



**HAL**  
open science

# Architecture of the multi-functional SAGA complex and the molecular mechanism of holding TBP

Adam Ben-Shem, Gabor Papai, Patrick Schultz

## ► To cite this version:

Adam Ben-Shem, Gabor Papai, Patrick Schultz. Architecture of the multi-functional SAGA complex and the molecular mechanism of holding TBP. *FEBS Journal*, 2021, 288 (10), pp.3135-3147. 10.1111/febs.15563 . hal-03454181

**HAL Id: hal-03454181**

**<https://hal.science/hal-03454181>**

Submitted on 29 Nov 2021

**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.

## TITLE

**Architecture of the multi-functional SAGA complex and the molecular mechanism of holding TBP**

## RUNNING TITLE

**Structure of yeast SAGA**

Adam Ben-Shem<sup>1,2,3</sup>, Gabor Papai<sup>1,2,3</sup> and Patrick Schultz<sup>1,2,3,4</sup>

<sup>1</sup> Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France,

<sup>2</sup> Centre National de la Recherche Scientifique, UMR7104, Institut National de la Santé et de la Recherche Médicale, U1258, Université de Strasbourg.

<sup>3</sup> Equipe labellisée Ligue Contre le Cancer.

<sup>4</sup> Corresponding author. Patrick Schultz ([patrick.schultz@iqbmc.fr](mailto:patrick.schultz@iqbmc.fr))

## ABBREVIATIONS

ChIP	chromatin immunoprecipitation
Cryo-EM	cryo electron microscopy
DUB	deubiquitination
FAT	FRAP, ATM, TRRAP
GTF	general transcription factor
HAT	histone acetyl transferase
HEAT	Huntingtin, elongation factor 3 ( <b>E</b> F3), protein phosphatase 2A ( <b>P</b> P2 <b>A</b> ), and the yeast kinase <b>I</b> TOR1
HF	histone fold
HIT	head interaction with Tra1
NTD	N-terminal domain
PIC	preinitiation complex
PIKK	Phosphatidylyl 3 kinase - related kinase
PoI II	RNA polymerase II
SAGA	Spt-Ada-Gcn5-acetyltransferase
SLIK	SAGA-like
TAF	TBP associated factor

TAND Taf1 N-terminal Domain  
TBP TATA-box binding protein  
TSS transcription start site  
UAS upstream activating sequences

## **KEYWORDS**

Transcription, co-activators, SAGA, TBP loading onto promoters, cryo-EM

## **CONFLICTS OF INTEREST**

The authors declare no conflict of interest

## **ABSTRACT**

In Eukaryotes transcription of protein encoding genes is initiated by the controlled deposition of the TATA-box binding protein TBP onto gene promoters, followed by the ordered assembly of a preinitiation complex. The SAGA co-activator is a 19-subunit complex that stimulates transcription by the action of two chromatin-modifying enzymatic modules, a transcription activator binding module, and by delivering TBP. Recent cryo electron microscopy structures of yeast SAGA with bound nucleosome or TBP reveal the architecture of the different functional domains of the co-activator. An octamer of histone-fold domains is found at the core of SAGA. This octamer, which deviates considerably from the symmetrical analogue forming the nucleosome, establishes a peripheral site for TBP binding where steric hindrance represses interaction with spurious DNA. The structures point to a mechanism for TBP delivery and release from SAGA that requires TFIIA and whose efficiency correlates with the affinity of DNA to TBP. These results provide a structural basis for understanding specific TBP delivery onto gene promoters and the role played by SAGA in regulating gene expression. The properties of the TBP-delivery machine harbored by SAGA are compared with the TBP loading device present in the TFIID complex and show multiple similitudes.

## **INTRODUCTION**

The TATA-box Binding Protein (TBP) is universally required for the synthesis of RNA molecules encoded by eukaryotic genes [1]. The interaction of TBP with the gene promoter initiates the ordered assembly of a transcription Pre-Initiation Complex (PIC) that positions the RNA polymerase enzyme over the transcription start site (TSS). For protein encoding genes, seven General Transcription Factors (GTF) compose the PIC together with RNA polymerase II (Pol II) [2]. Regulated TBP delivery, turnover and dissociation from gene promoters are key steps that modulate PIC assembly and gene expression. To adapt mRNA synthesis levels to stress or environmental cues, sequence-dependent activator proteins bind upstream of the

core promoter and recruit multi-functional co-activator complexes. Co-activators modify the structure of chromatin, act as bridging or stabilizing factors with GTFs, and help depositing TBP on gene promoters.

Specific TBP-containing complexes initiate transcription for each class of RNA polymerases [3], inhibit transcription [4], or dissociate TBP from the promoter [5]. The TFIID and SAGA co-activators facilitate Pol II transcription by loading TBP onto gene promoters [6, 7]. TBP is embedded in the TFIID complex, together with 14 TBP Associated Factors (TAFs), but is unable to bind its target TATA-box and to initiate transcription [8]. These activities are restored by the action of the general transcription factor TFIIA. This auto-inhibition mechanism was attributed to the N-terminal TAND regions of the Taf1 subunit which interact with the DNA binding cleft of TBP (TAND1) and its TFIIA interaction site (TAND2) [9]. TFIIA and promoter DNA compete for these sites to relieve the promoter-binding inhibition. Genetic screens, biochemical analysis and genome-wide transcription assays showed that the SAGA (Spt-Ada-Gcn5-Acetyltransferase) complex also binds TBP [10-12]. SAGA is recruited to gene promoters by activators and facilitates the binding of TBP to the promoter. Initially described as influencing a subset of stress-responsive gene [13], SAGA was recently shown to act positively on the transcriptional outcome of most yeast genes [12].

Progresses in single particle electron cryogenic microscopy (cryo-EM) allowed important progresses in elucidating the structure of SAGA. The aim of this review is to summarize the recent cryo-EM studies of *S cerevisiae* SAGA in the presence of a nucleosome [14] and of *P. pastoris* SAGA bound to TBP [15] that revealed the molecular organization of SAGA at high resolution. The structural data enabled us to revise the functional properties of SAGA and to compare the molecular organization with its closest cousins, the yeast [16] and Human [17] TFIID complexes.

### **General organization of SAGA**

The purification and subunit composition of the yeast SAGA complex made the connection between several transcription related functions, revealed by two independent genetic screens. The analysis of suppressors of promoter mutations caused by the insertion of yeast transposable elements identified the Spt proteins [18], and a screen for factors that interact functionally with the yeast Gcn4 or the herpes virus VP16 transcription activators revealed the Ada proteins [19]. The Gcn5 Histone Acetyl Transferase (HAT) activity, adaptor proteins involved in transcription activation and bridging factors with the general transcription factor TBP, all belonging to the Ada or Spt family, were found united in the same molecular complex [20]. The interaction network between the 19 subunits that compose the ~1.6 MDa SAGA coactivator supported a modular organization into independent structural and functional entities [21, 22]. SAGA harbors two enzymatic activities that can be detached from the core

complex without affecting the remaining subunit interactions and that can be expressed separately as active entities [23, 24] (Fig.1A). The ubiquitin protease Ubp8 assembles into a deubiquitination (DUB) module with Sus1, Sgf73 and Sgf11 to remove ubiquitin from histone H2B modified on lysine 120. The crystal structure of the DUB module composed of the first 96 residues of Sgf73 together with its three partners showed a compact intertwined organization [25, 26]. A lysine acetyl transferase activity to modify nucleosomal histone tails is carried by the Gcn5 subunit which forms a stable HAT module with Ada2, Ada3 and Sgf29. Despite intense efforts the HAT module was refractory to crystallization and the crystal structures of an Ada2-Gcn5 module suggest that it likely adopts multiple conformations [27]. The large 400 kDa Tra1 subunit, involved in activator binding, links SAGA's activities with cellular signaling pathways. It forms a module by itself [28] and a mutation in Spt20 was shown to detach Tra1 from the other SAGA subunits [29]. A dedicated module responsible for TBP interaction contains members of Spt family; Spt3 and Spt8 were described as direct TBP interactors [10, 30] while Spt7 and Spt20 were found to play major architectural roles [31]. Finally, a set of 5 TAFs are shared with TFIID and were originally attributed a role in maintaining SAGA's architecture (Fig.1A) [32]. Four TAFs (Taf6, Taf9, Taf10 and Taf12) contain a sequence motif homologous to nucleosomal core histones which promotes specific heterodimer formation [33-35]. The Taf6-Taf9 heterodimer of histone folds (HF) is found in TFIID but Taf12 and Taf10 have different partners in SAGA and pair with Ada1 and Spt7, respectively. The Spt3 subunit harbours two HFs that can form an intramolecular HF dimer (Fig.2A). The Taf5 subunit contains WD40 repeats, which adopt a beta propeller fold and interact with the Taf6-Taf9 heterodimer [36], as well as an evolutionary conserved N-terminal domain [37].

