# UNIVERSITÉ DE STRASBOURG

*ÉCOLE DOCTORALE 414 - Sciences de la Vie et de la Santé*

**IGBMC - CNRS UMR 7104 - Inserm U 1258 - Université de Strasbourg**

# THÈSE présentée par :

**Paul BARDOT**

soutenue le : **28 juin 2018**

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

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

# Analyse de la composition et de la fonction de la machinerie basale de transcription au cours du développement et de la différenciation

**THÈSE dirigée par :**

| | |
|---|---|
| **M. TORA Làzlò** | Directeur de recherche, IGBMC, Université de Strasbourg, France |

**RAPPORTEURS :**

| | |
|---|---|
| **Mme. DOSTATNI Nathalie** | Directeur de recherche, Institut Curie, Paris 6, France |
| **M. TIMMERS Marc** | Professeur, German Cancer Consortium (DKTK) Partner Site Freiburg, German Cancer Research Center (DKFZ), Allemagne |

**AUTRES MEMBRES DU JURY :**

| | |
|---|---|
| **M. WEBER Michaël** | Directeur de recherche, Biotechnologie et signalisation cellulaire, Université de Strasbourg, France |

# ACKNOWLEDGEMENTS

La thèse est faite de haut et de bas qui nécessite de la volonté et surtout un soutien de nombreuses personnes. Je tiens ici à présenter les remerciements à toutes celles et tous ceux qui m'ont aidé, à leur échelle, à réussir ce défi.

First of all, I would like to thanks Nathalie Dostatni, Andreas Mayer, Marc Timmers and Michaël Weber for accepting and taking time to judge my thesis. I would like also to thank my mid-thesis committee for their kind guidance and precious advice: Gérard Gradwohl, Gabor Papai and Marc Timmers.

Cette thèse a pu avoir lieu grâce à la confiance que m'a accordée Làszlò, afin de continuer ce projet, et qui a toujours été bienveillant à mon égard. Je ne puis oublier Stéphane, qui m'a permis d'étudier la biologie du développement sous un angle nouveau et qui a fait preuve d'une grande disponibilité et d'une grande patience. Merci de m'avoir transmis ta rigueur et ton regard critique, que je n'aurai de cesse de continuer à approfondir par la suite. Je vous remercie également tous les deux, ainsi que Vincent, d'avoir pris le temps de relire mon manuscrit.

A l'IGBMC, j'ai pu profiter des services des différentes plateformes et je remercie ici, Betty, Amélie, Marion et toute l'équipe de culture cellulaire pour la préparation des milieux, pour leur sérieux et leur disponibilité. A l'animalerie, je remercie Sylvie et Martine qui ont toujours été d'une grande aide. Je remercie aussi Catherine Birck pour le service de filtration sur gel. A la cytométrie en flux, j'ai pu compter sur Claudine et Muriel. Les personnes de l'administration ont également été très précieuses, merci à Francine, Annick, France, Sandra et Mathilde.

Au sein du labo Tora, je souhaite remercier chacun de ses membres, présents et passés avec qui j'ai passé de bons moments. Plus que des collègues, ce sont des amitiés qui se sont forgées à travers le temps. Changwei et Sascha mes plus proches voisins de paillasse avec qui on pouvait discuter de tout. Eli, toujours bienveillante et d'une grand complicité, cela a été un plaisir de discuter avec toi, notamment à propos de culture. Tiago, Pooja, Ivanka, Gizem, Nikolaos, Federica, Véronique, Farah, Kenny, Vincent, Matthieu, Fang, Michael, Didier, Emma, David vous avez tous été des personnes très agréables avec qui j'ai apprécié passer du temps au laboratoire et en-dehors. Je n'oublierai pas la chance que j'ai eu d'évoluer dans un

environnement aussi agréable. Une pensée pour David Umlauf qui m'a supervisé lors de mon stage dans ce même laboratoire en master 1.

L'aventure a commencé au sein du laboratoire d'Olivier Pourquié, et je remercie tous les membres que j'ai connu à cette époque pour les bons moments passés. Notamment Olivier, Caroline et Goncalo avec qui c'est toujours un plaisir de se retrouver autour d'un verre.

Je n'oublie pas aussi tous ceux qui m'ont donné goût à la recherche. Marc Leborgne qui m'a ouvert les portes se son laboratoire à Lyon pour ma toute première expérience de recherche. Par la suite, je remercie Rémi Mounier et Bénédicte Chazaud de m'avoir accueilli dans leur laboratoire à l'institut Cochin où j'ai vraiment pris goût à la recherche.

Merci à tous mes collègues doctorants du SPB, de l'Addal, du Réseau BIOTechno pour ces moments partagés dans des projets associatifs, qui m'ont permis aussi de diversifier mes compétences et de rencontrer des personnes agréables, dynamiques et curieuses.

La thèse aura réellement été une course d'endurance au sens figuré, qualité que j'ai pu également développer au sens strict grâce à tous mes amis du club cycliste de l'UC Vendenheim Serge, Etienne, Loïc, Cindy, Yann, Gaël, Stéphane et tous les autres. Vous m'avez appris à ne jamais lâcher et à me dépasser, je n'en ressors que plus fort après ces quatre ans de partage d'une passion commune autour du vélo.

Je me rappellerai aussi tous les bons moments passés au cours de ma thèse avec mes amis du lycée, qui me connaissent bien : Adrien C., Adrien S., Alexis, Angéline, Benji, Mickaël, Rémi, Pauline

Depuis le début de mes études, mes parents m'ont toujours fait confiance, sans me pousser vers une voie quelconque, et m'ont soutenu dans tous les choix que j'ai pu faire. Ils savent plus que n'importe qui, combien ce diplôme est important à mes yeux. Merci de m'avoir permis d'étudier, de m'avoir aidé et d'être toujours présent quand il le faut. Merci également à mes frères et sœur pour les retrouvailles en famille qui ont contribué à mon équilibre personnel.

Enfin, plus personnellement, je tiens à remercier Vanessa qui m'a toujours soutenu et supporté en plus de m'avoir permis de vivre pleinement épanoui.

## ABSTRACT

During development, tightly regulated gene expression programs control cell fate and patterning. In eukaryotes, transcription initiation requires the assembly of the preinitiation complex at promoters. TFIID is the first general transcription factor to bind the promoter, and is essential for RNA Polymerase II recruitment. TFIID is an evolutionary conserved multi-subunit complex composed of TBP and 13 TAFs in metazoans. However, TFIID composition was shown to be variable. In HeLa cells, at least two types of TFIID complexes co-exist, and in addition several paralogs of TBP and TAFs exist. Most of the data concerning the composition of TFIID comes mainly from *C. elegans*, *Drosophila* and human cells for metazoans. However, little is known about the exact composition of TFIID *in vivo* in metazoans and its biological role. The first goal of my thesis was to characterize the precise composition of TAF10-containing complexes in metazoans, TFIID and SAGA, in the embryo and in different cellular contexts. TFIID and SAGA composition was analyzed in thymocytes, mES cells and in the whole embryo at E9.5 by immunoprecipitation followed by mass spectrometry. All the TFIID and SAGA subunits were detected with some differences depending on the cellular context. TFIID sub-complexes could also be detected in the nucleus of mES cells. I showed that TAF10 is essential for the assembly of TFIID and SAGA complexes. In order to determine the role of TAF10-containing complexes during development, *Taf10* was conditionally deleted in the embryonic mesoderm derivatives at E7.5. It resulted in an efficient TAF10 protein loss in the presomitic and lateral plate mesoderm from E8.5. The conditional deletion of *Taf10* resulted in a growth arrest at E9.5 and initial paraxial mesoderm differentiation was not prevented in the absence of TAF10 whereas lateral plate differentiation was altered. At E9.5, steady state mRNA levels were unchanged in the presomitic mesoderm, with only a minor subset of genes dysregulated. Our data suggest that the canonical TFIID and SAGA are dispensable for early paraxial mesoderm development, arguing against their proposed generic role in transcription. To better understand their role in transcription, newly-transcribed mRNA have been analyzed in mES cells after the inducible deletion of *Taf10*. The results obtained here indicate that mRNA synthesis is strongly impaired for most of the genes tested and suggest a global role for TFIID and SAGA in transcription initiation in mammals. It was observed that the steady-state mRNA levels were not altered to the same extent than mRNA synthesis indicating that there might be a compensation. Ongoing genome-wide analyses will provide a more comprehensive view of the transcription dependency on TFIID and SAGA.

