction controls the accessibility of the transcription machinery to template DNA. Co-activators have critical roles in this process by actively regulating chromatin accessibility. Many transcriptional coactivators are multisubunit complexes, organized into distinct structural and functional modules and carrying multiple regulatory activities. The first nuclear histone acetyltransferase (HAT) characterized was General Control Non-derepressible 5 (Gcn5). Gcn5 was subsequently identified as a subunit of the HAT module of the Spt-Ada-Gcn5-acetyltransferase (SAGA) complex, which is an experimental paradigm for multifunctional co-activators. We know today that Gcn5 is the catalytic subunit of multiple distinct co-activator complexes with specific functions. In this review, we summarize recent advances in the structure of Gcn5-containing co-activator complexes, most notably SAGA, and discuss how these new structural insights contribute to better understand their functions. This article is part of a Special Issue entitled: Gcn5: the quintessential histone acetyltransferase.

1. Introduction

Transcription initiation is a major regulatory step in eukaryotic gene expression. The accessibility of the transcription machinery to template DNA is controlled by the degree of chromatin compaction. Co-activators have critical roles in this process, bridging DNA-bound transcription factors to the general transcription machinery, while actively regulating chromatin accessibility. Many transcriptional coactivators are multisubunit complexes carrying multiple regulatory activities, including histone modification, ATP-dependent nucleosome remodeling or incorporation of histone variants. Combining different activities into a single entity may be important to coordinate their activities during the transcription cycle. Genetic, biochemical and structural studies consistently reported that the activities of coactivators are carried by distinct structural and functional modules. A given module can be shared between different complexes, which can therefore be considered as specific assemblies of individual modules, selected during evolution for their role in chromatin regulation. The shared subunits have very

different properties (structural, chromatin readers or writers...), raising questions on the advantages of their partition in different complexes and on the mechanisms controlling the biogenesis of these specific assemblies.

The first nuclear histone acetyltransferase (HAT) identified was General Control Non-derepressible 5 (Gcn5) [1,2]. Gcn5 was rapidly shown to function as a subunit of the Spt-Ada-Gcn5-acetyltransferase (SAGA) complex, which is a paradigmatic example of such multifunctional co-activators. We now know that Gcn5 is the catalytic component of multiple co-activator complexes with distinct functions. In this review, we summarize recent advances in our understanding of the structure/function relationship of these GCN5-containing co-activator complexes.

☆ This article is part of a Special Issue entitled: Gcn5: the quintessential HAT edited by Dr. Vikki Weake.

* Corresponding authors at: Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67404 Illkirch, France.

E-mail addresses: devys@igbmc.fr (D. Devys), laszlo@igbmc.fr (L. Tora).

<https://doi.org/10.1016/j.bbagrm.2020.194614>

Received 15 May 2020; Received in revised form 28 July 2020; Accepted 29 July 2020

1874-9399/ © 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1

Summary of SAGA subunit composition across eukaryotic species. Shown are the names of each subunit from *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Homo sapiens*, and *Arabidopsis thaliana*, according to official nomenclatures. Names in brackets are alternative, commonly used names. Paralogous subunits are separated with a “/” sign and are indicated if they have been described as part of SAGA. In most cases, paralogous subunits are mutually exclusive within SAGA complexes. Individual module subunits are colored according to the code used in Fig. 1. “-”: no ortholog detected. “X”: ortholog exists but is not present in SAGA. [#]It is unclear whether HAF1 and HAF2 are *bona fide* SAGA components and genuine orthologs of Spt7 in plants. HFD: histone fold domain. ^{*}The TBP binding module comprises both Spt3 and Spt8 in *S. cerevisiae*, but presumably only Spt3 in plants and animals. ^oBased on HFD conservation it has been suggested that most plants have rather a TAF13 ortholog, than a SPT3 ortholog [132].

Functional tools	Orthologous SAGA complexes			
	<i>S. cerevisiae</i>	<i>D. melanogaster</i>	<i>H. sapiens</i>	<i>A. thaliana</i>
HAT module	Gcn5	KAT2 (GCN5)	KAT2A/KAT2B (GCN5/PCAF)	GCN5
	Ada2	Ada2b	TADA2b	ADA2b
	Ngg1 (Ada3)	Ada3	TADA3	ADA3
	Sgf29	Sgf29	SGF29	SGF29a/SGF29b
DUB module	Ubp8	dNonstop	USP22 (UBP22)	UBP22
	Sgf11	dSgf11	ATXN7L3	SGF11
	Sgf73	dATXN7	ATXN7/ATXN7L1/L2	-
	Sus1	dE(y)2	ENY2	ENY2
Core structural module	Taf5	WDA/TAF5L	TAF5L	TAF5
	Taf6	SAF6/TAF6L	TAF6L	TAF6a/TAF6b
	Taf9	TAF9	TAF9/TAF9b	TAF9
	Taf10	TAF10b	TAF10	TAF10
	Taf12	TAF12	TAF12	TAF12/TAF12b
	Spt7	Spt7	SUPT7L (STAF65G)	HAF1/HAF2 [#]
	Hfi1 (Ada1)	Ada1	TADA1	ADA1a/ADA1b
	Spt20	Spt20	SUPT20H	SPT20
TBP binding [*]	Spt3	Spt3	SUPT3H	TAF13 ^o
	Spt8	-	-	-
TF binding module	Tra1	Nipped-A (Tra1)	TRRAP	TRA1a/TRA1b
Splicing module	X	SF3B3	SF3B3	X
	X	SF3B5	SF3B5	X

2. SAGA complex

2.1. Modular organization of the multifunctional SAGA complex

Following its discovery as a HAT, the yeast Gcn5 protein was shown to be part of two multiprotein complexes called SAGA and ADA [3–5]. These two complexes contain several Ada proteins, which were originally identified in a yeast screen for adaptor proteins that facilitate transcription activation by acidic activation domains [6–9]. Both SAGA and ADA complexes share Ada2, Ada3 and Gcn5 (Ada4), which were

known to interact physically [10,11]. In contrast, Ada1 and Ada5 (Spt20) are only found in SAGA. Both Ada2 and Ada3 were shown to be required for the HAT activity of Gcn5 and mutations in the corresponding genes cause similar phenotypes [12]. The association of these three proteins confers Gcn5 the ability to acetylate nucleosomal histones, whereas Gcn5 alone acetylates only free histones [2,3,13]. Later work identified Sgf29, a fourth component of the HAT module [14,15]. The discovery of these four proteins in these two complexes, in which they form a shared HAT module, was the first step towards the understanding of the modular organization of the SAGA complex (Table 1).

Genetic and functional overlap between Ada and proteins isolated as suppressors of Ty and solo δ insertion mutations in yeast (Spt) led to the identification of additional SAGA components. Indeed, all Spt proteins that are functionally linked to the TATA-binding protein (TBP/Spt15), namely Spt3, Spt7, Spt8 and Spt20 (Ada5), were found to physically interact with Gcn5, Ada2 and Ada3 [3,16]. These SAGA subunits were classified in two groups according to the phenotypes of their mutants and to their roles in complex assembly [4,17]. Mutations in *S. cerevisiae* SPT20/ADA5, SPT7, and ADA1 cause severe growth phenotypes and are crucial for SAGA structural integrity [4,17]. Spt3 and Spt8 mutants exhibit milder growth phenotypes, are not crucial for SAGA integrity, and interact genetically with TBP in *S. cerevisiae* [17–19]. These landmark studies gave the first hints suggesting that distinct SAGA subunits may assemble in three discrete functional modules: a HAT module (Ada2, Ada3 and Gcn5), a TBP-interaction module (Spt3 and Spt8) and a structural module (Ada1, Spt7 and Spt20) (Table 1).

Further analyses of native SAGA complexes purified from *S. cerevisiae* identified a subset of TBP-associated factors (TAFs) shared with the TFIID complex (Taf5, Taf6, Taf9, Taf10 and Taf12) as integral components of SAGA [20]. Four of these are histone-fold (HF) containing TAFs (Taf6, Taf9, Taf10 and Taf12) [21,22] were proposed to have a structural role within SAGA. Indeed, Taf6, Taf9, Taf12, together with Ada1 were shown to assemble *in vitro* in a histone octamer-like complex [23]. Four additional HFs were further identified in other SAGA subunits (Ada1, Spt7 and two in Spt3) [24–26], suggesting that a histone octamer-like structure could also be formed by four HF pairs: Taf6-Taf9, Ada1-Taf12, Taf10-Spt7, Spt3-Spt3 (Table 1).

A fourth functional module composed of a single large subunit was identified with the co-purification of Tra1 within SAGA fractions [27,28]. Finally, an extensive proteomic analysis of TAF-containing complexes revealed that *S. cerevisiae* SAGA has a second enzymatic activity, carried out by the ubiquitin protease Ubp8 [15]. To be fully active on its substrate, mono-ubiquitinated histone H2B, Ubp8 has to associate with three other proteins, Sgf73, Sgf11 and Sus1, forming the fifth functional module of SAGA [29–32].

2.2. Evolutionary conservation of SAGA composition

Gcn5 homologues in fungi, metazoan and plants are highly conserved, except that metazoan Gcn5 homologues possess an N-terminal extension ([33] and references therein). In addition, whereas most metazoan and plant genomes encode one GCN5-type factor, vertebrates have a second paralogous gene encoding PCAF (p300/CBP-associated factor), which amino-acid sequence is 73% identical to GCN5 (Table 1) [34]. GCN5 and PCAF, as most native nuclear HATs, exist as components of large multisubunit complexes. GCN5-, or PCAF-containing complexes purified from all tested vertebrate organisms were very similar to the 2MDa yeast SAGA complexes [35–41]. Further biochemical studies confirmed that the modular organization is highly conserved throughout eukaryotes (Table 1) [42]. Nevertheless, notable differences exist between SAGA complexes purified from distinct species (Table 1). First, the metazoan and plant SAGA complexes do not contain homologues of yeast Spt8. Second, the metazoan SAGA complexes co-purify with two splicing factors, SF3B3 and SF3B5, suggesting that an

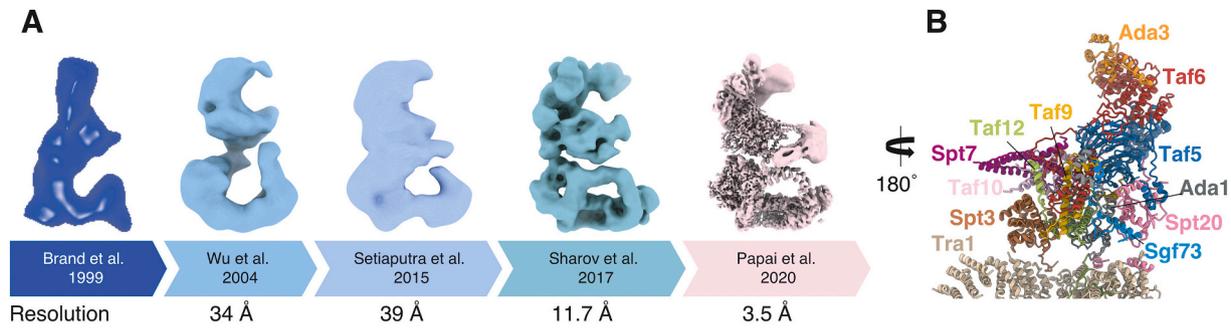


Fig. 1. A) Representation of the evolution of the resolution of the SAGA structures over the last 21 years. The structures were obtained from [46,47,51,52,131] (available EMD numbers: 6301, 3804, 10438). B) High resolution structure of the SAGA core and its connection to Tra1 shown as a ribbon model (PDB: 6TB4).