The structural description of the full SAGA complex first relied on low resolution electron microscopy models derived from negatively stained particles [38] which evidenced a two-lobed organization. Studies combining antibody labelling and analysis of tagged or mutated complexes mapped key subunits [39-41]. The first cryo-EM study taking full advantage of improvements in direct electron detectors uncovered the Tra1 subunit at 5.7 Å resolution as it fully occupies one lobe of the complex while the main lobe containing all other subunits was blurred [28]. The structure of Tra1 is virtually identical in isolation indicating that the molecular determinants of its intricate folding pattern are independent of the other SAGA subunits [42].

In our hands, purification of intact SAGA was a key factor to reach higher resolution. Not only is SAGA a scarce regulatory complex, probably 10-50 times less abundant than RNA polymerase II, but some of its subunits are prone to proteolysis. As an example, the cleavage of Spt7 by the Pep4 protease leading to the detachment of Spt8 and the generation of the SAGA like (SLIK) complex, produces structural and conformational heterogeneity [43, 44]. Zinc ions stabilize the UBP zinc finger fold present in Ubp8 which renders the DUB module sensitive

to chelating molecules. In addition, several reports mentioned the possible detachment of the DUB module from SAGA [45].

The recent cryo-EM studies showed that the main lobe, is organized around a central module which contains Spt20, Taf5 and the seven HF-containing subunits (Fig.1B). The two enzymatic DUB and HAT modules are anchored onto the central module and explore a large conformational space.

#### **Tra1, a PI3K-like kinase shared with NuA4.**

The Tra1 subunit is a major target of acidic transcription activators such as Gal4, VP16 or Gcn4 and contributes to the recruitment of SAGA to gene promoters [46-50]. Tra1 is a member of the Phosphatidylyl 3 kinase - related kinase (PIKK) family organized into three major protein domains: HEAT, FAT and Kinase (Fig.1A). The ~300 kDa N-terminal part of Tra1 is composed of HEAT repeats arranged into a solenoid that folds into an elaborate higher order architecture (Fig.1C). The N-terminal 250 residues form a protruding tail followed by a linear central solenoid overlaid by a ring-shaped solenoid that altogether adopt a cradle shape when viewed from the side. This architecture is stabilized by a network of interactions between the central and ring solenoids in the region where they cross each other. The FAT module encircles the kinase domain which is not functional in Tra1 as a result of a mutated active site [46].

Tra1 is coupled to the central module by three bridges involving Spt20, Taf12 and Spt3. The largest interface is established by the central part of Spt20 that associates with the Tra1 FAT domain through two helices. The deletion of this Spt20 Head Interacting with Tra1 (HIT) domain prevents the stable association of Tra1 to SAGA [29]. A second connection is formed by a long loop placed before the Taf12 HF domain which creates a thread at the surface of the FAT domain. A third bridge is created by the loop connecting the second and third helices of the C-terminal HF of Spt3 that inserts between two HEAT repeats of the Tra1 ring solenoid.

The complex folding of the HEAT solenoid of Tra1 and its connections with the main lobe still hold many open questions. Tra1 communicates with distal parts of SAGA as demonstrated by mutations that diminish the HAT activity or affect the stable association of the HAT module [50, 51]. These findings highlight still to be disclosed allosteric mechanisms that control communication across the whole structure. Comparing the structural organization of three PIKK proteins TOR, Tra1 and DNA-PK revealed a highly conserved sequence motif that organizes and connects different stretches of HEAT repeats. The structural similarity is particularly striking between yeast Tra1 and the human DNA-PK for which the overall architecture, the shape and size of the HEAT domain, as well as the positions of the FAT and kinase domains are almost identical. Phylogenetic analysis of PIKKs further supports the structural similarity between Tra1/TRRAP and DNA-PKcs, and confirms that they are the most closely related PIKKs [52]. Tra1 is also part of the essential NuA4 co-activator involved in

regulated gene expression but also in DNA repair. The finding that a massive structural element composed of multiple HEAT-like repeats is present in molecular complexes with critically distinct nuclear functions questions the role of these giant solenoid structures with respect to the organization of the nucleus into functional micro-domains.

### **Highly flexible enzymatic modules**

The DUB and HAT modules are firmly anchored onto the main lobe through specific attachment sites but preserve a high degree of freedom to explore a large conformational space in search of their chromatin substrates (Fig.1B). The two enzymatic modules adopt different strategies to remain bound to SAGA while maintaining their freedom of movement. The independent and functional DUB module is a stably folded structure [25, 26] that gains independence through an unstructured 164 residue-long linker which forms little contact with SAGA surface. The DUB module behaves like an independent entity and its structure is not modified upon binding to the modified nucleosome ligand [14, 53]. Nucleosome binding may result in even more flexible enzymatic modules [14]. The central part of Sgf73 forms the anchoring domain which docks into an interface composed exclusively of SAGA-specific elements, with no contribution from subunits shared with TFIID. The action of the proteasome regulatory particle and a specific interaction between the proteasomal ATPase, Rpt2 and Sgf73 is required to detach the DUB module from the SAGA complex [45]. The structures show that the anchoring domain is surrounded by Ada1 loops and by a Spt20 domain making it very difficult to detach the DUB module without major structural rearrangements.

In contrast to the DUB, the HAT module appears flexible by itself [27], with several sub-modules as well as a large proportion of intrinsically unfolded sequences. The HAT module is not deeply embedded within SAGA core, but is inserted through two helical domains, at the surface of the Taf6 HEAT repeats. This interface is used for homo-dimerization in TFIID and only the unique organization of Taf6 HEAT and Taf5 N-Terminal Domain (NTD) found in SAGA creates the specific binding site for Ada3 (see below).

### **Histone fold assemblies in SAGA and TFIID**

A sequence motif homologous to the histone fold is present in several subunits of the TFIID and SAGA complexes [34]. Five heterodimers can be formed *in vitro* from HF domains present in nine TFIID subunits (Taf4-12, Taf6-9, Taf11-13 and Taf10 having two partners Taf8 and Taf3) [34]. In SAGA, seven subunits assemble into four HF heterodimers namely Taf6-9, Taf10-Spt7, Taf12-Ada1 and Spt3 which contains two HF domains forming an intramolecular dimer (Fig.2A). The assembly of higher ordered structures from several HF heterodimers was suggested at their discovery [33], but their exact composition remained debated. An octamer containing two copies of the Taf6-9 and Taf4-12 heterodimers was reported [54]. This

hypothesis was attractive since these subunits were found to be present as two copies in TFIID [8] and also because the two-fold symmetry of the histone octamer would have been preserved. The cryo-EM structures of SAGA and TFIID confirmed the presence of higher ordered HF structures held together by four-helix bundle interactions between adjacent HF pairs as they are employed by nucleosomal histones. SAGA accommodates a non-symmetric octamer including the seven HF proteins (Fig.2A). The three-lobed human [17] and yeast [16] TFIID structures contain both an octamer and a hexamer of HF proteins located in two distinct lobes. The flexibility of TFIID limits the currently attainable resolution and the identification of the proteins forming each HF assembly relies on an integrative approach including cross-linking experiments. The octamer present in the yeast Taf1-lobe, closely related to the human lobe A, is most likely composed of the Taf4-12, Taf6-9, Taf10-3 and Taf11-13 pairs. The hexameric structure, located in the yeast Twin lobe, also named lobe B in humans, is formed by the Taf4-12, Taf6-9 and Taf10-8 HF pairs (Fig.2A). A similar hexameric HF organization has been proposed in archaea where the small histone like HMfB proteins interact with DNA and form a continuous superhelical arrangement of (HMfB)<sub>2</sub> homodimers [55]. The presence of a hexamer in the Twin lobe questions the mechanism that prevents the accretion of the last heterodimer. Taf8 could play an essential role in this process during complex assembly as revealed by sub-complex production [56, 57]. We proposed that the addition of the Taf8-10 heterodimer introduces asymmetry into the two-fold symmetric 5 TAF subcomplex containing Taf5, Taf6-9 and Taf4-12, and controls the differentiation of the Twin- and Taf1-lobes within TFIID.