# RESUME EN FRANÇAIS

## Introduction

La transcription est une étape essentielle à la mise en place des programmes d'expression génique qui contrôlent les processus de développement. Chez les eucaryotes, la transcription des pré-ARNm est catalysée par l'ARN polymérase II (Pol II) associée à une machinerie basale. D'autres complexes tels que le Médiateur, les enzymes de remodelage de la chromatine et les facteurs de transcription sont impliqués dans la coordination de l'activation de la transcription. Chez les eucaryotes, un des complexes capable de modifier la chromatine est le complexe SAGA (Spt-Ada-Gcn5 acétyltransferase). L'initiation de la transcription au promoteur est contrôlée par l'assemblage séquentiel et ordonné des facteurs généraux de transcription en un complexe de pré-initiation (PIC). Parmi eux, TFIID (facteur IID) intervient dans la reconnaissance des séquences caractéristiques du promoteur. Chez la levure, TFIID a été montré comme étant nécessaire pour la transcription par l'ARN Pol II de la quasi-totalité du génome.

TFIID est un complexe multi-protéique composé de TBP (*TATA-binding protein*) et de 13 TAFs (*TBP-associated factors*) chez les métazoaires et de 14 Tafs chez la levure, et qui sont conservées à travers l'évolution. Le complexe TFIID a été initialement caractérisé chez plusieurs espèces dont la drosophile, la levure et les cellules humaines. Un grand nombre des sous-unités TAFs se caractérisent par la présence d'un domaine HFD (*Histone fold domain*) permettant leur dimérisation. Cette caractéristique en fait des protéines structurales importantes pour l'assemblage de TFIID. D'autres sous-unités paralogues de TBP et de certains TAFs ont également été découvertes et identifiées chez les métazoaires : TRF1/TLF1 (*TBP related factor/TBP like factor 1*), TRF2, TRF3/TBPL2, TAF4 et TAF4b, TAF7 et TAF7L et TAF9 et TAF9b. L'existence de machineries basales de transcription spécifiques de certains tissus a ainsi été démontrée dans les cellules germinales et dans les adipocytes. Par ailleurs, dans les cellules humaines cancéreuses HeLa, il a été montré qu'au moins deux types de complexes TFIID coexistent, avec ou sans TAF10. Toutes ces données montrent que la composition de la machinerie basale de transcription peut varier selon le contexte cellulaire. Le modèle structural de TFIID obtenu d'après la surexpression de ses sous-unités dans les cellules consiste en un cœur structural composé de TAF4, TAF5, TAF6, TAF9, TAF12 complété par TAF2-TAF8-

TAF10, le complexe « huit-TAFs ». Enfin, le complexe « holo-TFIID » est formé par l'incorporation finale de TBP, TAF1, TAF3, TAF7, TAF11 et TAF13.

Le complexe SAGA partage certains TAFs avec TFIID comme TAF9, TAF10, TAF12 ainsi que les paralogues de TAF5 et TAF6 (au moins chez l'humain et la souris): TAF5L et TAF6L respectivement. Ces TAFs appartiennent au cœur structural de SAGA, qui est donc très similaire au cœur structural de TFIID. SAGA présente une organisation modulaire, à savoir le cœur structural, le module de dé-ubiquitinylation (permettant la dé-ubiquitinylation de l'histone H2B1), le module d'acétylation des histones (permettant l'acétylation de la lysine 9 de l'histone H3) et interagit avec TBP. SAGA partage aussi certaines de ses sous-unités avec d'autres complexes transcriptionnels. Récemment, SAGA a également été montré comme étant requis pour la transcription de la quasi-totalité du génome chez la levure. TFIID et SAGA sont donc deux complexes clés pour la transcription catalysée par l'ARN Pol II.

La plupart des TAFs ont un rôle dans l'assemblage des complexes TFIID et de SAGA, et ne possèdent pas d'activité enzymatique connue, à l'exception de TAF1 pour qui des activités controversées d'acétyltransférase, de kinase et d'ubiquitinylation ont été décrites. Chez les métazoaires, TFIID est important pour permettre l'activation de la transcription *in vitro*, et interagit physiquement avec des co-activateurs et des facteurs de transcription. Cela dit, chez les métazoaires, la fonction de TFIID *in vivo* est moins bien connue. La mutation nulle des sous-unités *Tbp*, *Taf7*, *Taf8* et *Taf10* chez la souris est létale, liée à un défaut d'implantation du blastocyste. Il a été montré que la transcription par l'ARN Pol I et l'ARN Pol III dans les blastocystes $Tbp^{-/-}$ était affectée mais pas la transcription catalysée par l'ARN Pol II. TAF10, une sous-unité ubiquitaire, est une protéine d'architecture nécessaire à l'assemblage de TFIID. Bien qu'ubiquitaire, TAF10 présente cependant un phénotype différent en fonction du contexte cellulaire et du stade de développement. Il a été montré que TAF10 est nécessaire à l'embryogenèse précoce chez la souris. En culture, les cellules de la masse interne $Taf10^{-/-}$ meurent massivement par apoptose alors que les cellules du trophectoderme (les trophoblastes) $Taf10^{-/-}$ sont viables. Cependant, l'endoréplication et la transcription par l'ARN Pol II sont bloquées dans les trophoblastes $Taf10^{-/-}$. TAF10 joue donc un rôle important lors de la transcription selon le contexte cellulaire. Par ailleurs, il a aussi été montré que la transcription se poursuit chez l'adulte en l'absence de TAF10 dans les cellules de l'épiderme et du foie. La

sous-unité TAF7 a également été montrée comme étant différemment requise lors de la maturation des cellules de thymocytes ainsi que pour la transcription en fonction du type cellulaire étudié. L'ensemble de ces résultats suggèrent donc une variabilité de la composition de la machinerie basale de transcription, selon le contexte cellulaire et le stade de développement.

La variabilité de la composition de TFIID ainsi que sa fonction *in vivo* reste mal connue. Afin de mieux comprendre le rôle de la machinerie basale de transcription, nous avons choisi d'utiliser le développement embryonnaire comme paradigme. En effet, au cours du développement, la cellule œuf est à l'origine des différents types cellulaires, aux fonctions très diverses et spécialisées, de l'organisme à partir d'une information génétique unique. Pour cela, le contrôle de l'expression des gènes joue un rôle essentiel. De plus, les programmes transcriptionnels contrôlent les processus cellulaires à l'œuvre dans l'embryon, telles que la prolifération et la différenciation des cellules ou la morphogenèse. Afin d'étudier la machinerie basale de transcription pendant le développement, nous avons choisi ici le mésoderme présomitique (PSM).