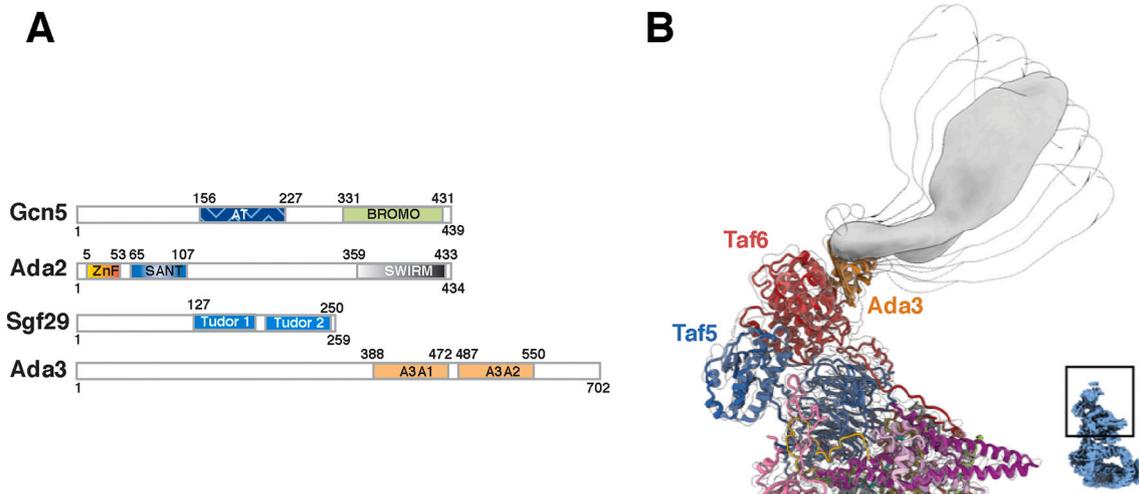


Fig. 2. A) Schematic representation of the four subunits of the *S. cerevisiae* SAGA HAT module. Amino acid coordinates of each factor are labelled together with the different functional domains described in each factor. AT: Acetyl-transferase; BROMO: bromodomain; ZNF: zinc-finger domain; SANT: Swi3, ADA2, N-Cor and TFIIB; Tudor1 and Tudor 2: double Tudor domain; A3A1 and A3A2: Ada3 anchoring domains. B) Schematic representation of the flexible HAT module and its incorporation in the high-resolution SAGA structure from Papai et al. [52]. Low resolution HAT densities (white) represent flexibility described in [52] while grey density represents the most frequent position. Colours of the subunit are the same as on Fig. 1B. The insert highlights the position of the HAT module within the overall structure of SAGA.

additional module exists in metazoan SAGA, but is absent from yeasts and plants [43,44]. Third, plants do not have an ortholog of Sgf73 suggesting that the DUB module is different from that in yeasts and animals [40,45].

2.3. Overview of SAGA topological organization

Structural information is crucial to understand why different activities are found together within a complex and how they are coordinated. Since its discovery, much work has been done to determine the overall architecture of SAGA and its different modules. Early structural analyses supported the idea that the identified functional entities form distinct structural modules (Fig. 1) [46,47]. However, the models of the relative positions of different subunits were sometimes contradicted by results obtained with orthogonal structural and biochemical approaches [48,49]. The rapid expansion of cryo-electron microscopy (cryo-EM) as a result of better electron detectors and image-processing techniques allowed the acquisition of high-quality images and high-resolution structures of very large complexes. Using cryo-EM, the structure of the large Tra1 subunit and its anchoring to the rest of the SAGA complex was determined a few years ago [50,51]. Over 20 years after SAGA discovery, two recent reports finally provided high-resolution models of SAGA at different stage of transcription regulation, either bound to TBP [52] or to nucleosomes [53] (Fig. 1). Notably, these studies describe a central core module with an atypical histone-

fold octamer-like conformation, providing a mechanism for how SAGA loads TBP to promoters. These structures also unequivocally locate the other functional modules and provide unprecedented molecular details into their connections with the central core, paving the way for a better understanding of how SAGA topology and modularity mediates its functions.

2.3.1. Structure/function relationship of the two enzymatic modules

It was recognized early that the two enzymatic subunits of SAGA, Gcn5 and Ubp8, require association with specific co-factors for their catalytic activation and for substrate recognition [13,14,49,54–57]. Both enzymatic modules are now known to be composed of four subunits making extensive contacts with each other. Determination of the low-resolution structure of the human HAT module and the atomic structure of the recombinant yeast DUB module, confirmed that the previously defined functional modules form discrete structural entities.

2.3.1.1. The histone acetyltransferase module. An almost identical HAT module is shared between all identified Gcn5-containing complexes (Table 1). Early work suggested that Ada2, Ada3 and Gcn5 assemble into a complex with coactivator function bridging activation domains with the general transcription machinery [10,58]. The same high sensitivity proteomic analyses that identified subunits of the SAGA DUB module identified an additional subunit, called Sgf29 [15]. Subsequently Sgf29 has been described as the fourth subunit of the

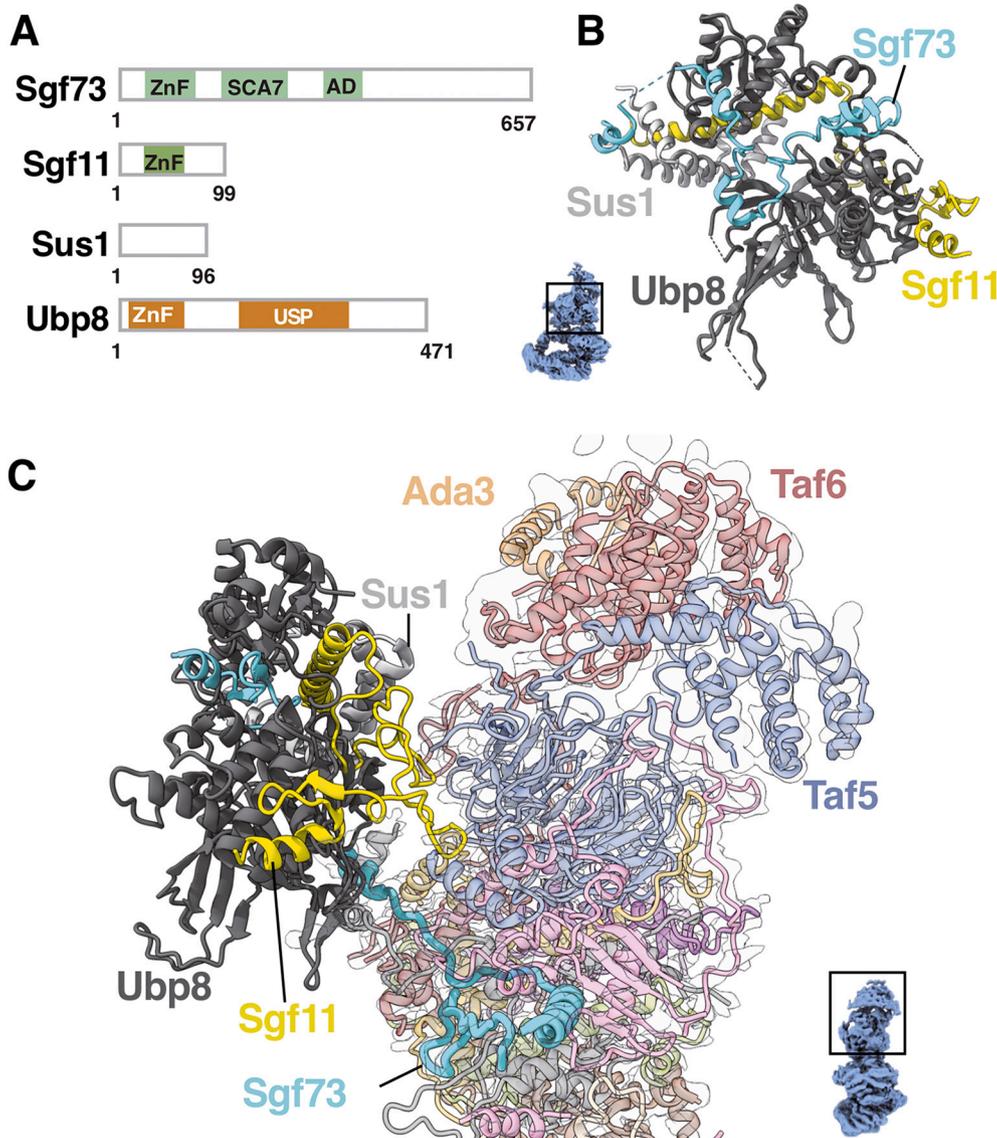


Fig. 3. A) Domain organization of the four subunits of the *S. cerevisiae* SAGA DUB module. ZnF: Zinc Finger domain, SCA7: SpinoCerebellarAtaxia 7, USP: Ubiquitin Specific Protease. B) The assembly lobe and the catalytic lobe of the *S. cerevisiae* DUB module are organized around the ZnF-UBP and the catalytic domain of Ubp8, respectively. Ubp8 (dark grey), Sgf11 (yellow), Sus1 (light grey) and Sgf73 (cyan) are shown in ribbon representation. The location of the represented structure is indicated by a rectangle on the overall structure of SAGA (EMDB: 10438). C) Sgf73 anchors the DUB module to the SAGA core. The composite image shows the crystal structure of the DUB module docked into the cryo-EM map of the *K. phaffii* combined with a ribbon model of the core module (6TBM). Colours of the subunit are the same as on Fig. 1B and panel B. In panels B) and C) the inserts highlight the regions of interest within the overall structure of SAGA.

HAT module [49]. Sgf29 contains a double Tudor domain and was shown to be important to recruit GCN5-containing complexes to H3K4 trimethylated chromatin [14,44,59].

Sequence analysis of the four SAGA HAT subunits revealed that all of them contain evolutionary conserved domains (Fig. 2A). All the eukaryotic Gcn5 homologues contain an acetyl transferase (AT) domain and a C-terminal bromodomain [33]. Bromodomains are approximately 110-amino-acid motifs found in many chromatin-associated proteins. Structural analysis revealed that the bromodomains of Gcn5 or PCAF have a specific affinity for acetyl-lysine-containing motifs, including acetylated histone H4, suggesting that bromodomains can interact with chromatin, but also with other proteins in an acetylation-dependent manner [60–63]. Indeed, specific point mutations suggest that the interaction of the Gcn5 bromodomain with acetyl-lysine is important for Gcn5-dependent transcription and histone acetylation at promoters [64].

Yeast Ada2 and metazoan ADA2b homologues all contain three well defined domains: a N-terminal ZZ type zinc finger domain, a SANT domain (found in Swi3, ADA2, N-Cor and TFIIB) involved in chromatin binding and a C-terminal SWIRM domain found in proteins implicated in chromatin remodeling and gene expression. Ada3 homologues are suggested to contain N- and C-terminal coiled-coil domains. Sgf29 homologues contain a N-terminal coiled-coil and a C-terminal double

Tudor domain, a structural motif involved in binding tri-methylated histone H3K4 (H3K4me3), which is a hallmark of active promoters [14,44].

An atomic model of most of these domains exist, including for the AT, bromo-, ZZ, SANT and double Tudor domains [65]. Crystal structures of the Ada2 ZZ and SANT domains in interaction with the Gcn5 AT domain, show that the SANT and ZZ domains of Ada2 interact with N- and C-terminal extensions of the Gcn5 AT domain. Furthermore, the Ada2 SANT stimulates the Gcn5 HAT activity by enhancing the binding of its cofactor acetyl-CoA [54]. Chemical cross-linking-mass spectrometry in combination with modeling allowed building a low-resolution structural model of the human SAGA HAT module [65]. In the tetrameric complex, GCN5 crosslinks with all the other three subunits, while ADA2b interacts only with ADA3 and GCN5, and SGF29 only with ADA3 and GCN5. Moreover, it was suggested that the HAT module may change its conformation when integrating in the holo-SAGA complex. (See more details in the article of Espinola Lopez and Tan in this Special Issue about the function of the SAGA HAT module).

Very recently, the core structure of the yeast SAGA complex was resolved at 3.3–3.5 Å resolution [52,53]. In these cryo-EM structures, the HAT module adopts multiple conformations and thus its structure could not be interpreted at high resolution (Fig. 2B). However, the bulk of the HAT module is visible as an elongated density at the opposite side

of the complex, when compared to the location of Tra1 (Fig. 2B). Nevertheless, Papai et al. [52] identified the docking site of the HAT module on the SAGA core and showed that two helical domains of Ada3 interact with the surface of the Taf6 HEAT-repeat containing domain. This observation is in good agreement with crosslinking data and structure-function experiments in yeast [48]. Overall, biochemical and structural approaches indicate that interactions between Ada3 and Taf6 mediate the docking of the HAT module to the rest of SAGA. Other HAT module subunits may also participate in the incorporation of the HAT module into SAGA, as the deletion of Ada2 was shown to result in the loss of the whole HAT module from SAGA [49]. It is noteworthy that Wang et al. [53] hypothesized that the HAT module of SAGA becomes flexible and/or is displaced upon nucleosome binding to better accommodate the HAT activity of SAGA for nucleosome modifications around the promoters of active genes. In conclusion, it seems that Ada3 plays a prominent role in integrating the HAT module into the holo-SAGA complex. In contrast, Gcn5 AT and bromodomain, as well as the double Tudor domains of Sgf29, are likely positioned on the outer side of the HAT module and recognize activating histone marks to maintain acetylation of nucleosomes at active promoters.