The HF-proteins found in SAGA and TFIID do not interact with DNA since the amino acid side chains that mediate contacts between nucleosomal histones and DNA are not conserved [58]. The N-terminal part of Taf4 is an exception to this rule and harbours non-sequence specific DNA binding activities [59]. The Taf4-12 pair accommodated in the Twin lobe, in addition to Taf2 located in the Taf2 lobe, provide TFIID with unique promoter binding properties that are absent in SAGA. The structure of promoter-bound human TFIID revealed in addition a strong TAF1 binding interface with the downstream core promoter DNA, next to the TAF2 interaction site [60].

### **TAF5 controls the higher order assembly of histone fold domains**

In all three instances of HF structures found in TFIID and SAGA, a common foundation organizes their higher ordered assembly. The Taf9 subunit contacts the side of the seven blade WD40 repeat of Taf5 by a helix adjacent to the HF and Taf6 associates with the centre of the  $\beta$ -propeller as was revealed by a crystal structure of the TAF5-TAF6-TAF9 subcomplex [36]. Successive HF pairs in the helical assembly that leads to hexamers or tetramers are held together by four-helix bundles composed of helices  $\alpha$ 2 and  $\alpha$ 3 of one histone fold from each



pair. The second pair always contains Taf12, which forms a four-helix bundle with Taf9, and its Taf4 partner is preserved in both the hexameric and octameric arrangements of TFIID. Interestingly, in all three HF assemblies Taf10 is always placed in the third HF pair, and was shown to contact Taf9 outside of a four-helix bundle in the high resolution SAGA structures [14, 15]. The Taf10 position, and its specific dimerization partner, appear strategic to control the incorporation of the last HF pair dedicated to TBP binding. In SAGA, Spt3 is incorporated when Taf10 pairs with Spt7, in TFIID, Taf11-13 is recruited when Taf10 dimerises with Taf3, and no further HF pair is added when Taf10 interacts with Taf8.

The nucleosomal histone octamer assembly proceeds from the internal H3-H4 tetramer and incorporates two outer H2A-H2B dimers. In SAGA and TFIID, the interaction of the Taf5 WD40 domain with the Taf6-9 heterodimer largely controls the assembly of the spiral of HF heterodimers. In the yeast complexes, the Taf6 HEAT repeats and its relative position to the Taf5-NTD correlates with the incorporation of one (SAGA) or two (TFIID) HF assemblies (Fig.2B, C). Yeast TFIID contains two Taf5 and Taf6 copies, held together by homo-dimerization of Taf5-NTDs and Taf6 HEAT repeats, thus exposing two Taf5 WD40 propellers for HF assembly [16]. TAF5-NTD dimerization was not observed in human TFIID and may reflect a weaker interaction interface in metazoan than in yeast cells [17]. Compared to TFIID, the Taf5-NTD in SAGA makes almost a 180 degrees turn around the WD40 domain and is placed next to the Taf6 HEAT repeats, thus preventing Taf5-Taf5 and Taf6-Taf6 homodimerization. As a consequence only one Taf5 WD40 domain is exposed resulting in the incorporation of a single octamer into SAGA. It is currently unclear whether the relative orientation of these domains, and therefore the assembly of SAGA versus TFIID is regulated, or whether they result from stochastic events during the assembly process. Interestingly the comparative analysis of post-translational modifications between TFIID and SAGA subunits revealed a SAGA-specific phosphorylation in a Taf5 loop next to the WD40 repeats which may play a role in orienting the Taf5 and Taf6 domains or preventing their homodimerization [61]. Alternatively, the binding interface of Ada3 with the Taf6 HEAT domain overlaps with the homodimerization interface, and the incorporation of the HAT module may drive the assembly of SAGA versus TFIID.

### **The TAF5 connection hub**

The Taf5 WD40 domain is a centrally located master hub that orchestrates the architecture of the entire main lobe. It not only controls the assembly of the octameric HF assembly through contacts with the Taf6-9 heterodimer, but further mediates multiple protein-protein interactions by contacting at least 11 different protein domains. As a striking example, Taf6 residues add a beta-strand to the last blade of the Taf5 WD40 propeller forming an inter-protein beta sheet. Histone fold-containing subunits located at a distance from Taf5, such as Spt7, Taf10 or Ada1, make specific contacts with the Taf5 WD40 domain through protein sequences predicted as

unstructured, which become exquisitely ordered and well resolved structural elements upon interaction. This dense network of interactions places Taf5 in a strategic position to influence the fate of TFIID or SAGA assembly, to bring together subunits required for SAGA assembly, but also to orchestrate informational flow between the multiple SAGA functions.

### **The TBP interaction module and the mechanism that prevents TBP from binding to DNA.**

The histone-fold octamer is not a mere protein-protein interaction interface designed for architectural purposes as originally thought, it is the core of the molecular machine that interacts with, and loads TBP onto gene promoters in a highly regulated manner. Purified SAGA is generally not associated with TBP, but a stable complex can be reconstituted from purified SAGA and TBP [15]. The cryo-EM structure of this complex disclosed the molecular details of the interaction of TBP with Spt3, which is the main component that drives the interaction and orients the DNA binding groove of TBP (Fig.3A). Spt3 interacts primarily with the C-terminal stirrup of TBP, which is almost completely buried, and contacts a large pocket formed by helices  $\alpha$ N and  $\alpha$ 2 from the C-terminal Spt3 HF and by helix  $\alpha$ C that is joined to the N-terminal Spt3 HF. The interaction of the C-terminal half of TBP with Spt3 is consistent with previous biochemical and genetic data [30, 62].

In contrast to Spt3, which is incorporated into the stable histone fold octamer, Spt8 is flexibly tethered to the tips of two long Spt7 helices extending at each end of the histone fold. The connection with Spt8 appears as a fuzzy density corresponding to a 30 residue long unstructured region in the C-terminus of Spt7 whose deletion results in loss of Spt8 [43, 44]. The flexible connection of Spt8 to the N-terminal half of TBP is likely to play a secondary role in positioning TBP or in competing with other TBP-associated proteins consistent with the observation that Spt3 deletion mutants showed a sharp drop in transcription while impairment of Spt8 produced little effect [12]. In the absence of TBP, Spt8 is flexible and was detected in a slightly different position suggesting a reorganization of Spt8 upon TBP binding to Spt3 [14]. TBP is held in a configuration that impairs binding to promoter DNA. Spt3 prevents the C-terminal stirrup of TBP from interacting with TFIIB while Spt8 interferes with TFIIA binding to the N-terminal stirrup. TBP is ready to bind DNA since the space between its DNA binding cleft and SAGA's main lobe is wide enough to accommodate a double-stranded DNA and since the cleft is not occupied by any well-folded protein density. Gel-shift and pull-down assays show that in these conditions TBP cannot bind to DNA, even containing a consensus TATA box. The structure shows that the path of DNA clashes with several Tra1 HEAT repeats and with parts of Spt3, in particular helix  $\alpha$ C rigidly bound to the N-terminal HF and protruding to the DNA-binding cavity of TBP, all located at a distance from the TBP DNA binding cleft. Steric hindrance provoked by distal structural elements is therefore the major mechanism that

prevents TBP from binding to DNA within the SAGA complex.