Le PSM est le siège d'un processus particulier dans l'embryon : la segmentation. Chez les vertébrés et la plupart des insectes, la segmentation est un processus séquentiel. Le PSM dérive du mésoderme paraxial qui forme deux bandes de tissu flanquant le tube neural. La segmentation du PSM au cours de l'embryogenèse, aussi appelée somitogenèse, est responsable de la formation des somites, précurseurs du squelette axial et des muscles du tronc et des membres. Chez la souris, une paire de somites est ajoutée toutes les deux heures à l'extrémité antérieure du PSM. La formation d'un nombre défini de segments de taille précise, spécifique à chaque espèce, est le résultat d'un contrôle spatio-temporel de l'expression des gènes. En effet, une horloge moléculaire et un front de détermination contrôlent la spécification de nouveaux segments. Le front de détermination est formé à l'intersection des gradients postéro-antérieurs de Wnt/FGF et d'un contre-gradient d'acide rétinoïque. Les forts niveaux de FGF maintiennent les cellules les plus postérieures à l'état de progéniteurs tandis qu'un contre-gradient d'acide rétinoïque induit la différenciation. Lorsqu'un groupe de cellules traversent le front de détermination, correspondant à un niveau seuil de gradient, ce groupe de cellules devient compétent à répondre au signal de l'horloge permettant la spécification d'un nouveau segment. L'horloge moléculaire est caractérisée par l'expression cyclique des gènes

des voies de signalisation Notch, Wnt et FGF. Certains de ces gènes codent des répresseurs qui permettent la mise en place de boucles de rétro-inhibition, à l'origine des oscillations moléculaires de l'expression des gènes. Ainsi, le temps nécessaire à la transcription des gènes impliqués dans des boucles de rétro-inhibition est un paramètre crucial à l'établissement de ces oscillations. Par conséquent, les étapes d'épissage et d'export nucléaire des ARNm suivant la transcription, ainsi que la demi-vie des ARNm constituent les paramètres de ces oscillations. En revanche, la transcription seule n'est pas un facteur limitant dans l'établissement des oscillations. L'expression dynamique des gènes liée à une transcription cyclique et l'expression classique des gènes dans le PSM, sous l'influence de cascades de signalisation, en font un paradigme intéressant pour l'étude de la transcription au cours du développement.

## Objectifs

La composition de TFIID est variable et a été principalement décrite chez la levure, la drosophile et les cellules humaines mais peu de choses sont connues quant à la fonction de ce complexe et de sa variabilité *in vivo* chez les métazoaires. En utilisant la sous-unité TAF10, partagée entre TFIID et SAGA, comme point de départ pour l'étude de la machinerie basale de transcription, les buts de ma thèse de doctorat étaient:

(1) De caractériser dans l'embryon sauvage la composition, qui n'a pas encore été déterminée de façon précise, de la machinerie basale de transcription, notamment TFIID et SAGA;
(2) De développer et de valider des approches alternatives permettant l'utilisation d'approches biochimiques classiques afin d'approfondir la caractérisation de TFIID et de SAGA dans différents contextes cellulaires;
(3) D'étudier le rôle de TFIID et SAGA au cours de la somitogenèse dans l'embryon à 9,5 j.p.c;
(4) D'analyser la contribution de TFIID et SAGA à la transcription catalysée par l'ARN Pol II dans un modèle mammifère.

## Résultats

Afin d'analyser les complexes TFIID et SAGA chez la souris, par immuno-précipitation couplée à la chromatographie liquide avec spectrométrie de masse en tandem, j'ai tout d'abord optimisé les conditions expérimentales. Pour cela j'ai testé l'efficacité de plusieurs anticorps,

pour la plupart initialement dirigés contre les épitopes de sous-unités humaines, à partir d'extraits cellulaires totaux de thymocytes. Les anticorps anti-TAF7 et anti-TBP ont été sélectionnés pour immuno-précipiter le complexe TFIID. Les anticorps anti-TRRAP et anti-SUPT3 ont été sélectionnés pour immuno-précipiter le complexe SAGA. La quantité de matériel a également été réduite de 4 mg à 0.7 mg afin de pouvoir réaliser ces expériences dans l'embryon de souris à 9,5 j.p.c (jours *post coitum*).

L'analyse à partir d'extraits cellulaires totaux d'embryons entiers à 9,5 j.p.c a révélé la présence du complexe TFIID canonique. Des résultats similaires ont été obtenus dans les thymocytes et dans les cellules souches embryonnaires murines (mES). Ces résultats clarifient la composition de TFIID dans l'embryon et indiquent qu'il s'agit du complexe canonique. De plus, pour la première fois dans l'embryon, la composition du complexe SAGA a été décrite. Toutes les sous-unités du complexe décrites dans la littérature ont été détectées en proportions différentes pour certaines entre les cellules de thymocytes, de mES et l'embryon. Ces données montrent que la composition de TFIID et SAGA est globalement conservée quel que soit le contexte cellulaire étudié ici avec quelques différences selon le type cellulaire.

La quantité de matériel à partir d'embryons étant limitée pour des approches biochimiques plus approfondies, les cellules mES ont été utilisées pour préparer des extraits nucléaires. L'analyse préliminaire par chromatographie d'exclusion à partir de ces échantillons a permis de détecter le complexe holo-TFIID mais aussi au moins deux autres sous-complexes à des poids moléculaires inférieurs à 1 MDa. Parmi les anticorps testés par western-blot, un premier sous-complexe contenant au moins TAF5, TAF8, TAF10 et TBP et le second de 670 kDa contenant au moins TAF5 et TBP ont été détectés. Ces résultats confirment la présence de sous-complexes de TFIID dont la composition reste à déterminer avec précision.

La délétion inductible du gène codant la sous-unité TAF10 dans l'embryon entier à 7,5 j.p.c et dans les cellules mES a permis la déplétion complète de la protéine dans l'embryon à 9,5 j.p.c et quasi totale dans les cellules mES deux et cinq jours plus tard. Les expériences d'immuno-précipitation contre TAF7 et TBP à partir d'extraits cellulaires totaux ont montré que les sous-unités de TFIID associées étaient faiblement détectées indiquant que TAF10 est requis pour l'assemblage complet de TFIID. Ces résultats ont été confirmés par la chromatographie d'exclusion à partir d'extraits nucléaires de cellules mES mutantes pour *Taf10*, pour lesquels le complexe holo-TFIID était encore détecté mais sans certaines sous-

unités tandis que les deux autres sous-complexes étaient toujours présents. Cela indique qu'en l'absence de TAF10, peu de complexes holo-TFIID sont retrouvés et que majoritairement il s'agit de complexes TFIID partiellement assemblés. Toutefois, il a été noté que l'assemblage de TFIID était plus fortement affecté dans l'embryon que dans les cellules mES en l'absence de TAF10, soulignant des différences liées au contexte cellulaire.

TAF10 a également été montré comme étant requis pour l'assemblage complet du complexe SAGA dans l'embryon et les cellules mES. Comme pour TFIID, l'assemblage du complexe était plus fortement affecté dans l'embryon que dans les cellules mES. Toutefois, malgré ce défaut d'assemblage, les niveaux d'acétylation de l'histone H3 sur la lysine 9 ainsi que les niveaux d'ubiquitine de l'histone H2B1 n'étaient pas affectés en l'absence de TAF10, suggérant que le complexe SAGA reste fonctionnel.