2.3.1.2. The histone deubiquitinase module. The yeast SAGA DUB module is formed by the assembly of Ubp8 with Sgf11, Sus1 and the N-terminal ~100 residues of Sgf73 (Fig. 3A, B; and see Table 1). Biochemical and structural studies demonstrated that Ubp8 is activated by interactions with the three other subunits and revealed a key role for Sgf73 in anchoring the DUB module within SAGA [32,48,55,57,66]. Crystal structures of recombinant DUB module revealed that it is organized in two lobes around two structural domains of Ubp8, the C-terminal catalytic USP domain and the Zinc Finger-Ubiquitin Binding Protein (ZnF-UBP) N-terminal domain (Fig. 3B) [55,57]. At the center of the assembly lobe, a long helix of Sgf11 contacts the Ubp8 ZnF-UBP domain on one side and the entire Sus1 protein on the other side. C-terminal to its alpha helix and joined by an extended linker, a ZnF domain in Sgf11 binds Ubp8 close to its catalytic domain, forming the catalytic lobe of the DUB module. The N-terminal region of Sgf73 is located between the two lobes and makes extensive contacts with the three other DUB subunits. The two ZnF domains from Sgf11 and Sgf73 make different contributions to the DUB module activity. The Sgf11-ZnF facilitates DUB module activity by maintaining the active conformation of Ubp8, whereas the Sgf73-ZnF is required for Sgf73 association with the DUB module and maintains the open conformation of the ubiquitin-binding pocket of Ubp8 [32,55,67,68]. In addition, the Sgf11 ZnF contains several arginine residues that play a crucial role in targeting the DUB module to nucleosomes [57,69]. Indeed, a crystal structure of a DUB module bound to ubiquitinated nucleosomes indicates that the arginine cluster in the Sgf11 ZnF binds the acidic patch of the H2A/H2B heterodimer. In contrast, Ubp8 mainly contacts ubiquitin, as well as the H2B C-terminal helix, which contains the ubiquitinated lysine [70]. Mutations of the basic residues in Sgf11 had more severe effects on the DUB activity when compared to mutations of Ubp8 residues interacting with nucleosomes. Thus, Sgf11 appears to make the most significant contribution to the specificity of substrate recognition.

An obvious advantage for the requirement of Ubp8 activity to associate with other factors would be to prevent untargeted de-ubiquitination activity. The mandatory incorporation of Ubp8 within SAGA might also restrict its substrate specificity and direct its action to specific genomic locations. However, several evidences suggest that a free DUB module can exist independently of the SAGA complex. For example, in *S. cerevisiae*, the 19S regulatory particle of the proteasome interacts *in vitro* with Sgf73 and stimulates the ATP-dependent dissociation of the DUB module from the SAGA complex [71]. Further work is needed to understand the dynamics of SAGA-DUB interaction *in vivo* and how it contributes to transcriptional regulation.

Several puzzling observations in different organisms further support

the existence of a functional DUB outside of SAGA. Genome-wide analyses of SAGA regulatory roles in *Drosophila* suggest that the DUB module binds and regulates a subset of promoters independently of SAGA [72]. Purifications of SAGA from two filamentous fungi, *Aspergillus nidulans* and *Fusarium fujikuroi*, apparently lack Ubp8 and Sgf11, although Sgf73 is present. Surprisingly, their genomes have orthologous genes encoding Ubp8 and Sgf11, suggesting that the DUB is either transiently associated with SAGA or functions completely independently [73,74]. Accordingly, work in model plant species demonstrated that the DUB module functions outside of SAGA. Indeed, the genomes of several plants, including *Arabidopsis thaliana*, *Oryza sativa* and *Vitis vinifera* lack a gene encoding a Sgf73 orthologue, but have orthologues of Ubp8, Sgf11, and Sus1 (called ENY2) (Table 1). In agreement with Sgf73's anchoring role, subunits of the DUB were not found associated with the SAGA complex in these plant species [40,41,45]. Surprisingly, despite the absence of the N-terminal ZnF from Sgf73, the plant DUB complex efficiently deubiquitinates histone H2B, although this region is required for the full activation of Ubp8/USP22 in yeast and human cells [32,56]. (See more details in the article of Grasser et al. in this Special Issue about the function of the DUB module of plants <https://doi.org/10.1016/j.bbagr.2020.194613>).

Supporting this possibility, work in human cells revealed the existence of catalytically active DUB modules lacking ATXN7, the orthologue of yeast Sgf73. Indeed, the human genome encodes two de-ubiquitinases, USP27X and USP51, which are closely related to USP22, the orthologue of yeast Ubp8. Each of these USP22 paralogs forms a small DUB complex with the human orthologue of Sgf11, ATXN7L3, and ENY2, the orthologue of yeast Sus1 [75]. The N-terminal ZnF domain of ATXN7L3 with its conserved arginine residues likely play a similar role to that of Sgf11 in docking these small DUB complexes to the H2A/H2B acidic patch. Indeed, these independent DUB modules, containing either USP27X or USP51, are catalytically active on mono-ubiquitinated H2B. Additional experiments are needed to understand how SAGA and these two small DUB complexes contribute to global and targeted de-ubiquitination of H2B.

In summary, the association of the DUB module with the SAGA complex show very different behaviors throughout evolution, perhaps depending on whether the DUB activity needs to be coordinated with other SAGA activities or not. This raises the question of the mechanisms leading to either a stable incorporation of the DUB module within SAGA or, on the contrary, to a fully independent action of the DUB module. Sgf73 has a dual role, with its N-terminal region being required for DUB module assembly as well as Ubp8 activation and its C-terminal region anchoring the DUB module within SAGA [32,76]. Recent cryo-EM structures of SAGA show that Sgf73 penetrates deeply into the SAGA central core module through its a central region (Fig. 3C) [52,53]. This region is connected to the Sgf73 N-terminal ZnF through a linker and the structured SCA7 domain, which was not resolved in the cryo-EM structure, but is dispensable for DUB interaction with SAGA [77,78]. The central region of Sgf73 forms two helices and several loops that make extensive contacts with several subunits of the core module. Particularly, the second anchoring helix is embedded in a chamber formed by Ada1 and the SEP (shp1-eyc-p47) domain of Spt20. Although the position of the DUB module relative to the core appears highly flexible and has few contacts with the SAGA core, it is tightly tethered to SAGA, as suggested by the multiple contacts that the Sgf73 anchoring domain makes with core SAGA subunits. However, both studies were performed in conditions in which SAGA subunits were recovered in stoichiometric amounts, including the DUB. Thus, whether a similar mechanism exists in species in which the DUB might interact only transiently with the rest of SAGA is an interesting question for future studies.

2.3.1.3. Structure of the activator binding module. Targeting chromatin-modifying activities to specific genes is important for the spatiotemporal regulation of transcription. Therefore, understanding

how transcription factors recruit SAGA to promoters is a central question. Seminal biochemical studies from the Workman, Green, and Hahn labs established that this function involves the largest subunit of SAGA, called Tra1 in yeast and TRRAP in humans. TRRAP was originally identified in mammalian cells as an interacting partner for the c-MYC and E2F proto-oncogenic transcription factors [79] and its yeast ortholog quickly identified in purified SAGA fractions [27,28]. Tra1 was then identified in the NuA4/TIP60 complex, which is another conserved transcriptional co-activator with acetyltransferase activity [80]. Tra1/TRRAP belongs to a family of atypical kinases, named the phosphoinositide 3 kinase-related kinases (PIKKs), but lacks the catalytic residues required for kinase activity and is therefore the sole inactive member within this family (reviewed in [81]).

Tra1 is essential for viability in *S. cerevisiae* and TRRAP is essential for early embryonic development in mice [28,82]. Genetic analyses using partial loss of function or conditional mutant alleles have implicated Tra1/TRRAP in a number of important processes, such as cell cycle progression and stem cell maintenance. Elegant biochemical studies have shown direct interaction between Tra1/TRRAP and several transcription activators, indicating that its primary role is to target co-activator complexes to specific promoters [79,83–89]. In summary, its large size, lack of kinase activity, and ability to directly contact activators suggested that Tra1/TRRAP serves as a scaffold for the assembly and recruitment of co-activator complexes to chromatin. However, for many years, it remained difficult to identify the specific roles of Tra1/TRRAP within either SAGA or NuA4/TIP60, in part because the genetic dissection of this very large protein (~420 kDa) is challenging. However, recent work brought unprecedented insights into our understanding of Tra1/TRRAP structure-function relationship.

First, advances in cryo-EM solved the structure of yeast Tra1 at high resolution [50,51]. Like all PIKKs, Tra1/TRRAP has a characteristic domain architecture, which includes repeated superhelical structural motifs forming the HEAT and FAT domains, preceding the highly conserved, PI3K-related, kinase domain and a short C-terminal FATC motif [79,90]. Structural analyses identified three distinct regions within the N-terminal array of HEAT repeats, the ‘Finger’, ‘Ring’, and ‘Clasp’ domains, which fold into an α -solenoid superhelical structure. The remaining FAT, kinase, and FATC domains form a globular region, termed ‘Head’, which adopts a similar conformation to other PIKKs.

Second, the fission yeast *Schizosaccharomyces pombe* provides a unique opportunity to study Tra1 topology and function, because it has two paralogous proteins, Tra1 and Tra2, and each has non-redundant functions that are specific for SAGA and NuA4, respectively [91]. In addition, a *tra1* Δ deletion mutant is viable, facilitating detailed structure-function analyses. Biochemical and phenotypic analyses in *S. pombe* revealed that, within SAGA, Tra1 has specific regulatory roles

and does not scaffold the entire complex but, rather, controls the incorporation of the DUB module. In contrast, Tra2 contributes to the overall integrity of the NuA4 complex [92]. These observations agree with the most recent structures of yeast SAGA and NuA4, which showed that Tra1 is positioned differently relative to the other subunits. Indeed, within SAGA, Tra1 localizes to the periphery of the complex, whereas it occupies a more central position within NuA4, contacting subunits from different modules [51,93]. We therefore speculate that the single Tra1/TRRAP protein found in most other eukaryotic organisms has distinct architectural roles between each complex and functions as a scaffold only within the NuA4/TIP60 complex.

The latest structures of SAGA show that Tra1 interacts with the rest of the complex through narrow surface interactions [51–53]. Biochemical evidence indicates that a small, 50-residue region of the Tra1 FAT domain is both necessary and sufficient to anchor Tra1 within in SAGA in both *S. pombe* and *S. cerevisiae* [92]. This region of Tra1 forms a concave structure that clamps a short helix from the Spt20 subunit [52,53]. Remarkably, the homologous region from *S. pombe* Tra2 diverged such that it cannot interact with Spt20. Transcriptomic and phenotypic analyses of mutants specifically disrupting this interaction confirmed that Tra1 contacts SAGA mainly through this region of Spt20 *in vivo* [92]. Nevertheless, the most recent structures of SAGA clearly show that other subunits, Taf12 and Spt3, make direct contacts with Tra1. It is possible that SAGA exists in different conformations, depending on the number of contacts between Tra1 and these subunits. Whether this flexibility allows allosteric regulation between modules is an exciting question for future studies.

To conclude, integrating structural, biochemical, and functional approaches established that Tra1/TRRAP is a pseudoenzyme acting as a protein-protein interaction hub during transcription initiation. Although the molecular details remain elusive, the current model for Tra1/TRRAP function suggests that it conveys the signal from promoter-bound transcription factors to SAGA and NuA4/TIP60, perhaps controlling their chromatin-modifying activities and pre-initiation complex assembly to elicit specific transcriptional responses.