### **Mode of TBP release and loading onto promoters**

The general transcription factor TFIIA plays a crucial role in unlocking TBP and releasing DNA binding inhibition. When the SAGA-TBP complex is incubated with a TATA-box containing DNA together with TFIIA, TBP is able to access DNA and is detached from SAGA as a ternary DNA-TBP-TFIIA complex [15]. The release of TBP from SAGA enables subsequent PIC assembly by liberating the TBP C-terminal stirrup for TFIIB interaction. Interestingly, TBP release from SAGA correlates with the affinity of TBP for its DNA substrate. Indeed, TBP binds to DNA less efficiently, but still detaches from SAGA, when the TATA element is degraded by two mismatches, whereas no TBP release is observed when the DNA does not contain any TATA or TATA-like elements.

The exact mechanism for the TFIIA-stimulated concerted DNA binding and TBP release from SAGA is not known but must include conformational changes that are built-in the octameric HF structure. TFIIA is likely to first displace Spt8 from TBP, but then needs to modify the orientation of TBP to place its DNA binding cleft in a favourable position to promote initial binding of DNA. Most importantly, shifting of the Spt3 helix  $\alpha$ C that obstructs the path of DNA is critically required. The extremity of helix  $\alpha$ C, close to TBP's DNA binding cleft, and its preceding helix  $\alpha$ 3 are poorly resolved in the cryo-EM map suggesting that they are flexible to facilitate the infiltration of DNA. Interestingly, Spt3 in isolation does not interact with TBP in solution and only the conformation adopted within SAGA allows TBP binding [11]. It is therefore tempting to speculate that TFIIA binding or DNA infiltration switches the conformation of Spt3 towards the one found in isolation and would therefore favour the full release of TBP from SAGA. Since helix  $\alpha$ C is firmly joined to helix  $\alpha$ 3 in the N-terminal Spt3 HF, this crucial step requires a movement of the whole Spt3 intramolecular HF dimer which is feasible only because the octamer is not as perfectly symmetric as in nucleosomes, but is being deformed as the HF spiral progresses towards its Spt3 end. Close inspection of the octamer shows that while all other HF pairs are oriented as in the nucleosome, the Spt3 pair is tilted by 20 degrees and is not immobilized by a tight four-helix bundle. This 20° tilt, shears the four-helix bundle and nearly frees Spt3 from its association with the HF octamer thus providing sufficient degrees of freedom to reorient TBP. It is noteworthy that, probably to secure incorporation of the weakened Spt3 into SAGA, the C-terminal tail of Spt3 inserts into the octamer center, a void volume in the nucleosome, and bonds with all other histone folds in SAGA. The deformed SAGA octamer appears to be geared specifically towards a delicate balance between a rigid Spt3 firmly embedded into the octamer and an excessively flexible Spt3 unable to position TBP.

### **A shared TBP interaction and release mechanism between SAGA and TFIID.**

SAGA and TFIID contain a conserved TBP-binding machine constituted by a HF octamer, TBP interaction partners with high sequence similarity [35], and a DNA-binding autoinhibition mechanism that is released by TFIIA. Altogether these properties suggest that the mechanism of TBP binding and loading onto the promoter are comparable in both complexes. Currently, cryo-EM resolution in the yeast Taf1-lobe and the human lobe A is not sufficient to determine to what extent the TBP interaction mode is similar in promoter unbound TFIID and in SAGA (Fig.3A and B). A model derived from the analysis of a ternary human TAF11-TAF13-TBP complex combining TAF11-13 crystal coordinates, SAXS data, biochemical analyses and cross-linking mass-spectrometry data, identified a conserved C-terminal TBP-interaction domain in TAF13 [63], corresponding to the  $\alpha$ C helix of Spt3. This study also suggested that TAF11-13 interacts with the DNA binding surface of TBP arguing for a significantly different TBP binding mode in TFIID as compared to SAGA. Since the whole HF octamer is required for productive interaction of TBP with Spt3, it is possible that TBP binds differently to isolated TAF11-13 than to full TFIID, and the crystal structure may have trapped a TBP release intermediate rather than the TFIID-bound state.

The low TBP occupancy in purified SAGA preparations points to further differences in TBP binding to TFIID vs SAGA, and argues for a weaker interaction of TBP with Spt3 than with Taf11-13, which may provoke the dissociation of TBP from SAGA during purification. This hypothesis is supported by a yeast strain expressing a mutant Spt3 that binds stronger to TBP [10], since SAGA purified from this strain contains a higher TBP stoichiometry. However, a reconstituted SAGA-TBP complex is very stable and withstands chromatographic purification or pull-down assays [15], suggesting that the low TBP content in purified SAGA may arise from alternative mechanisms. The dynamics of TBP deposition and TBP turnover on specific classes of Pol II promoters affects the fraction of time TBP interacts with either TFIID or SAGA. In this respect, the highest TBP turnover, as measured by differential chromatin immunopurification experiments, was found on TATA-containing promoters bound by SAGA [64]. High TBP turnover also correlates with Mot1-dependent promoters which opens the possibility that additional factors modulate TBP occupancy on SAGA. Finally, the distance between the coactivator binding site and the TSS may influence the amount of TBP bound to TFIID or SAGA. In yeast, SAGA is mainly found on Upstream Activating Sequences (UAS) placed 200 to 300 bp upstream of the TSS. Therefore SAGA is likely to move away once TBP is loaded on the promoter (Fig.3C). In contrast TFIID interacts upstream and downstream of the TSS and is likely to remain close to the promoter bound TBP (Fig.3D).

SAGA and TFIID employ similar mechanisms, with minor adjustments, to prevent TBP from binding to spurious DNA. Steric hindrance, brought about by the position and orientation of TBP within the molecular complex, is the mechanism prevailing in SAGA. The compact

conformation of yeast TFIID, achieved by Taf5-NTD dimerization, is likely to play a similar role in hiding TBP's DNA binding cleft. With its more open conformation this effect is reduced in human TFIID. Alternatively, the TAF1 TAND domain [65] and the TAF11-13 HF pair [63] have been shown to inhibit TBP from binding to DNA and may play a role in TATA-box mimicry and/or in DNA binding cleft obstruction. High resolution structures of yeast Taf1-lobe or human Lobe A need to confirm that these effects, demonstrated on partial complexes, also operate in the context of the full co-activators. In SAGA, the DNA binding cleft of TBP is not occupied by a folded protein domain mimicking the structure of DNA, but it cannot be excluded that the unfolded helix  $\alpha$ C end of the N-terminal Spt3 HF, which protrudes towards the DNA-binding cavity of TBP, can interfere with TBP binding to DNA.

Finally, both SAGA-TBP and TFIID require TFIIA to unlock TBP and load it onto the promoter, thus indicating that a similar mechanism is used. The combined action of TFIIA and promoter sequences results in the dissociation of a TBP-TFIIA-DNA complex from SAGA. Human TFIID was shown to undergo a series of rearrangements resulting in a fully engaged state in which a TFIIA-TBP-DNA structure is tethered to TFIID by interactions of TFIIA with TAF4 and TAF12, while TBP does not contact any TFIID subunit. Such massive rearrangements were not observed upon binding of yeast TFIID to promoter DNA, and the TFIIA-TBP-DNA complex was not observed. The most plausible interpretation is that in the absence of identified yeast core promoter elements, TBP is indeed loaded onto the DNA with TFIIA and the whole TFIIA-TBP-DNA complex can slide along the core TFIID.