Afin d'étudier le rôle de TFIID et SAGA au cours du développement, nous avons choisi de déléter *Taf10* dans le mésoderme paraxial où la transcription est dynamique avec certains gènes qui ont une expression cyclique. L'analyse de la délétion du gène codant la sous-unité TAF10 a été réalisée. La délétion de *Taf10* ubiquitaire inductible ou conditionnelle dans les dérivés du mésoderme à partir de 7,5 j.p.c, entraine la disparition totale de la protéine et provoque un sévère ralentissement de la croissance embryonnaire entre 9,5 j.p.c et 10,5 j.p.c. La délétion conditionnelle de *Taf10* dans les dérivés du mésoderme n'affecte pas la formation et la différenciation précoce des somites à 9,5 j.p.c. Toutefois, la délétion de *Taf10* n'est pas viable à long terme, puisqu'elle conduit à la mort de l'embryon vers 10,5 j.p.c, probablement en raison de la dégénérescence du placenta et de l'allantoïde chez les mutants *T-Cre ;Taf10*$^{flox/flox}$. L'analyse transcriptomique du PSM entre les embryons sauvages et mutants a montré que l'expression de la plupart des gènes n'est pas perturbée en l'absence de TAF10. Cependant, certains gènes dont *Cdkn1a* (*p21*) et *Cdkn1c* (*p57*) codant des inhibiteurs du cycle cellulaire, sont surexprimés chez les mutants, suggérant que l'arrêt de la croissance des embryons mutants est lié au blocage du cycle cellulaire et donc de la prolifération cellulaire. Ces résultats mettent en évidence que l'expression de la plupart des gènes est normale en l'absence de TAF10, alors que des travaux précédents ont montré que TAF10 est nécessaire à la transcription dans l'embryon. Toutefois, la transcription reste fonctionnelle seulement pendant un intervalle de temps précis en l'absence de TAF10 et les embryons mutants meurent à partir de 10,5 j.p.c. Cela suggère donc que la transcription peut dans certaines conditions

fonctionner avec une machinerie basale alternative sans TAF10 dans l'embryon. Cette fenêtre de temps représente donc une opportunité pour étudier la machinerie qui permet la transcription dans l'embryon en l'absence de TAF10. Un autre fait marquant est l'effet différent produit par la délétion de *Taf10* selon les tissus. En effet, le mésoderme de la plaque latérale (LPM), qui est également ciblé par la délétion conditionnelle de *Taf10*, présente une mort cellulaire importante à 9,5 j.p.c alors que le mésoderme paraxial est normal. Cela suggère une variabilité de la machinerie de transcription en fonction du contexte comme cela a été démontré dans les cellules du trophectoderme et les cellules de la masse interne. L'ensemble des résultats confirment que la transcription dans l'embryon est particulière, avec des spécificités propres à certains tissus.

L'induction de la délétion de *Taf10* dans les cellules mES provoque un fort ralentissement de leur croissance trois jours après, récapitulant le phénotype observé dans l'embryon. Toutefois, aucune mort cellulaire massive n'a été détectée, suggérant une diminution de leur prolifération. L'analyse de la prolifération par incorporation d'EdU n'a pas permis de révéler une forte réduction de la prolifération trois jours après l'induction de la délétion, mais il n'est pas exclu qu'elle le soit par la suite, quatre ou cinq jours après l'induction de la délétion. Ces résultats indiquent que les cellules mES se comportent différemment des cellules de la masse interne, qui sont incapables de survivre dans les blastocystes *Taf10*$^{-/-}$ comme cela a été rapporté dans la littérature et permettent donc de servir comme un bon modèle d'étude pour l'analyse du rôle de TFIID et SAGA dans la transcription. Les cellules de tératocarcinome murines F9, en revanche, sont quasiment incapables de se maintenir et de proliférer en culture en l'absence de TAF10 confirmant le phénotype décrit dans la littérature et soulignant ainsi l'importance du contexte cellulaire dans les différences de phénotype obtenues avec la délétion de *Taf10*.

De façon surprenante, nous avons montré que TAF10 n'est pas indispensable pour l'expression globale des gènes dans le PSM, à l'exception de certains gènes. La délétion de *Taf10* dans les cellules mES impactent différemment les niveaux d'ARN totaux des gènes testés par RT-qPCR, avec l'augmentation des niveaux d'ARN totaux pour *Cdkn1a* et *Cdkn1c* notamment, qui récapitulent ce qui a été observé dans le PSM. De même, *Gas5* et *Taf1d*, dont les niveaux d'ARN totaux sont diminués dans le PSM sont également diminués dans les cellules mES mutantes. Il a été montré chez la levure qu'une diminution du taux de la synthèse des

ARNm peut être compensée par une diminution du taux de la dégradation des ARN, normalisant ainsi les niveaux des ARN totaux. Par conséquent, les niveaux d'ARN totaux peuvent ne pas toujours refléter l'état de la transcription. Afin d'analyser directement la transcription, j'ai adapté la technique de marquage métabolique des ARN nouvellement synthétisés aux cellules mES. Cette technique repose sur le marquage pendant un temps court, ici 10 minutes, des ARNm en cours de synthèse grâce à l'incorporation de l'analogue de l'uridine, le 4-thiouridine (4sU). Les ARN marqués sont ensuite biotinylés et purifiés avec des billes de streptavidine magnétiques. L'analyse par RT-qPCR d'un panel de gènes a montré que la transcription est fortement affectée pour la quasi-totalité des gènes, avec une réduction d'au moins 50 % du niveau des ARNm nouvellement synthétisés. TAF10 est donc requis à la transcription au moins pour ces gènes, et il apparait qu'un phénomène important de compensation a lieu dans les cellules mES. En effet, plus que la normalisation des niveaux d'ARN totaux, il y a une surcompensation liée à une probable plus forte diminution du taux de dégradation des ARN par rapport à leur taux de synthèse. Une stabilisation accentuée des transcrits résultant en une accumulation plus forte des transcrits pourrait expliquer que les niveaux d'ARN totaux de certains gènes comme *Cdkn1a et Cdkn1c* sont fortement augmentés.

Dans le but d'étudier l'état de la transcription de façon globale au niveau de tout le génome, j'ai tenté dans un premier temps d'utiliser la technique *Transient Transcriptome sequencing*. Cette technique repose sur le marquage métabolique avec le 4sU des ARN nouvellement synthétisés avec une étape préliminaire de fragmentation des ARN. Cette étape permet de ne purifier uniquement les fragments d'ARNm marqués au 4sU et donc de s'affranchir du biais lié aux fragments d'ARNm déjà synthétisés avant le début du marquage et qui ne correspondent pas à de la transcription naissante. Malgré plusieurs tentatives, le manque de reproductibilité du profil des ARNm fragmentés associé à la faible efficacité de purification des ARNm nouvellement synthétisés, n'ont pas rendu possible leur séquençage. C'est pourquoi, la technique de 4sU-sequencing des ARNm purifiés après leur marquage par le 4sU a été utilisée. Les résultats étant en cours d'analyse, je ne dispose pas lors de la rédaction de ce manuscrit d'informations concernant l'état global de la transcription dans les cellules mES après la délétion de *Taf10*.

## Conclusions

Au cours de ma thèse, j'ai pu clarifier la composition des complexes TFIID et SAGA dans l'embryon ainsi que les différences de composition en fonction du contexte cellulaire. J'ai également montré que la sous-unité TAF10 est requise pour l'assemblage complet des complexes TFIID et SAGA. De plus, j'ai montré que le rôle général de TFIID et SAGA peut être nuancé dans l'embryon. Enfin, j'apporte ici des éléments indiquant que l'initiation de la transcription est sévèrement affectée en l'absence de TAF10, suggérant que le rôle global de TFIID et SAGA dans la transcription pourrait être conservé dans les cellules de mammifères.