2.3.1.4. Structure of the SAGA core. One of the most spectacular breakthroughs from the recent cryo-EM studies is the high-resolution structure of the SAGA core module, whose organization remained elusive [52,53]. The overall topology of SAGA is formed by two lobes. One lobe is almost exclusively constituted by the single Tra1 protein (see above) whereas all other subunits are located in the main lobe. In its center, the core module of SAGA is formed by an intricate assembly of 9 subunits which tethers all other modules to the complex. Each module attaches to different regions of the SAGA core with different degrees of flexibility (Figs. 2–4). In contrast to Tra1, the

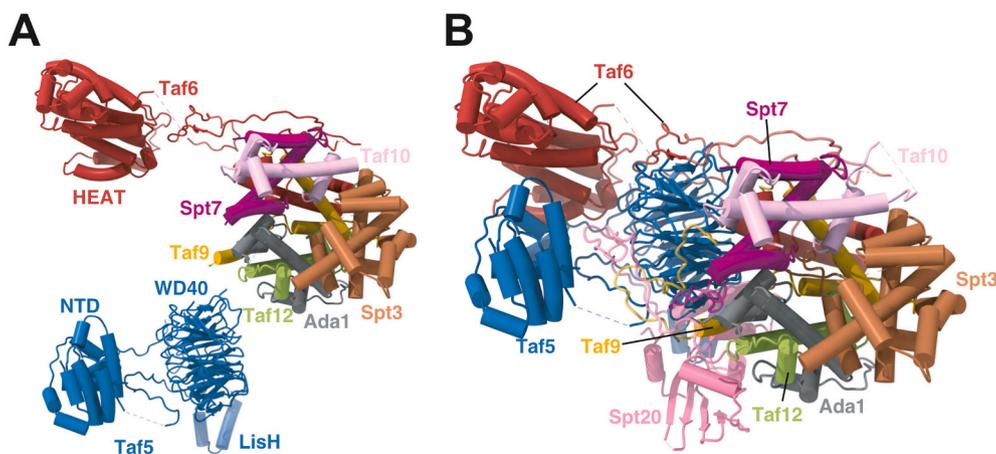


Fig. 4. Taf5-Taf6 heterodimer organizes the central core module. A) Individual models of two key structural elements of the SAGA core are shown. The NTD, WD40 and LisH domains of Taf5 are shown in blue. The N-terminal histone fold domain of Taf6 (red) is part of the octamer-like structure and is connected to its HEAT domain through a long linker. For easier visualisation, only the histone folds of other subunits are indicated. These two structures were artificially isolated from the SAGA structure for visualisation and do not correspond to existing cellular entities. B) Topology of the SAGA core module highlighting the domain organization of the Taf5-Taf6 heterodimer and their connection to the octamer-like structure. This view shows the central core

structure of the two enzymatic modules could not be resolved at high resolution in SAGA cryo-EM maps, because of the high flexibility of the HAT and DUB modules. One possible explanation for this phenomenon is that both modules can adopt multiple orientations to accommodate their substrate in the context of nucleosomal chromatin. Densities of the HAT and DUB modules were lost upon SAGA binding to nucleosomes containing mono-ubiquitinated histone H2B, suggesting increased flexibility of these modules [53]. This was suggested to facilitate the coordination of the different SAGA activities (histone acetylation, de-ubiquitination, and TBP-binding) on promoters with different characteristics. Supporting this possibility, previous work suggests that the DUB module modestly stimulates Gcn5 activity on mononucleosomal substrates [48].

The core module of SAGA is organized around several key structural elements (Fig. 4A,B). These include a histone octamer-like fold, the WD40 repeat domain at the C-terminus of Taf5, the N-terminal domain of Taf5 and the HEAT repeat domain of Taf6. Analogous to the canonical histone octamer in nucleosomes, a potential octamer-like structure could be modelled within the cryo-EM map of the human TFIID complex [94]. This octamer contains the Taf6-Taf9, Taf4-Taf12, Taf3-Taf10 and Taf11-Taf13 heterodimers, which interact through histone-fold (HF) domains. Yeast SAGA shares Taf6, Taf9, Taf10 and Taf12 with TFIID and has specific subunits with HFs, including Ada1, Spt7 and Spt3 [25,26]. Therefore, a long-standing hypothesis was that SAGA contains an octamer-like structure at its core, formed by the Taf6-Taf9, Spt7-Taf10, Ada1-Taf12 pairs and the two HFs from Spt3. Both recent SAGA structures visualized this octamer-like structure for the first time, at the edge of the SAGA central lobe (Fig. 4B). In a canonical histone octamer, the HF pairs are distributed along a spiral in the following order, H2A-H2B, H3-H4, H3-H4 and H2A-H2B, forming a symmetrical structure. The SAGA octamer starts with the Taf6-Taf9 pair, linked to Taf12-Ada1, then to Taf10-Spt7, and ends with the two HF domains of Spt3. At the end of the octamer, the packing of the Spt3 HF pair deviates from the symmetrical organization of the canonical histone octamer. As a consequence of the unconventional link between Taf10 and Spt3, the Spt3 HF pair is relatively unconstrained from the octamer, but is tethered by a C-terminal extension contacting all other HF pairs at the center of the octamer. This specific organization is of crucial importance for TBP-binding as revealed by the structure of the SAGA complex bound by TBP [52]. The C-terminal region of TBP interacts with the Spt3 HF pair whereas the TBP N-terminus binds Spt8 which is flexibly tethered to Spt7 in the SAGA core. The TBP DNA-binding surface faces the SAGA core and DNA binding in the available space is prevented by different structural elements. Biochemical experiments demonstrated that the general transcription factor TFIIA is needed for initial binding of TBP to DNA followed by the release of DNA-bound TBP from SAGA [52].

The WD40 repeats of yeast Taf5 form a seven blade β propeller which is located at the center of the core module. It interacts on one face with the Taf6-Taf9 and Taf10-Spt7 histone fold pairs and on the other face with both the N-terminal domain of Taf5 and the Taf6 HEAT repeats (Fig. 4B). The association between Taf5 and Taf6 in SAGA and their interactions with other core subunits result in a different orientation of the Taf5 N-terminal domain and the Taf6 HEAT repeats from that observed in TFIID. In TFIID, the N-terminal domain of Taf5 docks the octamer-like structure and is stabilized by the TFIID-specific subunit Taf4. In SAGA, the Taf5 N-terminal domain is located on the upper face of the WD40 structure and is stabilized by interactions with Spt20 and the Taf6 HEAT repeat domain (Fig. 4B). Thus, the Taf5-Taf6 heterodimer has a major structural function, making contacts with more than ten protein domains in the SAGA central module.

3. Additional Gcn5-containing complexes and their functions

3.1. Yeast ADA complex

Gcn5 is found in multiple, but related HAT complexes in different species in which it forms a HAT module together with Ada2 (ADA2a or ADA2b in metazoans), Ada3 (ADA3 or TADA3) and Sgf29 (SGF29). The yeast Gcn5 protein was identified as a component of the small 0.8 MDa ADA complex which contains all SAGA HAT module components, together with two other specific subunits, ADA HAT component 1 (Ahc1) and Ahc2, demonstrating that ADA functions as a distinct complex in *S. cerevisiae* [3,49,95]. Indeed, deletion of yeast Ahc1 results in the specific loss of the ADA complex from yeast extracts, but does not cause an Ada phenotype, which would depend on SAGA functions [95]. This ADA complex acetylates nucleosomal histone H3 *in vitro* [96]. In contrast with SAGA, the ADA complex does not contain the Tra1 subunit, which is proposed to recruit SAGA to target genes, suggesting other recruitment mechanisms for the ADA complex. The bromodomain of Gcn5, the tandem Tudor domains of Sgf29 and the Ada2b ZZ and SANT domains might recruit the ADA complex to chromatin.

3.2. Two Gcn5-containing small complexes (ADA and CHAT) in *Drosophila*

The identification of the ADA complex in yeast and of SAGA-independent DUB modules in several species, suggested that small Gcn5-containing complexes may also exist in metazoans. Two such complexes were recently characterized in *Drosophila*, corresponding either to a typical HAT module functioning as an independent complex called 'metazoan ADA complex', or to a HAT module associated with Chiffon, forming the Chiffon acetyltransferase (CHAT) complex [97,98]. Formation of these complexes is partially explained by the existence in *Drosophila* of two splice variants of Ada2b which differ by their C-terminal region, leading to their specific incorporation in different complexes [99,100]. The short Ada2b isoform is exclusively incorporated in the CHAT complex, but not in SAGA or ADA complexes. In insects, but not in other organisms, Chiffon is a bicistronic gene and the sole C-terminal product interacts with Gcn5 to incorporate in the CHAT complex [98]. In agreement with the insect specificity of the C-terminal domain of Chiffon, orthologues of Chiffon in yeast or mammals do not interact with Gcn5 and no CHAT type complexes could be isolated from yeast or vertebrates. The CHAT complex was shown to be required for histone H3 acetylation *in vivo* and was proposed to regulate gene expression in flies [98] (see also the article of Torres-Zelada and Weake in this Special Issue for more details on the function of the *Drosophila* CHAT complex <https://doi.org/10.1016/j.bbagr.2020.194610>).

In contrast, the long Ada2b isoform is excluded from CHAT, but appears to be an exclusive subunit of both SAGA and the recently described 'metazoan ADA complex' in *Drosophila* [97]. The fly ADA complex thus contains the long Ada2b isoform together with Gcn5, Ada3 and Sgf29, but does not contain any additional specific subunits. This tetrameric HAT module complex was suggested to be a discrete entity that does not result from *in vitro* dissociation of SAGA and to function independently of the large SAGA complex. ChIP-seq analyses revealed a few number of genomic sites bound by HAT subunits but not by SAGA-specific subunits, suggesting that the ADA complex, which has *in vitro* acetyltransferase activity on nucleosomal histones, could bind and regulate the expression of specific genes [97]. Thus, in addition to its function as part of the SAGA complex, the tetrameric Gcn5-containing HAT module can be biochemically isolated as an independent acetyltransferase complex and seems to function independently from the large 2 MDa SAGA complex. (See also the article of Torres-Zelada and Weake in this Special Issue for more details on the function of the *Drosophila* ADA complex <https://doi.org/10.1016/j.bbagr.2020.194610>).

3.3. Metazoan ATAC complexes

GCN5 was identified also as a subunit of a ten subunit coactivator HAT complex in *Drosophila* and mammals, named ATAC (Ada two A-containing) [101–103], but not in plants [41]. Interestingly, the yeast Ada2 protein has two paralogues in all metazoans, called ADA2a and ADA2b [104,105]. Similarly to *Drosophila*, ADA2b is part of the HAT module of all the metazoan SAGA complexes, and ADA2a is a subunit of the HAT module of all metazoan ATAC complexes together with GCN5, ADA3, and SGF29 ([42] and refs therein). The HAT modules of mammalian ATAC complexes exist with either GCN5 or its vertebrate paralogue, PCAF, in a mutually exclusive manner [103].

Although eukaryotic SAGA complexes preferentially acetylate histone H3K9 and H3K14 lysine residues [35,96,106,107], specificities of the metazoan ATAC complexes are less well understood [102,103,108]. After knock-down of ATAC subunits acetylation of H4K5, -K12, and -K16 decreases in *Drosophila* [109–111]. In contrast, analyses of changes in histone tail acetylation patterns upon knock-out or knock-down of specific ATAC subunits in mammalian cells suggested that ATAC preferentially acetylates histone H3 rather than H4 [102,103,108].

As the structure of the entire ATAC complexes is not yet known, and as structural information is only available about the different conserved domains of ATAC subunits (Fig. 5), in the next sections we will summarize the structure/function relationship of the existing data. It has been described that *in vitro* ADA2b, but not ADA2a, increases the acetylation of nucleosomes by GCN5 [112] and that the ADA2b-containing SAGA HAT module stimulated the HAT activity of GCN5 better than that of the ADA2a-containing ATAC HAT module [113]. Both ADA2a and ADA2b have ZZ, SANT and SWIRM domains. Analyses of the role of the SANT domain of yeast Ada2 revealed that the domain was required for full HAT activity of the associated Gcn5 [54]. Thus, it is possible that the ADA2a SANT domain, in contrary to the corresponding domain of ADA2b, interacts differently with GCN5 and does not enhance efficiently the binding of the acetyl-CoA to GCN5 as described for the SANT domain of yeast Ada2. Nevertheless, the catalytic activity of GCN5 is stimulated collectively by all subunits of the ADA2a-containing HAT module, and is further increased when the HAT module is incorporated in the ATAC holo-complex [113]. These observations suggest that the interactions of the HAT with the other functional module of ATAC complex would further accommodate the structure of the HAT module of ATAC for a full activity, as suggested for yeast SAGA [53].

In addition to the ATAC-specific HAT modules, the metazoan ATAC complexes all have a second module, composed of six *bona fide* subunits with distinct chromatin/transcription linked properties, such as chromatin recognition and binding (YEATS2 and ZZZ3), transcription regulation (NC2 β), nucleosome remodeling (WDR5) and a second putative HAT enzyme (ATAC2 or Cysteine-rich protein 2-binding protein, or CRP2BP) (Fig. 5) [102,103,108,114,115]. YEATS2 is an evolutionary conserved protein. *In vitro* the YEATS domain of YEATS2 recognizes acetylated and/or crotonylated H3K27 peptides [114,116]. Thus, it is thought that the YEATS domain of YEATS2 is a selective H3K27 acetyl (ac)/crotonyl (crot) chromatin reader, suggesting a model wherein YEATS2 would participate in recruiting the ATAC complex to H3K27ac-, or H3K27crot-enriched target genes to promote active transcription *via* maintaining promoter histone H3K9 acetylation levels [114]. YEATS2 was also described as a scaffolding subunit of the core module of ATAC, as its N-terminal end (first 386 amino acids) interacts with the HAT module of ATAC (GCN5, ADA2a, ADA3 and SGF29) and its C-terminal end (from amino acids 1328 to 1422) is required for interactions with WDR5, MBIP and NC2 β [102]. In addition, it was shown that the C-terminal region of YEATS2 contains a HFD through which it forms a histone fold pair with NC2 β and that the YEATS2-HFD/NC2 β histone fold pair can interact with TBP. Further *in vitro* transcription experiments suggested that the TBP-interacting NC2-like

module (YEATS2-HFD/NC2 β) present in ATAC complexes could be involved in both positive and negative regulation of transcription [102].