TFIID and SAGA thus share the capacity to interact with and to load TBP onto gene promoters, but their roles in gene expression are different. Pioneering genome wide ChIP experiments established the concept of housekeeping TFIID- and stress-induced SAGA- dependent genes [13], and this paradigm has been revised significantly by recent run on transcription, subunit deletion or anchor away experiments [12, 66, 67]. In *S. cerevisiae*, SAGA has more widespread roles in RNA polymerase II transcription than initially suggested, but has specific additional effects on highly regulated, and possibly highly expressed genes, with no clear cut-off between the two classes. With reduced intrinsic DNA binding properties, SAGA is likely to be recruited to UAS regions mainly by transcriptional activators induced by stress specific signalling pathways (Fig.3C). In contrast, TFIID binds strongly to DNA and the cryo-EM structure of yeast promoter-bound TFIID shows that a 35-40 bp long stretch of promoter DNA is held between the Twin- and Taf2-lobes by Taf4 and Taf2 respectively [16]. Human TFIID shows additional interactions of TAF1 with downstream promoter elements and of TBP with the upstream DNA [60]. TFIID's unique DNA binding properties, requiring 120-150 Å of free DNA for optimal interaction, may reduce activator dependency (Fig.3D). As this distance is larger than the internucleosomal linkers, whose length varies between 15-20 bp and corresponding to 50-70 Å, TFIID may search for nucleosome depleted regions found around

the promoters of most genes, including weakly expressed house-keeping genes, in order to interact tightly and find a suitable TBP deposition site. The wide range of SAGA, versus TFIID, gene dependencies could thus be explained by considering that each promoter has two components for TBP loading: a component proximal to the transcription start site which is driven by TFIID and a distal UAS-based component that is controlled by SAGA.

### **The versatile SAGA complex.**

SAGA assembles a unique blend of nuclear functions into a unique complex: chromatin modifications performed by two different enzymatic modules, readers of multiple histone modifications, activator binding, and TBP interaction and loading onto promoters. This exquisite assembly of functional modules serves multiple purposes in addition to transcription initiation and participates in the coordination of major nuclear activities. SAGA plays a major role in double strand (ds) DNA repair mechanisms. H2B deubiquitination impacts the early stages of various dsDNA repair pathways since the failure of SAGA to remove ubiquitin from H2BK120ub inhibits ATM- and DNAPK-induced  $\gamma$ H2AX formation, a hallmark of several dsDNA repair mechanisms [68], and a ds break-induced monoubiquitination-to-acetylation switch on histone H2BK120 was described to be mediated by the SAGA complex [69]. The H2B monoubiquitination balance maintained by the SAGA DUB module is also crucial for correct double strand break repair during homologous recombination [70].

SAGA has been involved in the positioning of active genes and in mRNA export through a functional cross-talk between the mRNA export machinery TREX-2 and the DUB module [71]. SAGA couples splicing and transcription as highlighted in a recent report showing that splicing proofreading is modulated by Spt8 and Spt3, thus revealing interactions between the TBP-binding module of SAGA and the spliceosomal ATPase Prp5p [72]. SAGA subunits further contribute to the asymmetric segregation of non-chromosomal DNA circles by mediating their interaction with nuclear pore complexes [73]. SAGA causes accumulation of circles and NPCs in the mother cell, and thereby promotes ageing.

The plethora of functions conveyed by SAGA, some of which were disclosed only recently, prefigures that additional regulatory functions involving fascinating molecular mechanisms are yet to be discovered. All these functions need to be envisaged and explored to explain the deleterious effects of SAGA mutations in cancer or in neurodegenerative diseases.

### **AUTHOR CONTRIBUTIONS**

PS, ABS and GP discussed the scope of the review, PS wrote the manuscript, ABS and GP revised the manuscript, and GP contributed to the artwork.

### **ACKNOWLEDGEMENTS**

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Centre National pour la Recherche Scientifique (CNRS), the Ligue contre le Cancer, the ANR-15-CE11-0022-01 and the grant ANR-10-LABX-0030-INRT under the frame program Investissements d'Avenir ANR-10- IDEX-0002-02.

## REFERENCES

1. Hahn, S. & Young, E. T. (2011) Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators, *Genetics*. **189**, 705-36.
2. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II, *Cell*. **56**, 549-561.
3. Vannini, A. & Cramer, P. (2012) Conservation between the RNA polymerase I, II, and III transcription initiation machineries, *Mol Cell*. **45**, 439-46.
4. Meisterernst, M. & Roeder, R. G. (1991) Family of proteins that interact with TFIID and regulate promoter activity, *Cell*. **67**, 557-567.
5. Poon, D., Campbell, A. M., Bai, Y. & Weil, P. A. (1994) Yeast Taf170 is encoded by MOT1 and exists in a TATA box-binding protein (TBP)-TBP-associated factor complex distinct from transcription factor IID, *JBiol Chem*. **269**, 23135-23140.
6. Dudley, A. M., Rougeulle, C. & Winston, F. (1999) The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step in vivo, *Genes Dev*. **13**, 2940-5.
7. Reinberg, D., Horikoshi, M. & Roeder, R. G. (1987) Factors involved in specific transcription in mammalian RNA polymerase II. Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site, *J Biol Chem*. **262**, 3322-30.
8. Sanders, S. L., Garbett, K. A. & Weil, P. A. (2002) Molecular characterization of *Saccharomyces cerevisiae* TFIID, *Mol Cell Biol*. **22**, 6000-13.
9. Bai, Y., Perez, G. M., Beechem, J. M. & Weil, P. A. (1997) Structure-function analysis of TAF130: identification and characterization of a high-affinity TATA-binding protein interaction domain in the N terminus of yeast TAF(II)130, *Mol Cell Biol*. **17**, 3081-3093.
10. Eisenmann, D. M., Arndt, K. M., Ricupero, S. L., Rooney, J. W. & Winston, F. (1992) SPT3 interacts with TFIID to allow normal transcription in *Saccharomyces cerevisiae*, *Genes Dev*. **6**, 1319-1331.
11. Sermwittayawong, D. & Tan, S. (2006) SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment, *Embo J*. **25**, 3791-800.
12. Baptista, T., Grunberg, S., Minoungou, N., Koster, M. J. E., Timmers, H. T. M., Hahn, S., Devys, D. & Tora, L. (2017) SAGA Is a General Cofactor for RNA Polymerase II Transcription, *Mol Cell*. **68**, 130-143 e5.
13. Huisinga, K. L. & Pugh, B. F. (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*, *Mol Cell*. **13**, 573-85.
14. Wang, H., Dienemann, C., Stutzer, A., Urlaub, H., Cheung, A. C. M. & Cramer, P. (2020) Structure of the transcription coactivator SAGA, *Nature*. **577**, 717-720.
15. Papai, G., Frechard, A., Kolesnikova, O., Crucifix, C., Schultz, P. & Ben-Shem, A. (2020) Structure of SAGA and mechanism of TBP deposition on gene promoters, *Nature*. **577**, 711-716.
16. Kolesnikova, O., Ben-Shem, A., Luo, J., Ranish, J., Schultz, P. & Papai, G. (2018) Molecular structure of promoter-bound yeast TFIID, *Nature communications*. **9**, 4666.
17. Patel, A. B., Louder, R. K., Greber, B. J., Grunberg, S., Luo, J., Fang, J., Liu, Y., Ranish, J., Hahn, S. & Nogales, E. (2018) Structure of human TFIID and mechanism of TBP loading onto promoter DNA, *Science*. **362**.
18. Winston, F., Chaleff, D. T., Valent, B. & Fink, G. R. (1984) Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*, *Genetics*. **107**, 179-97.