Publication:

P. Bardot, S. D. Vincent, M. Fournier, A. Hubaud, M. Joint, L. Tora, and O. Pourquié, "The TAF10-containing TFIID and SAGA transcriptional complexes are dispensable for early somitogenesis in the mouse embryo.," *Development*, vol. 144, no. 20, pp. 3808–3818, Oct. 2017.

Communications:

Bardot P, Platania A, Pourquié O, Tora L and Vincent S. Characterization and functional analyses of the basal transcription machinery during development and cellular differentiation. Communication sous forme d'affiche. Tri-regional stem cell & developmental biology meeting. 9 Décembre 2016, IGBMC, Illkirch.

Paul Bardot, Olivier Pourquié, László Tora, Stéphane D. Vincent. Analysis of the TFIID and SAGA transcriptional complexes composition and function during development. Communication orale. Journées du campus d'Illkirch. 28-29 Mars 2017, ESBS, Illkirch.

Paul Bardot, Olivier Pourquié, László Tora, Stéphane D. Vincent. Analysis of the TFIID and SAGA transcriptional complexes composition and function during development. Présentation sous forme d'affiche. Keystone symposia, « Gene Control in Development and Disease (X6) ». 23-28 Mars 2018, Whistler.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ANNEXES

# LIST OF ABBREVIATIONS

**4OHT**: 4-hydroxy tamoxifen

**4sU**: 4-thiouridine

**4TU**: 4-tiouracile

**Ac**: acetylation

**AF**: Activation function containing region

**ARC**: Activator Recruited Factor

**ATAC**: Ada2a containing complex

**ATP**: Adenosine Tri Phosphate

**BrdU**: 5-Bromo-2-deoxyuridine

**BRE**: TFIIB Recognition Element

**CC**: Close Complex

**CDK**: Cyclin Dependent Kinase

**cDTA**: comparative dynamic transcriptome analysis

**CHD**: Chromodomain Helicase DNA-binding

**ChEC-seq**: Chromatin Endogenous Cleavage sequencing

**ChIP-seq**: Chromatin Immuno-precipitation Sequencing

**Cre**: Cre recombinase

**CTD**: Carboxyl Terminal Domain

**DNA**: Deoxyribo Nucleic Acid

**DPE**: Downstream Promoter Element

**DRIP**: VDR-interacting proteins

**DSIF**: DRB sensitivity-inducing factor

**DTT**: dithiothreitol

**DUB**: De Ubiquitinylation

**EC**: Elongation Complex

**EDTA**: Ethylene Diamine Tetra Acetic Acid

**EdU**: 5-ethynyl-2-deoxyuridine

**EM**: Electron Microscopy

**EtOH**: Ethanol

**FGF**: Fibroblast Growth Factor

**GTF**: General Transcription Factor

**HAT**: Histone Acetyle Transferase

**HCl**: Chloridric acidHDAC: Histone Deacetylase

**HFD**: Histone Fold Domain

**Inr**: Initatior

**IP**: Immuno-precipitation

**ITC**: Initial Transcribing Complex

**LFNG**: Beta-1,3-N-acetylglucosaminyltransferase lunatic fringe

**LIF**: Leukemia Inhibitory Factor

**LPM** : Lateral Plate Mesoderm

**Me**: Methylation

**mES**: mouse embryonic stem cell

**MgCl2**: Magnesium Chloride

**miRNA**: micro RNA

**MTE**: Motif Ten Element

**MudPIT**: Multidimensional Protein Identification Technology

**NaCl**: Sodium Chloride

**NELF**: Negative Elongation Factor

**NFR**: Nucleosome Free Region

**NLS**: Nuclear Localizing Sequence

**NSAF**: Normalized Spectral Abundance Factor

**OC**: Open Complex

**PBS**: Phosphate Buffer Saline

**PIC**: Pre Initiation Complexe

**PM**: Paraxial Mesoderm

**PSMx**: Peptide Spectrum Match

**PSM**: Pre Somitic Mesoderm

**RARE**: Retinoic Acid Response Element

**Rcf**: Relative centrifugal force

**RNA Pol**: Ribo Nucleic Acid Polymerase

**RNA**: Ribo Nucleic Acid

**RPB**: RNA Pol II subunit

**SAGA**: Spt Ada Gcn5 Acetyltransferase

**SLIK**: SAGA Like complex

**SPT**: Suppressor of Ty

**SWI/SNF**: Switch/Sucrose Non Fermentable

**TAF**: TATA-Binding Protein Associated Factors

**TBP**: TATA-Binding Protein

**TBPL**: TBP Like Factor

**TCT**: TcT-E element

**TFII**: Transcription Factor II

**TFTC**: TBP Free TAF complex

**TLF**: TBP Like Factor

**TRAP**: Thyroid Hormone Receptor Associated Proteins

**TRF**: TBP Related Factor

**TSS**: Transcription Start Site

**UAS**: Upstream Activator Sequence

**UTR**: UnTranslated Region

**WHHERE**: Wdr5, HDAC1, HDAC2, RERE/ATROPHIN2

# INTRODUCTION

In 1959, Weiss and Gladstone described an RNA polymerase activity from rat liver nuclei (Weiss et al. 1959). Transcription was defined as the conversion of DNA into RNA, and thus represents the first step of genome expression. Chromatographic analyses from purified nuclei from sea urchin and rat liver, and the transcription inhibition mediated by α-amanitin, brought the first evidences of the existence of three RNA Polymerases (RNA Pol) that catalyze transcription in the nucleus (**figure 1**): RNA Pol I, RNA Pol II and RNA Pol III (except plants that have five RNA Polymerases, reviewed in (Duda 1976)) (Kedinger et al. 1970; Roeder et al. 1969). While RNA Pol I transcribes large ribosomal RNAs, RNA Pol II transcribes protein coding genes, messenger RNAs, and also non-coding RNAs, and RNA Pol III transcribes small RNAs (tRNA, 5S RNA).