ZZZ3, another component of the core module of the ATAC complex, harbors a SANT and a ZZ domain (Fig. 5). Using multiple orthogonal approaches, it was demonstrated that the ZZ domain at the C-terminus of ZZZ3 binds selectively acetylated histone H3K4 marks [115]. Experiments with loss-of-function ZZ-mutants pointed to an important role of the ZZ domain in the association of the ATAC complex with chromatin. Therefore, it was suggested that the ZZ domain of ZZZ3 cooperates with other readers in the core ATAC module to facilitate its chromatin recruitment and, potentially, the catalytic activity of the HAT module.

MBIP (MUK-binding inhibitory protein) was originally identified as a mitogen-activated protein kinase (MAPK) upstream kinase (MUK)-interacting protein in a two-hybrid screen and shown to inhibit the action of MUK in activating the c-Jun N-terminal kinase (JNK) [117]. MBIP contains two tandemly orientated leucine-zipper-like motifs with a cluster of basic amino acids located between the two motifs (Fig. 5). Subsequently, MBIP homologues have been identified in all metazoan ATAC complexes. The *Drosophila* homolog of MBIP was found in a fusion with the MoaE protein, a subunit of molybdoprotein synthase, and thus, it was suggested that the MoaE/MBIP fusion protein plays a non-catalytic role in MAPK signaling in *Drosophila* as part of the ATAC complex [118]. Interestingly, several genome-wide association studies (GWAS) analyzing the dysfunction and cancers of brain, lung and thyroid, identified amplifications, mutations or deletions in the *MBIP* gene [119–121]. As MBIP is a *bona fide* subunit of ATAC and thus seems to be an important subunit for the stability of the complex, these GWAS studies may suggest that ATAC- and/or MBIP-mediated functions are crucial for the correct homeostasis of these organs.

The *bona fide* subunit of ATAC, WDR5 (WD repeat-containing protein 5), is a highly conserved metazoan protein involved in development, epigenetic control, transcription, and cell division. WDR5 has been extensively studied and shown to be a moonlighting protein, as it is also a subunit of many functionally different multiprotein complexes involved in histone methylation (MLL/Set1 or COMPASS), histone acetylation (NSL) with positive and negative (NuRD) chromatin linked functions ([122] and refs therein). Early studies led to the suggestion that WDR5 itself might be an H3K4-methylation reader [123]. WDR5 contains WD40 repeats adopting a seven-bladed β -propeller structure (Fig. 5). Indeed, it seems that most of chromatin regulatory complexes have dedicated WD40 family members, such as for example Taf5 in yeast TFIID and SAGA, or TAF5 and TAF5L in metazoan TFIID and SAGA complexes, respectively, which all have seven β -propeller-containing domains [42]. The finding that TAF5, but also its paralogue TAF5L, interacts through their WD40 domains with the chaperonin TRiC/CCT, a large hetero-oligomeric ring-shaped complex, in the cytoplasm [124], suggests that the WD40-TRiC/CCT chaperonin interaction mechanism may be an important regulatory step during quality control and consequent biogenesis of chromatin regulatory complexes containing WD40 proteins. As WD40 propeller domain-containing proteins seem to occupy central positions in regulatory complexes, in addition they may also serve as structural hubs connecting many subunits of a given complex. The WD40 domain of Taf5 in yeast SAGA is a good example of this capacity, as it associates with at least ten different protein domains [52].

All ATAC2 (also called Cysteine-rich protein 2-binding protein or CRP2BP) homologues contain an acetyltransferase domain related to the GNAT family (Fig. 5). In addition, insect ATAC2 homologues contain a PhD domain at their N-terminus. However, the alignment of about the first 200 amino acids of insect and vertebrate ATAC2 proteins indicates that the vertebrate proteins do not contain a typical PhD domain at this position, but rather a PhD-like zinc-finger domain [103]. Indeed, this PhD-like domain was later found to be associated with a winged-helix (WH) domain in different protein families, called PWAPA

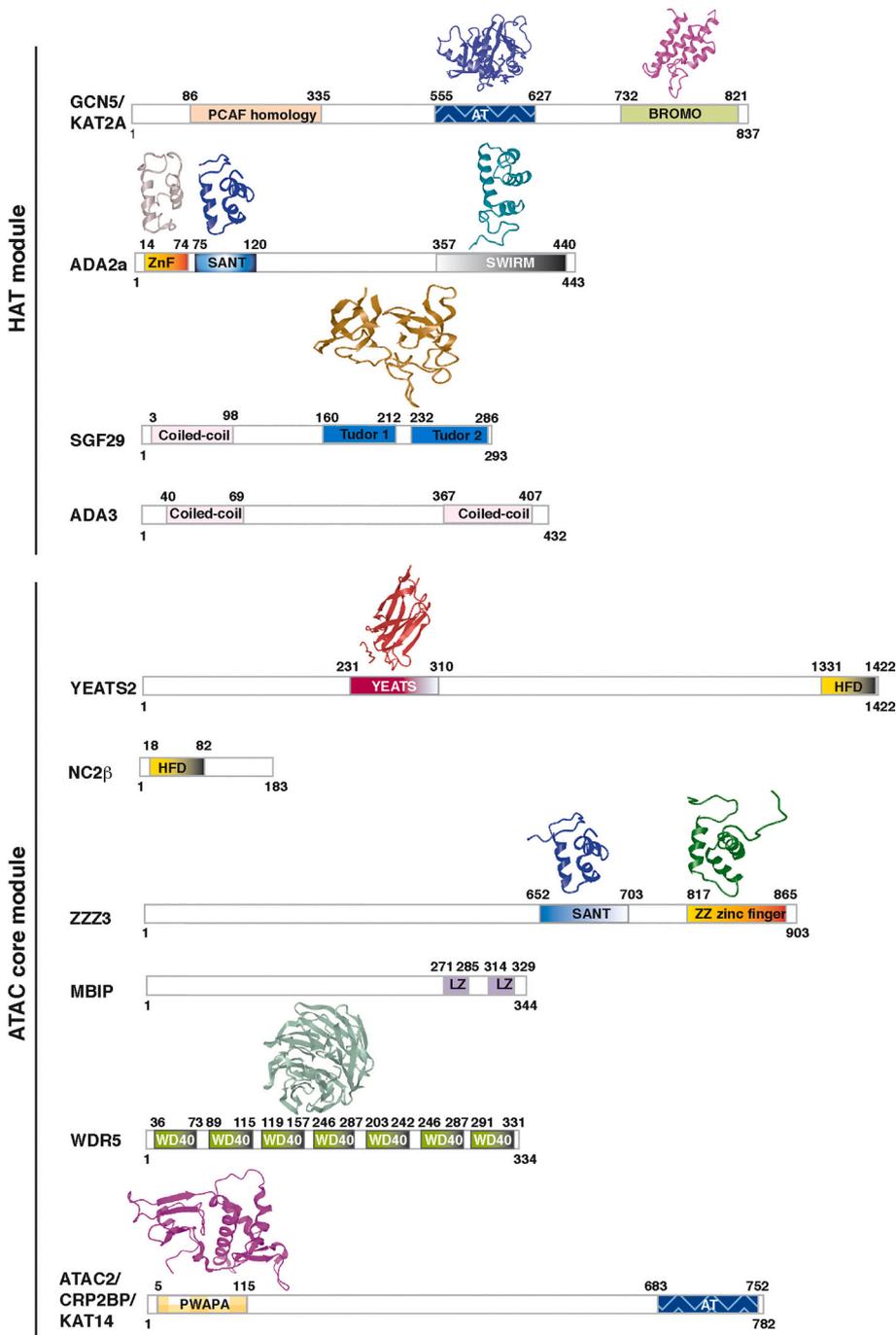


Fig. 5. Schematic representation of the ten subunits of the human ATAC complex. Amino acid coordinates of each factor are labelled together with the different functional domains described in each factor. The known high-resolution crystal structure of the different domains of the corresponding factors is indicated. The PDB coordinates of the different domains identified in ATAC subunits or homologues domains are the following: acetyl transferase domain (AT): 1Z4R; bromo domain: 3D7C; SWIRM: 2AQF; Ada2a ZZ finger (ZnF): 1X41; ZZZ3 domain: 2YUM; double Tudor domain: 3MEW; SANT domain: 2EQR; the 7 WD40 repeats of WDR5: 2H9M; YEATS2 domain: 5IQL; PWAPA domain: 3RSN.

domain for PhD-like/WH domain in human *ASH2* histone methyltransferase complex subunit, human polycomb-like proteins 1–3 (*PCL1–3*) and human *ATAC2* [125] (Fig. 5). Furthermore, it has been suggested that PWAPA domain-containing proteins may recognize modified histone tails and/or may function as DNA-binding domains. However, the specificity of these activities remains to be discovered. Vertebrate *ATAC2* (*CRP2BP*) homologues were suggested to contain also an IMP dehydrogenase (*IMPDH*)/GMP reductase domain, but no functional experiments have been carried out with this putative homology domain [126]. Using distinct recombinant *ATAC2/CRP2BP* protein preparations, different authors reported that *ATAC2/CRP2BP* acetylates i) all core histones with H3 being the most prominent [126], ii) H4 [110], iii) H4 on K5 [127], or iv) did not exhibit any significant HAT activity [103,108]. Thus, at present the histone acetylation specificity of *ATAC2/CRP2BP* in isolation is not well understood, however

its HAT activity may be stimulated and/or the specificity modified by its incorporation into the ATAC complex. In line with this suggestion the acetylation of histone H3K9, H4K5, H4K12, and H4K16 was reduced to at least 50% in human cells in which *ATAC2/CRP2BP* has been knocked down. Note however that in these knock-down cells it has not been investigated whether and how the composition of the ATAC complex has been altered and whether the *ADA2a*-containing HAT module is still active. Moreover, it has been recently reported that actin binds directly to *ATAC2/CRP2BP*, and thus may modulate its histone acetyl transferase activity *in vitro* and in cells [127]. As a matter of fact, *CRP2BP* has also been described as a cofactor of the serum response factor to regulate smooth muscle gene expression [126].

4. Concluding remarks

In all species examined so far, Gcn5/KAT2A is part of several multisubunit complexes that share a common tetrameric HAT module. Of these, SAGA is the only one for which we now have a high-resolution view of its structural organization, although, unluckily, its HAT module still remains blurry. The topology of the different functional modules is now well understood in yeast. SAGA is organized around a central core module, to which the other modules are connected with varying strengths. The core module is organized around a Taf5-Taf6 heterodimer that has a prominent scaffolding role for the entire complex. The Taf5-Taf6 heterodimer anchors the HAT module and contacts the DUB module and a histone octamer-like structure. This octamer has crucial roles in the interaction and positioning of Spt20, which binds Tra1, and Spt3, which binds TBP.

One striking observation from these studies is that the TBP-loading and activator-binding modules of SAGA appear stably positioned relative to the core. In marked contrast, the two enzymatic modules are extremely flexible, suggesting that substrate recognition and/or catalytic activity involve structural rearrangements. Specifically, the conformation of the DUB is itself rather stable, but might require flexibility to reach a nucleosome, perhaps to accommodate the highly variable lengths of nucleosome-free regions. The HAT module shows massive structural heterogeneity, both between its four subunits and relative to the rest of SAGA. The molecular nature and functional relevance of this conformational variability remains unknown and is an important question for future studies of SAGA mechanism of action.

Interestingly, a recent structure of the COMPASS complex, which catalyzes H3K4 methylation of H2B mono-ubiquitinated nucleosomes, demonstrated the importance of structural rearrangements for substrate recognition and catalysis [128]. As compared to the structure of COMPASS alone, Worden et al. [128] observed conformational changes for most subunits upon binding to mono-ubiquitinated nucleosomes. For example, several residues from the N-terminal N-Set domain of the Set1 catalytic subunit switched from a disordered conformation to a helix, which makes critical contacts with the acidic surface of H2A on one side and with ubiquitin on the other side.