19. Berger, S. L., Pina, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L., Triezenberg, S. J. & Guarente, L. (1992) Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains, *Cell*. **70**, 251-265.
20. Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., Berger, S. L. & Workman, J. L. (1997) 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**, 1640-1650.
21. Han, Y., Luo, J., Ranish, J. & Hahn, S. (2014) Architecture of the *Saccharomyces cerevisiae* SAGA transcription coactivator complex, *EMBO J.* **33**, 2534-46.
22. Lee, K. K., Sardi, M. E., Swanson, S. K., Gilmore, J. M., Torok, M., Grant, P. A., Florens, L., Workman, J. L. & Washburn, M. P. (2011) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes, *Mol Syst Biol.* **7**, 503.
23. Balasubramanian, R., Pray-Grant, M. G., Selleck, W., Grant, P. A. & Tan, S. (2002) Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation, *J Biol Chem.* **277**, 7989-95.
24. Kohler, A., Schneider, M., Cabal, G. G., Nehrbass, U. & Hurt, E. (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export, *Nat Cell Biol.* **10**, 707-15.
25. Kohler, A., Zimmerman, E., Schneider, M., Hurt, E. & Zheng, N. (2010) Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module, *Cell*. **141**, 606-17.
26. Samara, N. L., Datta, A. B., Berndsen, C. E., Zhang, X., Yao, T., Cohen, R. E. & Wolberger, C. (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module, *Science*. **328**, 1025-9.
27. Sun, J., Paduch, M., Kim, S. A., Kramer, R. M., Barrios, A. F., Lu, V., Luke, J., Usatyuk, S., Kossiakov, A. A. & Tan, S. (2018) Structural basis for activation of SAGA histone acetyltransferase Gcn5 by partner subunit Ada2, *Proc Natl Acad Sci U S A.* **115**, 10010-10015.
28. Sharov, G., Voltz, K., Durand, A., Kolesnikova, O., Papai, G., Myasnikov, A. G., Dejaegere, A., Ben Shem, A. & Schultz, P. (2017) Structure of the transcription activator target Tra1 within the chromatin modifying complex SAGA, *Nat Commun.* **8**, 1556.
29. Elias-Villalobos, A., Toullec, D., Faux, C., Seveno, M. & Helmlinger, D. (2019) Chaperone-mediated ordered assembly of the SAGA and NuA4 transcription co-activator complexes in yeast, *Nat Commun.* **10**, 5237.
30. Eisenmann, D. M., Chapon, C., Roberts, S. M., Dollard, C. & Winston, F. (1994) The *Saccharomyces cerevisiae* SPT8 gene encodes a very acidic protein that is functionally related to SPT3 and TATA-binding protein, *Genetics*. **137**, 647-657.
31. Wu, P. Y. & Winston, F. (2002) Analysis of Spt7 function in the *Saccharomyces cerevisiae* SAGA coactivator complex, *Mol Cell Biol.* **22**, 5367-79.
32. Hoffmann, A., Oelgeschlager, T. & Roeder, R. G. (1997) Considerations of transcriptional control mechanisms: do TFIID-core promoter complexes recapitulate nucleosome-like functions?, *Proc Natl Acad Sci USA.* **94**, 8928-8935.
33. Hoffmann, A., Chiang, C. M., Oelgeschlager, T., Xie, X., Burley, S. K., Nakatani, Y. & Roeder, R. G. (1996) A histone octamer-like structure within TFIID, *Nature*. **380**, 356-9.
34. Gangloff, Y., Romier, C., Thuault, S., Werten, S. & Davidson, I. (2001) The histone fold is a key structural motif of transcription factor TFIID, *Trends Biochem Sci.* **26**, 250-7.
35. Birck, C., Poch, O., Romier, C., Ruff, M., Mengus, G., Lavigne, A. C., Davidson, I. & Moras, D. (1998) Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family, *Cell*. **94**, 239-49.
36. Antonova, S. V., Haffke, M., Corradini, E., Mikucianas, M., Low, T. Y., Signor, L., van Es, R. M., Gupta, K., Scheer, E., Vos, H. R., Tora, L., Heck, A. J. R., Timmers, H. T. M. & Berger, I. (2018) Chaperonin CCT checkpoint function in basal transcription factor TFIID assembly, *Nat Struct Mol Biol.* **25**, 1119-1127.