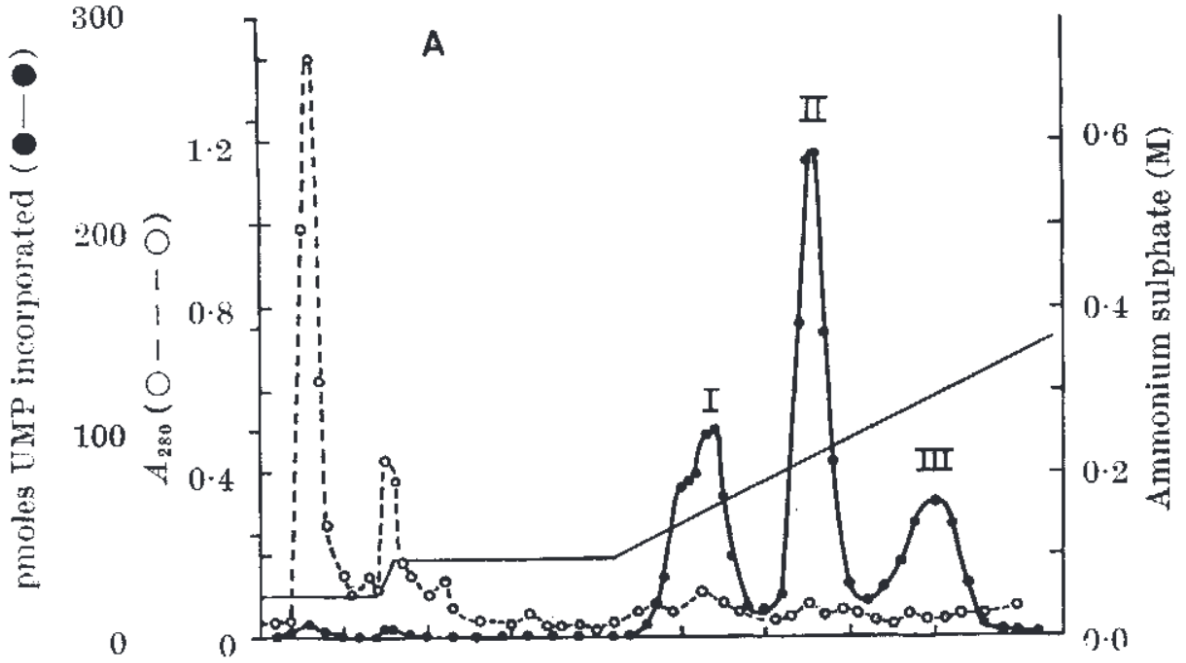



**Figure 1: Resolution of the three eukaryotic RNA Polymerases.** Activity measurement (units/µg protein) based on the incorporation of Uridine Mono Phosphate (UMP) into RNA/10min/ml of the fractions eluted by chromatography obtained from soluble enzyme preparation from sea urchin nuclei gastrula (52h development) (Roeder and Rutter 1969).

Transcription is a multi-step process composed of: (1) initiation with the recruitment of the RNA Pol II to the promoter, (2) elongation with productive mRNA synthesis and (3) termination that corresponds to the release of RNA Pol from DNA. Furthermore, transcription

represents a critical step for gene regulation, and plays major roles in the development of an organism by controlling many cellular processes as well as generating the different cell types.

The focus of my thesis concerns the general mechanisms of RNA Pol II transcription initiation. Firstly, I will describe how DNA is packaged in the eukaryotic nuclei and its implication for transcription. Then, I will detail the molecular mechanisms that govern RNA Pol II transcription with a special emphasis on the general transcription factor TFIID and the co-activator SAGA. Particularly, I will detail the importance of the components of the transcriptional apparatus in the control of gene expression *in vivo* and in a developmental context. Finally, I will present the new paradigm that I used to study RNA Pol II transcription in my thesis.

# I. Chromatin organization of the genome

## 1. Chromatin organization

Genetic information in eukaryotes is packaged into the nucleus of every cell. In the nucleus, DNA adopts a chromatin structure (**figure 2**), where it is wrapped around proteins called histones. Histone octamers form nucleosomes, the basic organization unit of chromatin. Each octamer contains two copies of histones: H2A, H2B, H3 and H4. Histone proteins form heterodimer through their Histone Fold Domain (HFD), which consists in three α-helices (α1, α2, and α3) connected by short loops L1 and L2 (**figure 3**) [reviewed in (Arents et al. 1991)]. Nucleosomes are separated between each other by a DNA linker. In addition, the histone H1 binds the nucleosome and allows a higher order of chromatin organization into 30 nm fibers and chromosomal 300 nm fibers. Chromatin is described as differentially condensed. Firstly, heterochromatin represents the highly condensed form and contains both non-coding gene regions (constitutive heterochromatin) and genes that are not expressed (facultative heterochromatin) (Bannister et al. 2011). Secondly, euchromatin is a decondensed chromatin that contain expressed genes. Chromatin organization is very dynamic and can change from one cell type to another, especially during development. Chromatin organization is of high importance as it impacts gene expression by modulating the accessibility of the transcription machinery on DNA. As a consequence, coding regions must be free of nucleosomes prior gene activation. The initial view of the chromatin organization has evolved, since the description of

at least five chromatin domains in *Drosophila* (Filion et al. 2010) and a more complex spatial organization of the genome with chromosome territories, A/B compartments, topologically associating domains, and chromatin loops [reviewed in (Serizay et al. 2018)].

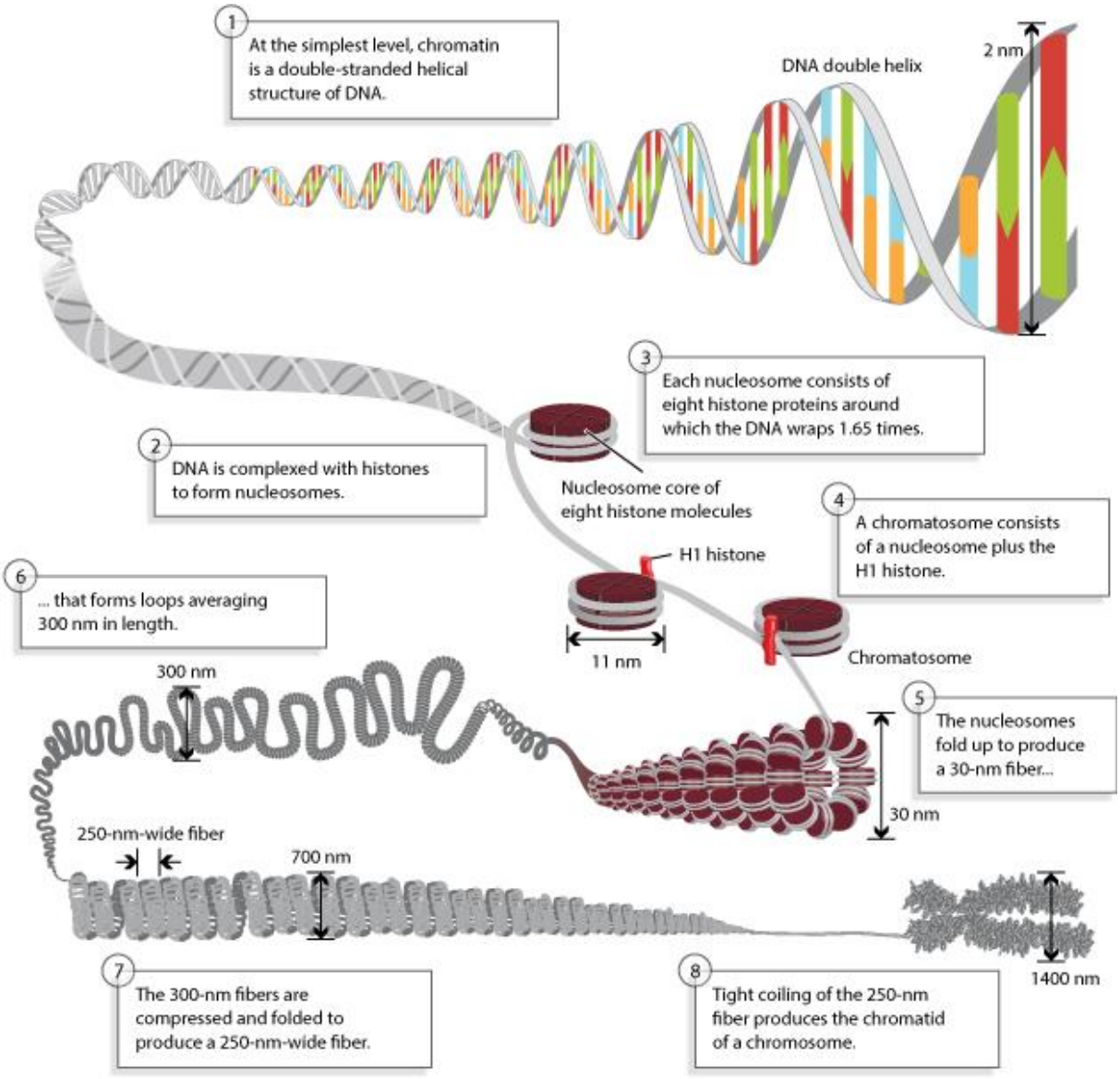



**Figure 2: Chromatin organization of the genome.** DNA is packaged in the nucleus under chromatin where it is wrapped around histone proteins. Multiple organisation levels structure the chromatin. Adapted from Pierce, Benjamin. Genetics: A Conceptual Approach, 2nd ed.