An integrated view of the structure of all SAGA modules at different stages of transcription will be crucial to answer long-standing questions about multifunctional co-activator complexes. In particular, despite much work over the past two decades, little is known about the physiological relevance of SAGA conformational heterogeneity and whether allosteric regulation exists within the complex. Rather, reductionist approaches suggest that each module can function independently of the others in gene regulation. However, the evolutionary conservation of SAGA composition argues for the importance of associating different activities, which appear to synergistically control global RNA polymerase II-dependent transcription [129].

The overlapping as well as differential HAT activities carried by the related HAT modules of either SAGA or ATAC, and the role of metazoan SAGA and ATAC complexes regulating potentially distinct gene regulatory pathways involved in signaling, development or cell cycle progression has been elegantly reviewed [130]. Although much progress has been made, many functional aspects carried out by related GCN5-containing complexes remain unanswered. In order to better understand the functional differences and similarities between these complexes, several questions require further investigation. For example, we still do not understand how the same HAT, Gcn5/KAT2A, can have different activities, depending on whether it is incorporated into SAGA, ATAC or other related complexes. ADA2a and ADA2b play probably an important role in this specialization, but the molecular mechanisms involved remain unknown. Hopefully, within a few years, continuing progress in structural and functional analyses will continue to bring answers to these questions.

Declaration of competing interest

None.

Acknowledgements

This manuscript is dedicated to Susan Abmayr, whose important contribution to the above discussed topics is unforgettable. We thank P. Schultz and A. Ben-Shem for stimulating and insightful discussions. We apologize to our colleagues whose work could not be cited due to space limitations. This study was supported by NIH RO1 grant (1R01GM131626-01 to L.T.), by the Agence Nationale de la Recherche (ANR-15-CE12-0009-01 and ANR-15-CE11-0022-03 to D.H., AAPG2019 PICen to G.P. and L.T., ANR-PRCI-AAPG2019 EpiCAST to L.T., ANR-18-CE12-0026 to D.D.), grant ANR-10-LABX-0030-INRT and a French State fund managed by the ANR under the frame program Investissements d'Avenir ANR-10-IDEX-0002-02 to IGBMC.

Author contributions

DH, PG, DD, and LT wrote the review. DH made the Table. PG, DD, and LT made the figures following advice from DH.

References

- J.E. Brownell, J. Zhou, T. Ranalli, R. Kobayashi, D.G. Edmondson, S.Y. Roth, C.D. Allis, Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation, *Cell* 84 (1996) 843–851.
- M.-H. Kuo, J.E. Brownell, R.E. Sobel, T.A. Ranalli, R.G. Cook, D.G. Edmondson, S.Y. Roth, C.D. Allis, Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines, *Nature* 383 (1996) 269–272.
- P.A. Grant, L. Duggan, J. Cote, S.M. Roberts, J.E. Brownell, R. Candau, R. Ohba, T. Owen-Hughes, C.D. Allis, F. Winston, S.L. Berger, J.L. Workman, Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex, *Genes Dev.* 11 (1997) 1640–1650.
- J. Horiuchi, N. Silverman, B. Piña, G.A. Marcus, L. Guarente, ADA1, a novel component of the ADA/GCN5 complex, has broader effects than GCN5, ADA2, or ADA3, *Mol. Cell. Biol.* 17 (1997) 3220–3228.
- A. Saleh, V. Lang, R. Cook, C.J. Brandl, Identification of native complexes containing the yeast coactivator/repressor proteins NGG1/ADA3 and ADA2, *J. Biol. Chem.* 272 (1997) 5571–5578.
- S.L. Berger, B. Pina, N. Silverman, G.A. Marcus, J. Agapite, J.L. Regier, S.J. Triezenberg, L. Guarente, Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains, *Cell* 70 (1992) 251–265.
- B. Piña, S. Berger, G.A. Marcus, N. Silverman, J. Agapite, L. Guarente, ADA3: a gene, identified by resistance to GAL4-VP16, with properties similar to and different from those of ADA2, *Mol. Cell. Biol.* 13 (1993) 5981–5989.
- G.A. Marcus, N. Silverman, S.L. Berger, J. Horiuchi, L. Guarente, Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors, *EMBO J.* 13 (1994) 4807–4815.
- G.A. Marcus, J. Horiuchi, N. Silverman, L. Guarente, ADA5/SPT20 links the ADA and SPT genes, which are involved in yeast transcription, *Mol. Cell. Biol.* 16 (1996) 3197–3205.
- J. Horiuchi, N. Silverman, G.A. Marcus, L. Guarente, ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex, *Mol. Cell. Biol.* 15 (1995) 1203–1209.
- R. Candau, S.L. Berger, Structural and functional analysis of yeast putative adaptors. Evidence for an adaptor complex in vivo, *J. Biol. Chem.* 271 (1996) 5237–5245.
- P.A. Grant, D.E. Sterner, L.J. Duggan, J.L. Workman, S.L. Berger, The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes, *Trends Cell Biol.* 8 (1998) 193–197.
- R. Balasubramanian, M.G. Pray-Grant, W. Selleck, P.A. Grant, S. Tan, Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation, *J. Biol. Chem.* 277 (2002) 7989–7995.
- C. Bian, C. Xu, J. Ruan, K.K. Lee, T.L. Burke, W. Tempel, D. Barsyte, J. Li, M. Wu, B.O. Zhou, B.E. Fleharty, A. Paulson, A. Allali-Hassani, J.Q. Zhou, G. Mer, P.A. Grant, J.L. Workman, J. Zang, J. Min, Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation, *EMBO J.* 30 (2011) 2829–2842.
- S.L. Sanders, J. Jennings, A. Canutescu, A.J. Link, P.A. Weil, Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry, *Mol. Cell. Biol.* 22 (2002) 4723–4738.
- S.M. Roberts, F. Winston, SPT20/ADA5 encodes a novel protein functionally related to the TATA-binding protein and important for transcription in

- Saccharomyces cerevisiae, *Mol. Cell. Biol.* 16 (1996) 3206–3213.
- [17] D.E. Sterner, P.A. Grant, S.M. Roberts, L.J. Duggan, R. Belotserkovskaya, L.A. Pacella, F. Winston, J.L. Workman, S.L. Berger, Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction, *Mol. Cell. Biol.* 19 (1999) 86–98.
- [18] D.M. Eisenmann, K.M. Arndt, S.L. Ricupero, J.W. Rooney, F. Winston, SPT3 interacts with TFIID to allow normal transcription in *Saccharomyces cerevisiae*, *Genes Dev.* 6 (1992) 1319–1331.
- [19] D.M. Eisenmann, C. Chapon, S.M. Roberts, C. Dollard, F. Winston, The *Saccharomyces cerevisiae* SPT8 gene encodes a very acidic protein that is functionally related to SPT3 and TATA-binding protein, *Genetics* 137 (1994) 647–657.
- [20] P.A. Grant, D. Schieltz, M.G. Pray-Grant, D.J. Steger, J.C. Reese, J.R. Yates 3rd, J.L. Workman, A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation, *Cell* 94 (1998) 45–53.
- [21] A. Hoffmann, C.-M. Chiang, T. Oelgeschläger, X. Xie, S.K. Burley, Y. Nakatani, R.G. Roeder, A histone octamer-like structure within TFIID, *Nature* 380 (1996) 356–359.
- [22] X. Xie, T. Kokubo, S.L. Cohen, U.A. Mirza, A. Hoffmann, B.T. Chait, R.G. Roeder, Y. Nakatani, S.K. Burley, Structural similarity between TAFs and the heterotetrameric core of the histone octamer, *Nature* 380 (1996) 316–322.
- [23] W. Selleck, R. Howley, Q. Fang, Y. Podolny, M.G. Fried, S. Buratowski, S. Tan, A histone fold TAF octamer within the yeast TFIID transcriptional coactivator, *Nat. Struct. Mol. Biol.* 8 (2001) 695–700.
- [24] C. Birck, O. Poch, C. Romier, M. Ruff, G. Mengus, A.-C. Lavigne, I. Davidson, D. Moras, Human TAFII28 and TAFII18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family, *Cell* 94 (1998) 239–249.
- [25] Y.G. Gangloff, S.L. Sanders, C. Romier, D. Kirschner, P.A. Weil, L. Tora, I. Davidson, 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, *Mol. Cell. Biol.* 21 (2001) 1841–1853.
- [26] Y.G. Gangloff, S. Werten, C. Romier, L. Carre, O. Poch, D. Moras, I. Davidson, The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs, *Mol. Cell. Biol.* 20 (2000) 340–351.
- [27] P.A. Grant, D. Schieltz, M.G. Pray-Grant, J.R. Yates 3rd, J.L. Workman, The ATM-related cofactor Tra1 is a component of the purified SAGA complex, *Mol. Cell* 2 (1998) 863–867.
- [28] A. Saleh, D. Schieltz, N. Ting, S.B. McMahon, D.W. Litchfield, J.R. Yates, S.P. Lees-Miller, M.D. Cole, C.J. Brandl, Tra1p is a component of the yeast AdaSpt transcriptional regulatory complexes, *J. Biol. Chem.* 273 (1998) 26559–26565.
- [29] J.A. Daniel, M.S. Torok, Z.W. Sun, D. Schieltz, C.D. Allis, J.R. Yates 3rd, P.A. Grant, Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription, *J. Biol. Chem.* 279 (2004) 1867–1871.
- [30] K.W. Henry, A. Wyce, W.S. Lo, L.J. Duggan, N.C. Emre, C.F. Kao, L. Pillus, A. Shilatfard, M.A. Osley, S.L. Berger, Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8, *Genes Dev.* 17 (2003) 2648–2663.
- [31] K. Ingvarsdottir, N.J. Krogan, N.C. Emre, A. Wyce, N.J. Thompson, A. Emili, T.R. Hughes, J.F. Greenblatt, S.L. Berger, H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex, *Mol. Cell. Biol.* 25 (2005) 1162–1172.
- [32] A. Kohler, M. Schneider, G.G. Cabal, U. Nehrbass, E. Hurt, Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export, *Nat. Cell Biol.* 10 (2008) 707–715.
- [33] Z. Nagy, L. Tora, Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation, *Oncogene* 26 (2007) 5341–5357.
- [34] X.-J. Yang, V.V. Ogryzko, J.-i. Nishikawa, B.H. Howard, Y. Nakatani, A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A, *Nature* 382 (1996) 319–324.
- [35] M. Brand, K. Yamamoto, A. Staub, L. Tora, Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction, *J. Biol. Chem.* 274 (1999) 18285–18289.
- [36] E. Wiczorek, M. Brand, X. Jacq, L. Tora, Function of TAFII-containing complex without TBP in transcription by RNA polymerase II, *Nature* 393 (1998) 187–191.
- [37] E. Martinez, T.K. Kundu, J. Fu, R.G. Roeder, A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID, *J. Biol. Chem.* 273 (1998) 23781–23785.
- [38] V.V. Ogryzko, T. Kotani, X. Zhang, R.L. Schiltz, T. Howard, X.J. Yang, B.H. Howard, J. Qin, Y. Nakatani, Histone-like TAFs within the PCAF histone acetylase complex, *Cell* 94 (1998) 35–44.
- [39] V.M. Weake, S.K. Swanson, A. Mushegian, L. Florens, M.P. Washburn, S.M. Abmayr, J.L. Workman, A novel histone fold domain-containing protein that replaces TAF6 in *Drosophila* SAGA is required for SAGA-dependent gene expression, *Genes Dev.* 23 (2009) 2818–2823.
- [40] F. Moraga, F. Aquea, Composition of the SAGA complex in plants and its role in controlling gene expression in response to abiotic stresses, *Front. Plant Sci.* 6 (2015) 865.
- [41] A. Pfab, A. Bruckmann, J. Nazet, R. Merkl, K.D. Grasser, The adaptor protein ENY2 is a component of the deubiquitination module of the Arabidopsis SAGA transcriptional co-activator complex but not of the TREX-2 complex, *J. Mol. Biol.* 430 (2018) 1479–1494.
- [42] D. Helmlinger, L. Tora, Sharing the SAGA, *Trends Biochem. Sci.* 42 (2017) 850–861.
- [43] R. Stegeman, P.J. Spreacker, S.K. Swanson, R. Stephenson, L. Florens, M.P. Washburn, V.M. Weake, The spliceosomal protein SF3B5 is a novel component of *Drosophila* SAGA that functions in gene expression independent of splicing, *J. Mol. Biol.* 428 (2016) 3632–3649.
- [44] M. Vermeulen, C.H. Eberl, F. Matarese, H. Marks, S. Denissov, F. Butter, K.K. Lee, J.V. Olsen, A.A. Hyman, H.G. Stunnenberg, M. Mann, Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers, *Cell* 142 (2010).
- [45] R. Srivastava, K.M. Rai, B. Pandey, S.P. Singh, S.V. Sawant, Spt-Ada-Gcn5-Acetyltransferase (SAGA) complex in plants: genome wide identification, evolutionary conservation and functional determination, *PLoS One* 10 (2015) e0134709.
- [46] M. Brand, C. Leurent, V. Mallouh, L. Tora, P. Schultz, Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC, *Science* 286 (1999) 2151–2153.
- [47] P.Y. Wu, C. Ruhlmann, F. Winston, P. Schultz, Molecular architecture of the *S. cerevisiae* SAGA complex, *Mol. Cell* 15 (2004) 199–208.
- [48] Y. Han, J. Luo, J. Ranish, S. Hahn, Architecture of the *Saccharomyces cerevisiae* SAGA transcription coactivator complex, *EMBO J.* 33 (2014) 2534–2546.
- [49] K.K. Lee, M.E. Sardi, S.K. Swanson, J.M. Gilmore, M. Torok, P.A. Grant, L. Florens, J.L. Workman, M.P. Washburn, Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes, *Mol. Syst. Biol.* 7 (2011) 503.
- [50] L.M. Diaz-Santín, N. Lukoyanova, E. Acıyan, A.C. Cheung, Cryo-EM structure of the SAGA and NuA4 coactivator subunit Tra1 at 3.7 angstrom resolution, *Elife* 6 (2017).
- [51] G. Sharov, K. Voltz, A. Durand, O. Kolesnikova, G. Papai, A.G. Myasnikov, A. Dejaegere, A. Shem, P. Schultz, Structure of the transcription activator target Tra1 within the chromatin modifying complex SAGA, *Nat. Commun.* 8 (2017) 1556.
- [52] G. Papai, A. Frechard, O. Kolesnikova, C. Crucifix, P. Schultz, A. Ben-Shem, Structure of SAGA and mechanism of TBP deposition on gene promoters, *Nature* 577 (2020) 711–716.
- [53] H. Wang, C. Dienemann, A. Stützer, H. Urlaub, A.C.M. Cheung, P. Cramer, Structure of the transcription coactivator SAGA, *Nature* 577 (2020) 717–720.
- [54] J. Sun, M. Paduch, S.A. Kim, R.M. Kramer, A.F. Barrios, V. Lu, J. Luke, S. Usatyuk, A.A. Kossiakoff, S. Tan, Structural basis for activation of SAGA histone acetyltransferase Gcn5 by partner subunit Ada2, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) 10010–10015.
- [55] A. Kohler, E. Zimmermann, M. Schneider, E. Hurt, N. Zheng, Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module, *Cell* 141 (2010) 606–617.
- [56] G. Lang, J. Bonnet, D. Umlauf, K. Karmodiya, J. Koffler, M. Stierle, D. Devys, L. Tora, The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements, *Mol. Cell. Biol.* 31 (2011) 3734–3744.
- [57] N.L. Samara, A.B. Datta, C.E. Berndsen, X. Zhang, T. Yao, R.E. Cohen, C. Wolberger, Structural insights into the assembly and function of the SAGA deubiquitinating module, *Science* 328 (2010) 1025–1029.
- [58] K.J. Pollard, C.L. Peterson, Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression, *Mol. Cell. Biol.* 17 (1997) 6212–6222.
- [59] N. Vosnakis, M. Koch, E. Scheer, P. Kessler, Y. Mély, P. Didier, L. Tora, Coactivators and general transcription factors have two distinct dynamic populations dependent on transcription, *EMBO J.* 36 (2017) 2710–2725.
- [60] C. Dhalluin, J.E. Carlson, L. Zeng, C. He, A.K. Aggarwal, M.M. Zhou, Structure and ligand of a histone acetyltransferase bromodomain, *Nature* 399 (1999) 491–496.
- [61] B.P. Hudson, M.A. Martinez-Yamout, H.J. Dyson, P.E. Wright, Solution structure and acetyl-lysine binding activity of the GCN5 bromodomain, *J. Mol. Biol.* 304 (2000) 355–370.
- [62] D.J. Owen, P. Ornaghi, J.C. Yang, N. Lowe, P.R. Evans, P. Ballarío, D. Neuhaus, P. Filetici, A.A. Travers, The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p, *EMBO J.* 19 (2000) 6141–6149.
- [63] X.-J. Yang, Lysine acetylation and the bromodomain: a new partnership for signaling, *Bioessays* 26 (2004) 1076–1087.
- [64] P. Syntichaki, I. Topalidou, G. Thireos, The Gcn5 bromodomain co-ordinates nucleosome remodelling, *Nature* 404 (2000) 414–417.
- [65] N.T. Nguyen-Huynh, G. Sharov, C. Potel, P. Fichter, S. Trowitzsch, I. Berger, V. Lamour, P. Schultz, N. Potier, E. Leize-Wagner, Chemical cross-linking and mass spectrometry to determine the subunit interaction network in a recombinant human SAGA HAT subcomplex, *Protein Sci.* 24 (2015) 1232–1246.
- [66] K.K. Lee, L. Florens, S.K. Swanson, M.P. Washburn, J.L. Workman, The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex, *Mol. Cell. Biol.* 25 (2005) 1173–1182.
- [67] N.L. Samara, A.E. Ringel, C. Wolberger, A role for intersubunit interactions in maintaining SAGA deubiquitinating module structure and activity, *Structure* 20 (2012) 1414–1424.
- [68] M. Yan, C. Wolberger, Uncovering the role of Sgf73 in maintaining SAGA deubiquitinating module structure and activity, *J. Mol. Biol.* 427 (2015) 1765–1778.
- [69] C. Koehler, J. Bonnet, M. Stierle, C. Romier, D. Devys, B. Kieffer, DNA binding by Sgf11 protein affects histone H2B deubiquitination by Spt-Ada-Gcn5-acetyltransferase (SAGA), *J. Biol. Chem.* 289 (2014) 8989–8999.
- [70] M.T. Morgan, M. Haj-Yahya, A.E. Ringel, P. Bandi, A. Brik, C. Wolberger, Structural basis for histone H2B deubiquitination by the SAGA DUB module, *Science* 351 (2016) 725–728.