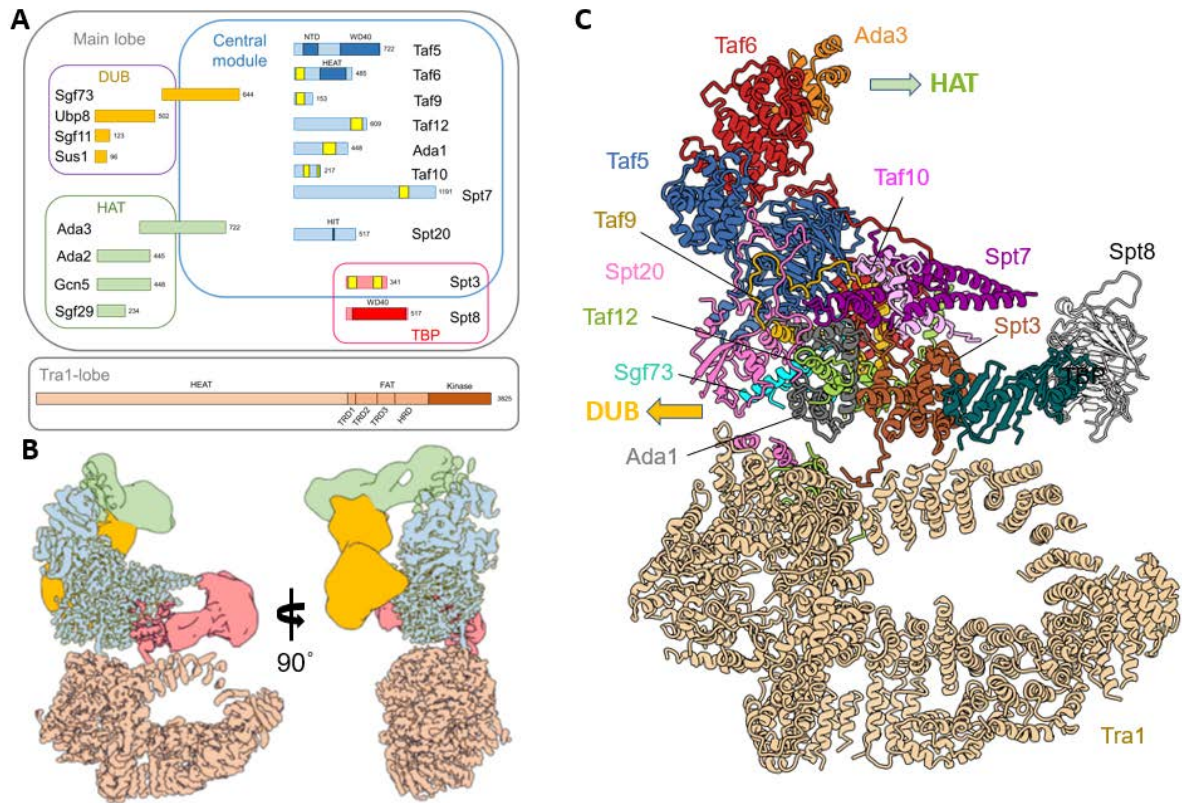


37. Romier, C., James, N., Birck, C., Cavarelli, J., Vivares, C., Collart, M. A. & Moras, D. (2007) Crystal structure, biochemical and genetic characterization of yeast and *E. cuniculi* TAF(II)5 N-terminal domain: implications for TFIID assembly, *J Mol Biol.* **368**, 1292-306.
38. Brand, M., Leurent, C., Mallouh, V., Tora, L. & Schultz, P. (1999) Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC, *Science.* **286**, 2151-3.
39. Wu, P. Y., Ruhlmann, C., Winston, F. & Schultz, P. (2004) Molecular architecture of the *S. cerevisiae* SAGA complex, *Mol Cell.* **15**, 199-208.
40. Setiাপutra, D., Ross, J. D., Lu, S., Cheng, D. T., Dong, M. Q. & Yip, C. K. (2015) Conformational flexibility and subunit arrangement of the modular yeast spt-ada-gcn5 acetyltransferase complex, *J Biol Chem.* **290**, 10057-70.
41. Durand, A., Bonnet, J., Fournier, M., Chavant, V. & Schultz, P. (2014) Mapping the deubiquitination module within the SAGA complex, *Structure.* **22**, 1553-9.
42. Diaz-Santin, L. M., Lukoyanova, N., Aciyan, E. & Cheung, A. C. (2017) Cryo-EM structure of the SAGA and NuA4 coactivator subunit Tra1 at 3.7 angstrom resolution, *eLife.* **6**.
43. Pray-Grant, M. G., Schieltz, D., McMahon, S. J., Wood, J. M., Kennedy, E. L., Cook, R. G., Workman, J. L., Yates, J. R., 3rd & Grant, P. A. (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway, *Mol Cell Biol.* **22**, 8774-86.
44. Spedale, G., Mischerikow, N., Heck, A. J., Timmers, H. T. & Pijnappel, W. W. (2010) Identification of Pep4p as the protease responsible for formation of the SAGA-related SLIK protein complex, *J Biol Chem.* **285**, 22793-9.
45. Lim, S., Kwak, J., Kim, M. & Lee, D. (2013) Separation of a functional deubiquitylating module from the SAGA complex by the proteasome regulatory particle, *Nat Commun.* **4**, 2641.
46. McMahon, S. B., Van Buskirk, H. A., Dugan, K. A., Copeland, T. D. & Cole, M. D. (1998) The novel ATM-related protein TRRAP is an essential cofactor for the c- Myc and E2F oncoproteins, *Cell.* **94**, 363-374.
47. Saleh, A., Schieltz, D., Ting, N., McMahon, S. B., Litchfield, D. W., Yates, J. R., Lees-Miller, S. P., Cole, M. D. & Brandl, C. J. (1998) Tra1p is a component of the yeast Ada.Spt transcriptional regulatory complexes, *JBiol Chem.* **273**, 26559-26565.
48. Brown, C. E., Howe, L., Sousa, K., Alley, S. C., Carrozza, M. J., Tan, S. & Workman, J. L. (2001) Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit, *Science.* **292**, 2333-7.
49. Bhaumik, S. R. & Green, M. R. (2001) SAGA is an essential in vivo target of the yeast acidic activator Gal4p, *Genes Dev.* **15**, 1935-45.
50. Knutson, B. A. & Hahn, S. (2011) Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes, *Mol Cell Biol.* **31**, 818-31.
51. Mutiu, A. I., Hoke, S. M., Genereaux, J., Hannam, C., MacKenzie, K., Jobin-Robitaille, O., Guzzo, J., Cote, J., Andrews, B., Haniford, D. B. & Brandl, C. J. (2007) Structure/function analysis of the phosphatidylinositol-3-kinase domain of yeast tra1, *Genetics.* **177**, 151-66.
52. Elias-Villalobos, A., Fort, P. & Helmlinger, D. (2019) New insights into the evolutionary conservation of the sole PIKK pseudokinase Tra1/TRRAP, *Biochem Soc Trans.* **47**, 1597-1608.
53. Morgan, M. T., Haj-Yahya, M., Ringel, A. E., Bandi, P., Brik, A. & Wolberger, C. (2016) Structural basis for histone H2B deubiquitination by the SAGA DUB module, *Science.* **351**, 725-8.
54. Selleck, W., Howley, R., Fang, Q., Podolny, V., Fried, M. G., Buratowski, S. & Tan, S. (2001) A histone fold TAF octamer within the yeast TFIID transcriptional coactivator, *Nat Struct Biol.* **8**, 695-700.
55. Mattioli, F., Bhattacharyya, S., Dyer, P. N., White, A. E., Sandman, K., Burkhart, B. W., Byrne, K. R., Lee, T., Ahn, N. G., Santangelo, T. J., Reeve, J. N. & Luger, K. (2017) Structure of histone-based chromatin in Archaea, *Science.* **357**, 609-612.
56. Trowitzsch, S., Viola, C., Scheer, E., Conic, S., Chavant, V., Fournier, M., Papai, G., Ebong, I. O., Schaffitzel, C., Zou, J., Haffke, M., Rappsilber, J., Robinson, C. V., Schultz, P., Tora, L. & Berger, I. (2015) Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules, *Nat Commun.* **6**, 6011.

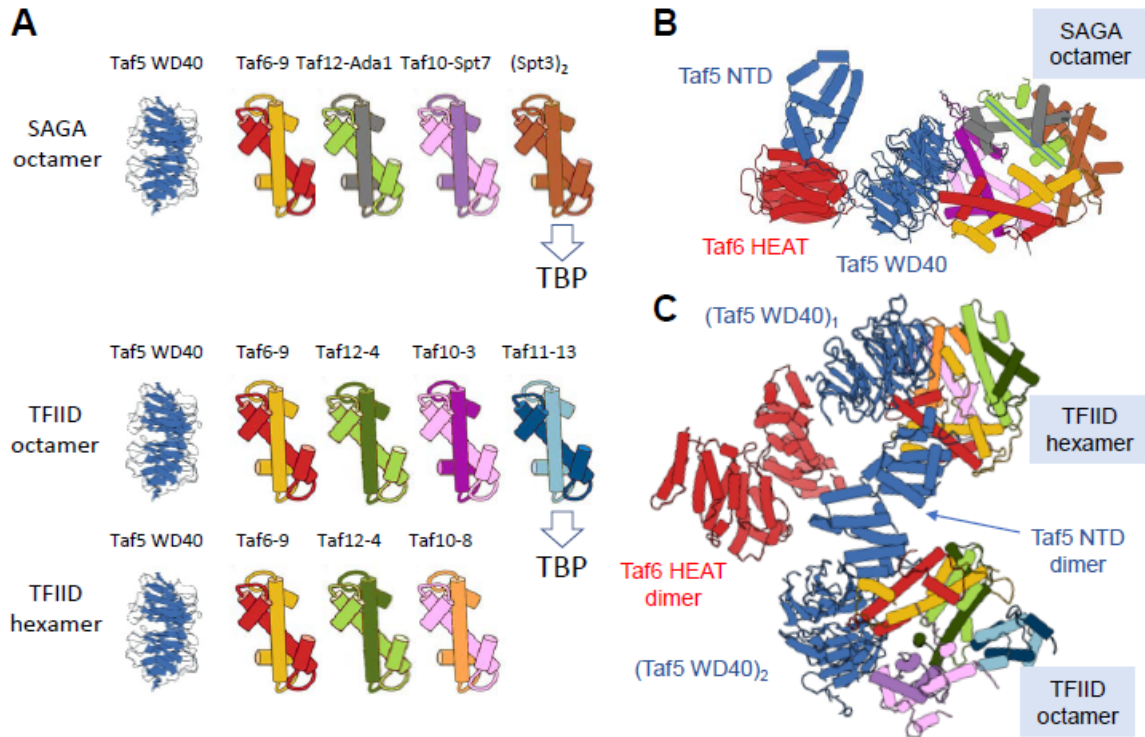
57. Bieniossek, C., Papai, G., Schaffitzel, C., Garzoni, F., Chaillet, M., Scheer, E., Papadopoulos, P., Tora, L., Schultz, P. & Berger, I. (2013) The architecture of human general transcription factor TFIID core complex, *Nature*. **493**, 699-702.
58. Werten, S., Mitschler, A., Romier, C., Gangloff, Y. G., Thuault, S., Davidson, I. & Moras, D. (2002) Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20), *J Biol Chem*. **277**, 45502-9.
59. Gazit, K., Moshonov, S., Elfakess, R., Sharon, M., Mengus, G., Davidson, I. & Dikstein, R. (2009) TAF4/4b x TAF12 displays a unique mode of DNA binding and is required for core promoter function of a subset of genes, *J Biol Chem*. **284**, 26286-96.
60. Louder, R. K., He, Y., Lopez-Blanco, J. R., Fang, J., Chacon, P. & Nogales, E. (2016) Structure of promoter-bound TFIID and model of human pre-initiation complex assembly, *Nature*. **531**, 604-9.
61. Mischerikow, N., Spedale, G., Altelaar, A. F., Timmers, H. T., Pijnappel, W. W. & Heck, A. J. (2009) In-depth profiling of post-translational modifications on the related transcription factor complexes TFIID and SAGA, *J Proteome Res*. **8**, 5020-30.
62. 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 Dev*. **22**, 2994-3006.
63. Gupta, K., Watson, A. A., Baptista, T., Scheer, E., Chambers, A. L., Koehler, C., Zou, J., Obong-Ebong, I., Kandiah, E., Temblador, A., Round, A., Forest, E., Man, P., Bieniossek, C., Laue, E. D., Lemke, E. A., Rappsilber, J., Robinson, C. V., Devys, D., Tora, L. & Berger, I. (2017) Architecture of TAF11/TAF13/TBP complex suggests novel regulation properties of general transcription factor TFIID, *eLife*. **6**.
64. van Werven, F. J., van Teeffelen, H. A., Holstege, F. C. & Timmers, H. T. (2009) Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome, *Nat Struct Mol Biol*. **16**, 1043-8.
65. Liu, D., Ishima, R., Tong, K. I., Bagby, S., Kokubo, T., Muhandiram, D. R., Kay, L. E., 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.
66. Donczew, R., Warfield, L., Pacheco, D., Erijman, A. & Hahn, S. (2020) Two roles for the yeast transcription coactivator SAGA and a set of genes redundantly regulated by TFIID and SAGA, *eLife*. **9**.
67. Petrenko, N., Jin, Y., Dong, L., Wong, K. H. & Struhl, K. (2019) Requirements for RNA polymerase II preinitiation complex formation in vivo, *eLife*. **8**.
68. Ramachandran, S., Haddad, D., Li, C., Le, M. X., Ling, A. K., So, C. C., Nepal, R. M., Gommerman, J. L., Yu, K., Ketela, T., Moffat, J. & Martin, A. (2016) The SAGA Deubiquitination Module Promotes DNA Repair and Class Switch Recombination through ATM and DNAPK-Mediated gammaH2AX Formation, *Cell Rep*. **15**, 1554-1565.
69. Clouaire, T., Rocher, V., Lashgari, A., Arnould, C., Aguirrebengoa, M., Biernacka, A., Skrzypczak, M., Aymard, F., Fongang, B., Dojer, N., Iacovoni, J. S., Rowicka, M., Ginalski, K., Cote, J. & Legube, G. (2018) Comprehensive Mapping of Histone Modifications at DNA Double-Strand Breaks Deciphers Repair Pathway Chromatin Signatures, *Mol Cell*. **72**, 250-262 e6.
70. Evangelista, F. M., Maglott-Roth, A., Stierle, M., Brino, L., Soutoglou, E. & Tora, L. (2018) Transcription and mRNA export machineries SAGA and TREX-2 maintain monoubiquitinated H2B balance required for DNA repair, *J Cell Biol*. **217**, 3382-3397.
71. Cabal, G. G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J. C., Hurt, E. C. & Nehrbass, U. (2006) SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope, *Nature*. **441**, 770-3.
72. Shao, W., Ding, Z., Zheng, Z. Z., Shen, J. J., Shen, Y. X., Pu, J., Fan, Y. J., Query, C. C. & Xu, Y. Z. (2020) Prp5-Spt8/Spt3 interaction mediates a reciprocal coupling between splicing and transcription, *Nucleic Acids Res*.
73. Denoth-Lippuner, A., Krzyzanowski, M. K., Stober, C. & Barral, Y. (2014) Role of SAGA in the asymmetric segregation of DNA circles during yeast ageing, *eLife*. **3**.