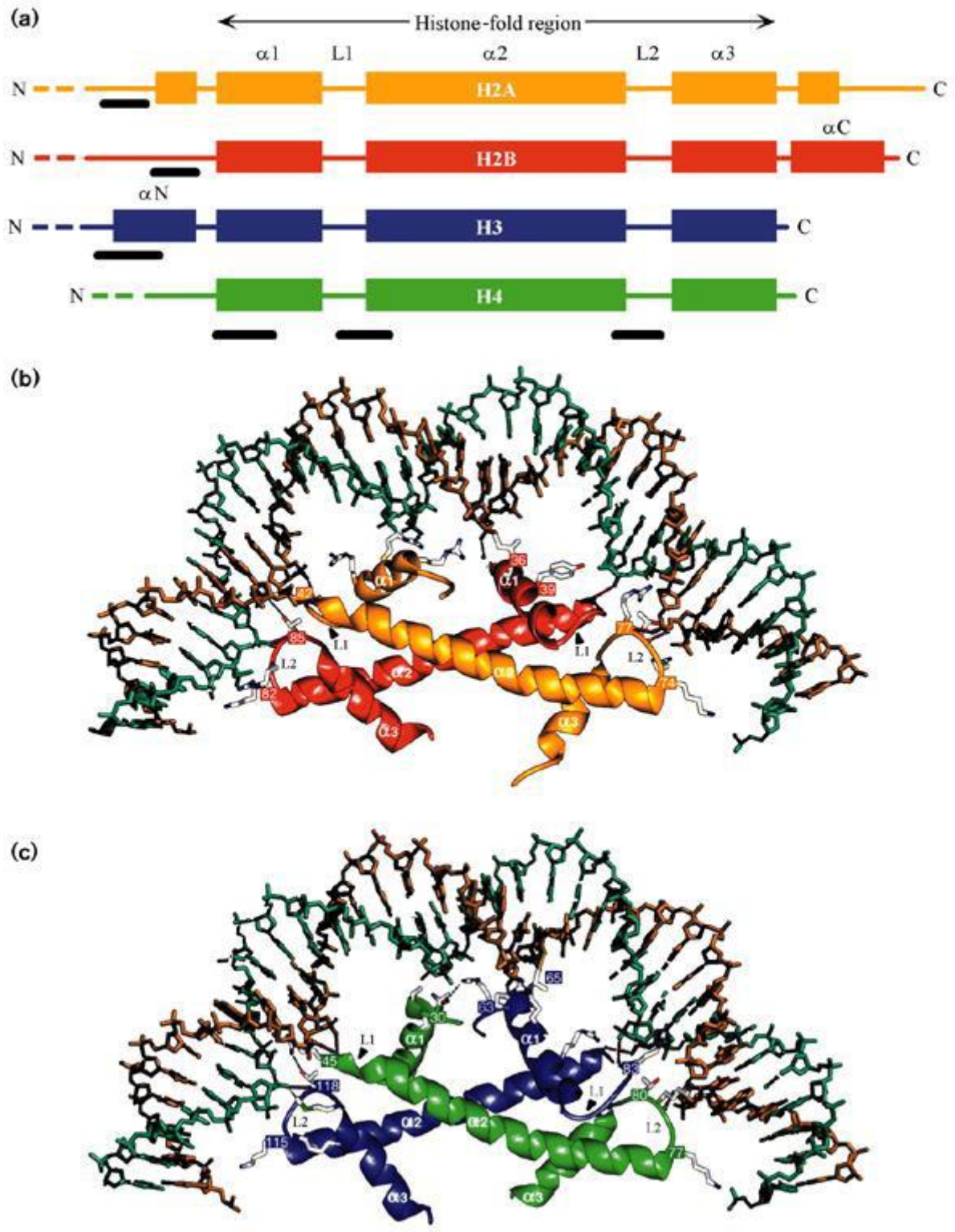



**Figure 3: Histone Fold Domain interactions.** (a) Scheme of the fours histone proteins with their HFD, helices correspond top rectangles and the black line represents the DNA interactions. Structure of (b) H2A-H2B dimer and (C) H3-H4 dimer (Dutnall and Ramakrishnan 1997).

## 2. Epigenetic modifications

Genome accessibility can be modulated by epigenetic modifications, that is to say, inherited modifications that do not alter the DNA sequence *per se* and that participate to the regulation of gene expression.

### a. DNA methylation

First of all, DNA can be directly methylated through the covalent transfer of a methyl group to the C-5 position of the cytosine ring, mostly at CpG dinucleotides [reviewed in (Li et al. 2014)]. Interestingly, DNA methylation of cytosine is not found at the same level in all species such as in *Drosophila*, where it was shown to be much lower (Boffelli et al. 2014), and nematodes for instance. Furthermore, it was shown that DNA can be also methylated on N(6)-adenine, as shown in several organisms (Greer et al. 2015; Liu et al. 2016; Wu et al. 2016; Zhang et al. 2015). In mammals, DNA methylation represents an epigenetic "lock" of gene expression. The methylated sequences include satellite DNAs, repetitive elements, non-repetitive intergenic DNA, and exons of genes. However, there are CpGs that remain unmethylated, and are found in CpG islands. Methylation is thought to play several roles in gene silencing such as X chromosome inactivation, for gene dosage compensation. Methylation of DNA controls the accessibility of transcription factors and can also be bound by specific factors that further recruit co-repressor complexes [reviewed in (Li et al. 2014)].

### b. Post-translational modifications of histone proteins

Amino-terminal tails of histones can also be modified and thus affect the nucleosome structure [reviewed in (Bannister et al. 2011)]. Several post-translational modifications of histones have been reported so far, and are listed in (**Table 1**) [reviewed in (Tessarz et al. 2014)]. Histone PTMs can affect directly chromatin compaction by modifying the chemical properties of their interactions with DNA. For instance, acetylation removes the positive charge from the histone tails rendering them neutral resulting in a more relaxed chromatin structure. PTMs are regulated by writer proteins that add PTMs on histones, while erasers remove them. PTMs are "read" by specific protein factors capable to recognize specific PTM or a combination of them, through specific domains [reviewed in (Lalonde et al. 2014)]. The recruitment of such proteins can also lead to chromatin modification and regulate transcription. Histone PTMs have

been proposed to constitute an additional biological "code" for regulating gene expression and epigenetic inheritance [reviewed in (Jenuwein et al. 2001)].