- [71] S. Lim, J. Kwak, M. Kim, D. Lee, Separation of a functional deubiquitylating module from the SAGA complex by the proteasome regulatory particle, *Nat. Commun.* 4 (2013) 2641.
- [72] X. Li, C.W. Seidel, L.T. Szerszen, J.J. Lange, J.L. Workman, S.M. Abmayr, Enzymatic modules of the SAGA chromatin-modifying complex play distinct roles in *Drosophila* gene expression and development, *Genes Dev.* 31 (2017) 1588–1600.
- [73] P. Georgakopoulos, R.A. Lockington, J.M. Kelly, The Spt-Ada-Gcn5 Acetyltransferase (SAGA) complex in *Aspergillus nidulans*, *PLoS One* 8 (2013) e65221.
- [74] S.M. Rosler, K. Kramer, I. Finkemeier, H.U. Humpf, B. Tudzynski, The SAGA complex in the rice pathogen *Fusarium fujikuroi*: structure and functional characterization, *Mol. Microbiol.* 102 (2016) 951–974.
- [75] B.S. Atanassov, R.D. Mohan, X. Lan, X. Kuang, Y. Lu, K. Lin, E. McIvor, W. Li, Y. Zhang, L. Florens, S.D. Byrum, S.G. Mackintosh, T. Calhoun-Davis, E. Koutelou, L. Wang, D.G. Tang, A.J. Tackett, M.P. Washburn, J.L. Workman, S.Y. Dent, ATXN7L3 and ENY2 coordinate activity of multiple H2B deubiquitinases important for cellular proliferation and tumor growth, *Mol. Cell* 62 (2016) 558–571.
- [76] K.K. Lee, S.K. Swanson, L. Florens, M.P. Washburn, J.L. Workman, Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik (SALSA) HAT complexes, *Epigenetics Chromatin* 2 (2009) 2.
- [77] J. Bonnet, Y.-H.H. Wang, G. Spedale, R.A. Atkinson, C. Romier, A. Hamiche, W.W. Pijnappel, H.T. Timmers, L. Tora, D. Devys, B. Kieffer, The structural plasticity of SCA7 domains defines their differential nucleosome-binding properties, *EMBO Rep.* 11 (2010) 612–618.
- [78] K. Kamata, A. Hatanaka, G. Goswami, K. Shinmyozu, J. Nakayama, T. Urano, M. Hatashita, H. Uchida, M. Oki, C-terminus of the Sgf73 subunit of SAGA and SLIK is important for retention in the larger complex and for heterochromatin boundary function, *Genes Cells* 18 (2013) 823–837.
- [79] S.B. McMahon, H.A.V. Buskirk, K.A. Dugan, T.D. Copeland, M.D. Cole, The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins, *Cell* 94 (1998) 363–374.
- [80] S. Allard, R.T. Utley, J. Savard, A. Clarke, P. Grant, C.J. Brandl, L. Pillus, J.L. Workman, J. Cote, NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p, *EMBO J.* 18 (1999) 5108–5119.
- [81] A. Elias-Villalobos, P. Fort, D. Helmlinger, New insights into the evolutionary conservation of the sole PIKK pseudokinase Tra1/TRRAP, *Biochem. Soc. Trans.* 47 (2019) 1597–1608.
- [82] Z. Herceg, W. Hulla, D. Gell, C. Cuenin, M. Leonart, S. Jackson, Z.Q. Wang, Disruption of Trrap causes early embryonic lethality and defects in cell cycle progression, *Nat. Genet.* 29 (2001) 206–211.
- [83] C.E. Brown, L. Howe, K. Sousa, S.C. Alley, M.J. Carrozza, S. Tan, J.L. Workman, Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit, *Science* 292 (2001) 2333–2337.
- [84] C. Bouchard, O. Dittrich, A. Kiermaier, K. Dohmann, A. Menkel, M. Eilers, B. Lüscher, Regulation of cyclin D2 gene expression by the Myc/Max/Mad network: Myc-dependent TRRAP recruitment and histone acetylation at the cyclin D2 promoter, *Genes Dev.* 15 (2001) 2042–2047.
- [85] J. Park, S. Kunjibettu, S.B. McMahon, M.D. Cole, The ATM-related domain of TRRAP is required for histone acetyltransferase recruitment and Myc-dependent oncogenesis, *Genes Dev.* 15 (2001) 1619–1624.
- [86] J. Klein, M. Nolden, S.L. Sanders, J. Kirchner, P.A. Weil, K. Melcher, Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA, *J. Biol. Chem.* 278 (2003) 6779–6786.
- [87] S.R. Bhaumik, T. Raha, D.P. Aiello, M.R. Green, In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer, *Genes Dev.* 18 (2004) 333–343.
- [88] J. Fishburn, N. Mohibullah, S. Hahn, Function of a eukaryotic transcription activator during the transcription cycle, *Mol. Cell* 18 (2005) 369–378.
- [89] W.M. Reeves, S. Hahn, Targets of the Gal4 transcription activator in functional transcription complexes, *Mol. Cell Biol.* 25 (2005) 9092–9102.
- [90] B.A. Knutson, S. Hahn, Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes, *Mol. Cell Biol.* 31 (2011) 818–831.
- [91] D. Helmlinger, S. Marguerat, J. Villen, D.L. Swaney, S.P. Gygi, J. Bahler, F. Winston, Tra1 has specific regulatory roles, rather than global functions, within the SAGA co-activator complex, *EMBO J.* 30 (2011) 2843–2852.
- [92] A. Elias-Villalobos, D. Toullec, C. Faux, M. Seveno, D. Helmlinger, Chaperone-mediated ordered assembly of the SAGA and NuA4 transcription co-activator complexes in yeast, *Nat. Commun.* 10 (2019) 5237.
- [93] X. Wang, S. Ahmad, Z. Zhang, J. Cote, G. Cai, Architecture of the *Saccharomyces cerevisiae* NuA4/TIP60 complex, *Nat. Commun.* 9 (2018) 1147.
- [94] A.B. Patel, R.K. Louder, B.J. Greber, S. Grünberg, J. Luo, J. Fang, Y. Liu, J. Ranish, S. Hahn, E. Nogales, Structure of human TFIIID and mechanism of TBP loading onto promoter DNA, *Science* 362 (2018) eaau8872.
- [95] A. Eberharther, D.E. Sterner, D. Schieltz, A. Hassan, J.R. Yates, S.L. Berger, J.L. Workman, The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*, *Mol. Cell Biol.* 19 (1999) 6621–6631.
- [96] P.A. Grant, A. Eberharther, S. John, R.G. Cook, B.M. Turner, J.L. Workman, Expanded lysine acetylation specificity of Gcn5 in native complexes, *J. Biol. Chem.* 274 (1999) 5895–5900.
- [97] J.H.M. Soffers, X. Li, A. Saraf, C.W. Seidel, L. Florens, M.P. Washburn, S.M. Abmayr, J.L. Workman, Characterization of a metazoan ADA acetyltransferase complex, *Nucleic Acids Res.* 47 (2019) 3383–3394.
- [98] E.F. Torres-Zelada, R.E. Stephenson, A. Alpsy, B.D. Anderson, S.K. Swanson, L. Florens, E.C. Dykhuizen, M.P. Washburn, V.M. Weake, The *Drosophila* Dbf4 ortholog Chiffon forms a complex with Gcn5 that is necessary for histone acetylation and viability, *J. Cell Sci.* 132 (2018) jcs.214072.
- [99] T. Pankotai, N. Zsindely, E.E. Vamos, O. Komonyi, L. Bodai, I.M. Boros, Functional characterization and gene expression profiling of *Drosophila melanogaster* short dADA2b isoform-containing dSAGA complexes, *BMC Genomics* 14 (2013) 44.
- [100] D. Qi, J. Larsson, M. Mannervik, *Drosophila* Ada2b is required for viability and normal histone H3 acetylation, *Mol. Cell Biol.* 24 (2004) 8080–8089.
- [101] S. Guelman, T. Suganuma, L. Florens, S.K. Swanson, C.L. Kiesecker, T. Kusch, S. Anderson, J.R. Yates 3rd, M.P. Washburn, S.M. Abmayr, J.L. Workman, Host cell factor and an uncharacterized SANT domain protein are stable components of ATAC, a novel dADA2A/dGcn5-containing histone acetyltransferase complex in *Drosophila*, *Mol. Cell Biol.* 26 (2006) 871–882.
- [102] Y.-L. Wang, F. Faiola, M. Xu, S. Pan, E. Martinez, Human ATAC is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein, *J. Biol. Chem.* 283 (2008) 33808–33815.
- [103] Z. Nagy, A. Riss, S. Fujiyama, A. Krebs, M. Orpinell, P. Jansen, A. Cohen, H.G. Stunnenberg, S. Kato, L. Tora, The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes, *Cell. Mol. Life Sci.* 67 (2010) 611–628.
- [104] T. Kusch, S. Guelman, S.M. Abmayr, J.L. Workman, Two *Drosophila* Ada2 homologues function in different multiprotein complexes, *Mol. Cell Biol.* 23 (2003) 3305–3319.
- [105] S. Muratoglu, S. Georgieva, G. Papai, E. Scheer, I. Enunlu, O. Komonyi, I. Cserpan, L. Ledebeva, E. Nabirochikina, A. Udvardy, L. Tora, I. Boros, Two different *Drosophila* ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes, *Mol. Cell Biol.* 23 (2003) 306–321.
- [106] Q. Jin, L.R. Yu, L. Wang, Z. Zhang, L.H. Kasper, J.E. Lee, C. Wang, P.K. Brindle, S.Y. Dent, K. Ge, Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation, *EMBO J.* 30 (2011) 249–262.
- [107] J. Bonnet, D. Devys, L. Tora, Histone H2B ubiquitination: signaling not scrapping, *Drug Discov. Today Technol.* 12 (2014) e19–e27.
- [108] S. Guelman, K. Kozuka, Y. Mao, V. Pham, M.J. Solloway, J. Wang, J. Wu, J.R. Lill, J. Zha, The double-histone-acetyltransferase complex ATAC is essential for mammalian development, *Mol. Cell Biol.* 29 (2009) 1176–1188.
- [109] T. Pankotai, O. Komonyi, L. Bodai, Z. Ujfaludi, S. Muratoglu, A. Ciurciu, L. Tora, J. Szabad, I. Boros, The homologous *Drosophila* transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions, *Mol. Cell Biol.* 25 (2005) 8215–8227.
- [110] T. Suganuma, J.L. Gutiérrez, B. Li, L. Florens, S.K. Swanson, M.P. Washburn, S.M. Abmayr, J.L. Workman, ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding, *Nat. Struct. Mol. Biol.* 15 (2008) 364–372.
- [111] A. Ciurciu, O. Komonyi, T. Pankotai, I.M. Boros, The *Drosophila* histone acetyltransferase Gcn5 and transcriptional adaptor ADA2a are involved in nucleosomal histone H4 acetylation, *Mol. Cell Biol.* 26 (2006) 9413–9423.
- [112] A.M. Gamper, J. Kim, R.G. Roeder, The STAGA subunit ADA2b is an important regulator of human GCN5 catalysis, *Mol. Cell Biol.* 29 (2009) 266–280.
- [113] A. Riss, E. Scheer, M. Joint, S. Trowitzsch, I. Berger, L. Tora, Subunits of ADA-two-A-containing (ATAC) or Spt-Ada-Gcn5-acetyltransferase (SAGA) coactivator complexes enhance the acetyltransferase activity of GCN5, *J. Biol. Chem.* 290 (2015) 28997–29009.
- [114] W. Mi, H. Guan, J. Lyu, D. Zhao, Y. Xi, S. Jiang, F.H. Andrews, X. Wang, M. Gagea, H. Wen, L. Tora, S.Y.R. Dent, T.G. Kutateladze, W. Li, H. Li, X. Shi, YEATS2 links histone acetylation to tumorigenesis of non-small cell lung cancer, *Nat. Commun.* 8 (2017) 1088.
- [115] W. Mi, Y. Zhang, J. Lyu, X. Wang, Q. Tong, D. Peng, Y. Xue, A.H. Tencer, H. Wen, W. Li, T.G. Kutateladze, X. Shi, The ZZ-type zinc finger of ZZZ3 modulates the ATAC complex-mediated histone acetylation and gene activation, *Nat. Commun.* 9 (2018) 3759.
- [116] D. Zhao, H. Guan, S. Zhao, W. Mi, H. Wen, Y. Li, Y. Zhao, C.D. Allis, X. Shi, H. Li, YEATS2 is a selective histone crotonylation reader, *Cell Res.* 26 (2016) 629–632.
- [117] K. Fukuyama, M. Yoshida, A. Yamashita, T. Deyama, M. Baba, A. Suzuki, H. Mohri, Z. Ikezawa, H. Nakajima, S. Hirai, S. Ohno, MAPK upstream kinase (MUK)-binding inhibitory protein, a negative regulator of MUK/dual leucine zipper-bearing kinase/leucine zipper protein kinase, *J. Biol. Chem.* 275 (2000) 21247–21254.
- [118] T. Suganuma, A. Mushegian, S.K. Swanson, L. Florens, M.P. Washburn, J.L. Workman, A metazoan ATAC acetyltransferase subunit that regulates mitogen-activated protein kinase signaling is related to an ancient molybdopterin synthase component, *Mol. Cell. Proteomics* 11 (2012) 90–99.
- [119] T. Harris, Q. Pan, J. Sironi, D. Lutz, J. Tian, J. Sapkar, R. Perez-Soler, S. Keller, J. Locker, Both gene amplification and allelic loss occur at 14q13.3 in lung cancer, *Clin. Cancer Res.* 17 (2010) 690–699.
- [120] Y. Hwangbo, Y.J. Park, Genome-wide association studies of autoimmune thyroid diseases, thyroid function, and thyroid cancer, *Endocrinol. Metab. (Seoul)* 33 (2018) 175–184.
- [121] E. Porcu, M. Medici, G. Pistis, C.B. Volpato, S.G. Wilson, A.R. Cappola, S.D. Bos, J. Deelen, M. den Heijer, R.M. Freathy, J. Lahti, C. Liu, L.M. Lopez, I.M. Nolte, J.R. O'Connell, T. Tanaka, S. Trompet, A. Arnold, S. Bandinelli, M. Beekman, S. Bohringer, S.J. Brown, B.M. Buckley, C. Camaschella, A.J. de Craen, G. Davies, M.C. de Visser, I. Ford, T. Forsen, T.M. Frayling, L. Fugazzola, M. Gogele, A.T. Hattersley, A.R. Hermus, A. Hofman, J.J. Houwing-Duistermaat, R.A. Jensen, E. Kajantie, M. Kloppenburg, E.M. Lim, C. Masciullo, S. Mariotti, C. Minelli, B.D. Mitchell, R. Nagaraja, R.T. Netea-Maier, A. Palotie, L. Persani, G.M. Piras, B.M. Psaty, K. Raikonen, J.B. Richards, F. Rivadeneira, C. Sala, M.M. Sabra,