74. Goddard, T. D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H. & Ferrin, T.E. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25.

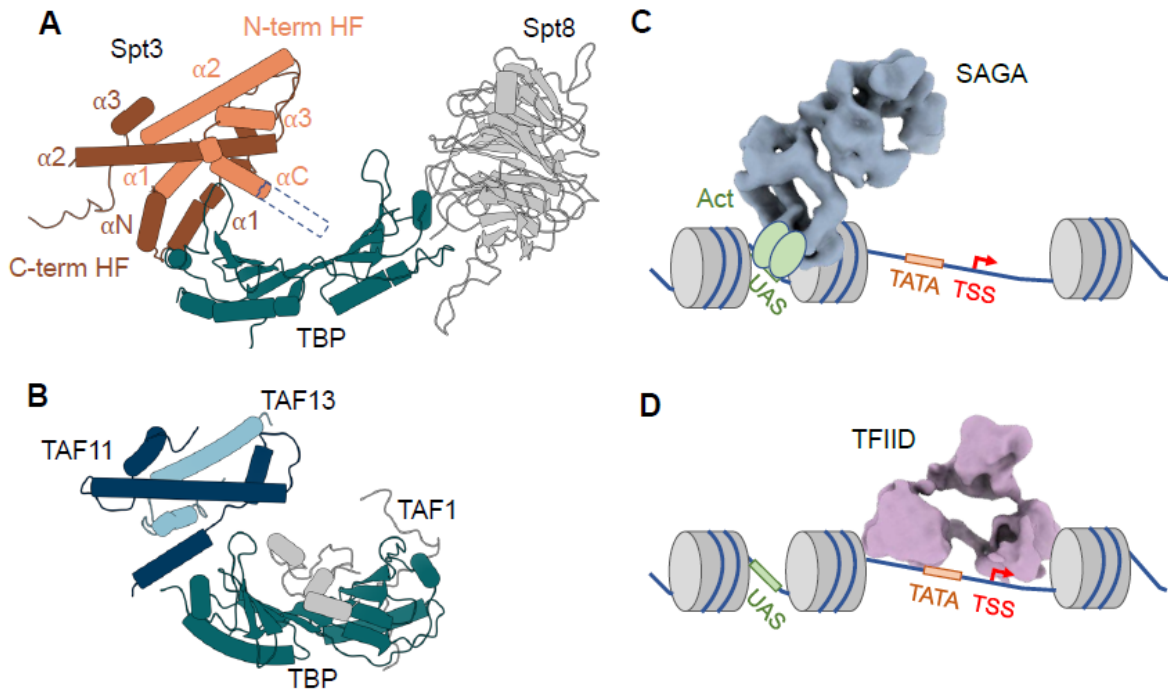
## FIGURE LEGENDS



**Figure 1 | Architecture of the SAGA-TBP complex. A.** SAGA subunits are organized into the Tra1-lobe and the main lobe. The main lobe is divided into two enzymatic HAT and DUB modules, a central module and a TBP-binding module. Overlaps appear as the central module is a scaffold for the assembly of all other modules. The Histone Fold domains, present in seven subunits, are highlighted as yellow boxes. Subunit length are indicated for the yeast *Pichia pastoris*. **B.** Front and side views of the yeast SAGA-TBP complex cryo-EM map in which the modules have been colored as in (A). **C.** Front views of the atomic model of the SAGA-TBP complex. Arrows indicate the anchor regions of the DUB and the HAT modules. All displayed images were generated using UCSF ChimeraX<sup>60</sup>



**Figure 2 | Histone fold arrangements in TFIID and SAGA.** **A.** Composition of the histone fold arrangements found in TFIID and SAGA. (Spt3)<sub>2</sub> represents the intramolecular histone fold dimer present in Spt3. **B., C.** Arrangement of Taf5 and histone-fold containing subunits in yeast SAGA (B) and TFIID (C). The Taf5-NTD and Taf6 HEAT repeats adopt different positions in yeast SAGA and TFIID.



**Figure 3 | TBP interactions and promoter binding.** **A.** Molecular interactions of TBP with the SAGA subunits Spt3 and Spt8. The partially unfolded  $\alpha C$  helix of the Spt3 N-terminal HF is represented by dotted lines. **B.** Model for the position of TBP within TFIID and its interaction with the TAF11-13 HF pair and the N-terminal TAND domains of TAF1 (grey), Adapted from [17]. **C.** Model for the interaction of SAGA with transcriptional activators (Act) bound to UAS sequences upstream of the TSS. **D.** Model for the interaction of TFIID with proximal promoter sequences surrounding the TSS.