- N. Sattar, B.M. Shields, N. Soranzo, J.M. Starr, D.J. Stott, F.C. Sweep, G. Usala, M.M. van der Klauw, D. van Heemst, A. van Mullem, S.H. Vermeulen, W.E. Visser, J.P. Walsh, R.G. Westendorp, E. Widen, G. Zhai, F. Cucca, I.J. Deary, J.G. Eriksson, L. Ferrucci, C.S. Fox, J.W. Jukema, L.A. Kiemeny, P.P. Pramstaller, D. Schlessinger, A.R. Shuldiner, E.P. Slagboom, A.G. Uitterlinden, B. Vaidya, T.J. Visser, B.H. Wolffenbuttel, I. Meulenbelt, J.I. Rotter, T.D. Spector, A.A. Hicks, D. Toniolo, S. Sanna, R.P. Peeters, S. Naitza, A meta-analysis of thyroid-related traits reveals novel loci and gender-specific differences in the regulation of thyroid function, *PLoS Genet.* 9 (2013) e1003266.
- [122] A.D. Guarnaccia, W.P. Tansey, Moonlighting with WDR5: a cellular multitasker, *J. Clin. Med.* 7 (2018) 21.
- [123] J. Wysocka, T. Swigut, T.A. Milne, Y. Dou, X. Zhang, A.L. Burlingame, R.G. Roeder, A.H. Brivanlou, C.D. Allis, WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development, *Cell* 121 (2005) 859–872.
- [124] S.V. Antonova, M. Haffke, E. Corradini, M. Mikuciunas, T.Y. Low, L. Signor, R.M. van Es, K. Gupta, E. Scheer, H.R. Vos, L. Tora, A.J.R. Heck, H.T.M. Timmers, I. Berger, Chaperonin CCT checkpoint function in basal transcription factor TFIID assembly, *Nat. Struct. Mol. Biol.* 25 (2018) 1119–1127.
- [125] I. Callebaut, J.P. Mornon, The PWAPA cassette: intimate association of a PHD-like finger and a winged-helix domain in proteins included in histone-modifying complexes, *Biochimie* 94 (2012) 2006–2012.
- [126] Y. Ma, Q. Li, A. Li, Y. Wei, P. Long, X. Jiang, F. Sun, R. Weiskirchen, B. Wu, C. Liang, J. Grötzinger, Y. Wei, W. Yu, M. Mercola, Y. Huang, J. Wang, Y. Yu, R.J. Schwartz, The CSR2BP histone acetyltransferase drives smooth muscle gene expression, *Nucleic Acids Res.* 45 (2016) 3046–3058.
- [127] T. Viita, S. Kyheroinen, B. Prajapati, J. Virtanen, M.J. Frilander, M. Varjosalo, M.K. Vartiainen, Nuclear actin interactome analysis links actin to KAT14 histone acetyl transferase and mRNA splicing, *J. Cell Sci.* 132 (2019).
- [128] E.J. Worden, X. Zhang, C. Wolberger, Structural basis for COMPASS recognition of an H2B-ubiquitinated nucleosome, *eLife* 9 (2020) e53199.
- [129] T. Baptista, S. Grunberg, N. Minoungou, M.J.E. Koster, H.T.M. Timmers, S. Hahn, D. Devys, L. Tora, SAGA is a general cofactor for RNA polymerase II transcription, *Mol. Cell* 68 (2017) 130–143 (e135).
- [130] G. Spedale, H.T. Timmers, W.W. Pijnappel, ATAC-king the complexity of SAGA during evolution, *Genes Dev.* 26 (2012) 527–541.
- [131] D. Setiaputra, J.D. Ross, S. Lu, D.T. Cheng, M.Q. Dong, C.K. Yip, Conformational flexibility and subunit arrangement of the modular yeast Spt-Ada-Gcn5 acetyltransferase complex, *J. Biol. Chem.* 290 (2015) 10057–10070.
- [132] S.V. Antonova, J. Boeren, H.T.M. Timmers, B. Snel, Epigenetics and transcription regulation during eukaryotic diversification: the saga of TFIID, *Genes Dev.* 33 (2019) 888–902.