**Table 1: Overview of the histone post-translational modifications.** Adapted from (Tessarz and Kouzarides 2014).

| **Post-translational modification** | **Histone (residue)** | **Proposed function** |
|---|---|---|
| Methylation | H1 (Arginine, Lysine, Tyrosine)<br>H2A (Arginine, Lysine, Glutamine, Tyrosine)<br>H2B (Arginine, Lysine)<br>H3 (Arginine, Lysine)<br>H4 (Arginine, Lysine) | Chromatin compaction, rDNA transcription, Transcription, |
| Acetylation | H1 (Lysine)<br>H2A (Lysine)<br>H2B (Lysine)<br>H3 (Lysine)<br>H4 (Lysine) | Chromatin compaction, DNA repair, DNA replication, Transcription |
| Phosphorylation | H1 (Lysine, Tyrosine)<br>H2B (Serine)<br>H3 (Threonine, Tyrosine)<br>H4 (Serine, Tyrosine) | Chromatin compaction, DNA repair, Transcription |
| Formylation | H1 (Lysine)<br>H2B (Lysine)<br>H3 (Lysine)<br>H4 (Lysine) | DNA repair |

| | | |
|---|---|---|
| Oxidation | H4 (Lysine) | |
| Crotonylation | H2A (Lysine)<br>H2B (Lysine) | Transcription |
| Hydroxylation | H1 (Lysine)<br>H2A (Tyrosine)<br>H2B (Lysine, Tyrosine)<br>H4 (Tyrosine) | |
| Ubiquitinylation | H2A (Lysine)<br>H2B (Lysine)<br>H3 (Lysine) | Transcription |
| Succinylation | H2A (Lysine)<br>H2B (Lysine)<br>H3 (Lysine)<br>H4 (Lysine) | |
| Citrullination | H1 (Arginine) | |
| Propionylation | H4 (Lysine) | |

# II. Mechanisms of eukaryotic RNA Pol II transcription initiation

1. Core promoter

Eukaryotic genes consist in a promoter region depleted of nucleosomes, and so-called Nucleosome Free Region (NFR), where transcription starts, defined as the Transcription Start Site (TSS), and a gene body that contains the open reading frame. The core promoter is defined as the minimal DNA region bound by the transcription machinery sufficient for basal transcription [reviewed in (Kadonaga 2012)]. It contains sequences (**figure 4**) that are bound by specific proteins that recruit RNA Pol II to the promoter.

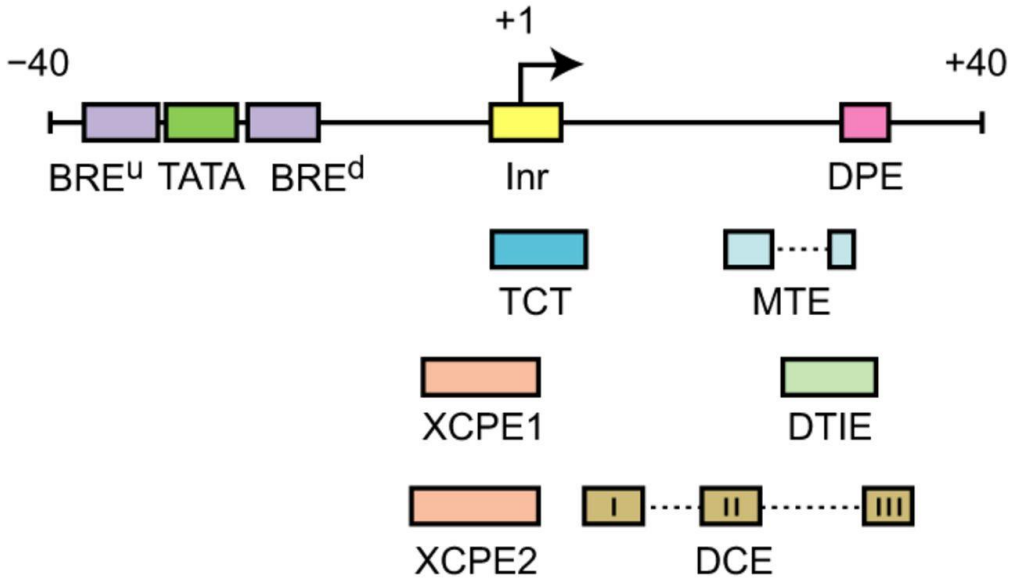



**Figure 4: RNA Pol II core promoter elements diversity.** The core promoter can display numerous elements, alone or in combination (Vo ngoc et al. 2017).

The TATA-box was the first core promoter element to be identified (Gannon et al. 1979) and is located 25-35 base pairs before the TSS but is not present in all metazoan genes (Jin et al. 2006; Kimura et al. 2006; Yang et al. 2007). Actually, only a minority of genes, about 20% in yeast, contain a TATA-box (Basehoar et al. 2004). So, higher eukaryotes genes harbor a combination of several different sequence elements such as for example the Initiator (Inr), the TFIIB Recognition Element (BRE), the Downstream Promoter Element (DPE), the Motif Ten Element (MTE) (**figure 4**) and others [reviewed in (Kadonaga 2012; Vo ngoc et al. 2017)].

## 2. Pre-Initiation Complex assembly at the promoter

### a. RNA Polymerase II : a multi-subunit complex

RNA Pol II enzyme is responsible for directing the synthesis of mRNA and has been shown to be conserved from yeast to human [reviewed in (Young 1991)]. This enzyme has a mass >0.5 MDa and chromatographic analyses revealed that it consists in a multi-subunit complex composed of 12 polypeptides (Bartholomew et al. 1986; Edwards et al. 1991) [reviewed in (Young 1991)]. The subunits RPB5, 6, 8 and 10 are shared by all three RNA Polymerases while RNA Pol II is characterized by the presence of RPB4, 7, 9 and RPB1 which harbors the Carboxyl Terminal Domain (CTD) (Carles et al. 1991; Hampsey 1998; Wild et al.

2012; Woychik et al. 1990). The resolution of its structure revealed four structural domains: the core, the clamp, the shelf, and the jaw lobe (Cramer et al. 2001).

The CTD of RBP1 consists in tandemly repeated heptapeptides and contains the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS). The number of repeated hexapeptides depends on the species and range from 26-27 times in Saccharomyces up to 52 times in mouse [reviewed in (Young 1991)]. The CTD can be post-translationally modified by phosphorylation and also glycosylation (Kelly et al. 1993) [reviewed in (Young 1991)]. The CTD is under an unphosphorylated state when RNA Pol II is recruited to the Pre Initiation Complex (PIC), while serine-5 phosphorylation is associated with transcription initiation and serine-2 phosphorylation with elongation (Bartholomew et al. 1986; Cadena et al. 1987; Chesnut et al. 1992; Laybourn et al. 1990).

b. The General Transcription Factors

Transcription initiation is an orchestrated process that requires the assembly of the Pre-Initiation Complex (PIC) that recruits the RNA Pol II enzyme to initiate mRNA synthesis at the promoter in NFRs delimited by an upstream -1 and a downstream +1 nucleosome (Jiang et al. 2009). *In vitro* transcription within cellular-free systems showed that RNA Pol II is not sufficient to direct accurate transcription from a DNA template, and that additional factors are required: the General Transcription Factors (GTFs) (Luse et al. 1980; Weil et al. 1979) [reviewed in (Thomas et al. 2006)]. Those factors have been identified from chromatography analyses of crude cell extracts and used for incubation with purified RNA Pol II. Initially, four nuclear factors have been identified for accurate RNA Pol II transcription initiation from four enzymatically active fractions (A, B, C and D) followed by the characterization of additional GTF from the sub-fractionation of the C active fraction (Samuels et al. 1982; Weil et al. 1979). Those RNA Pol II GTFs were named according their fraction elution, as Transcription Factor II (TFII). They include TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH (Flores et al. 1989, 1992; Ge et al. 1996; Sawadogo et al. 1985). The GTFs are evolutionary conserved large multi-subunit complexes, except TFIIB which is a single polypeptide chain.

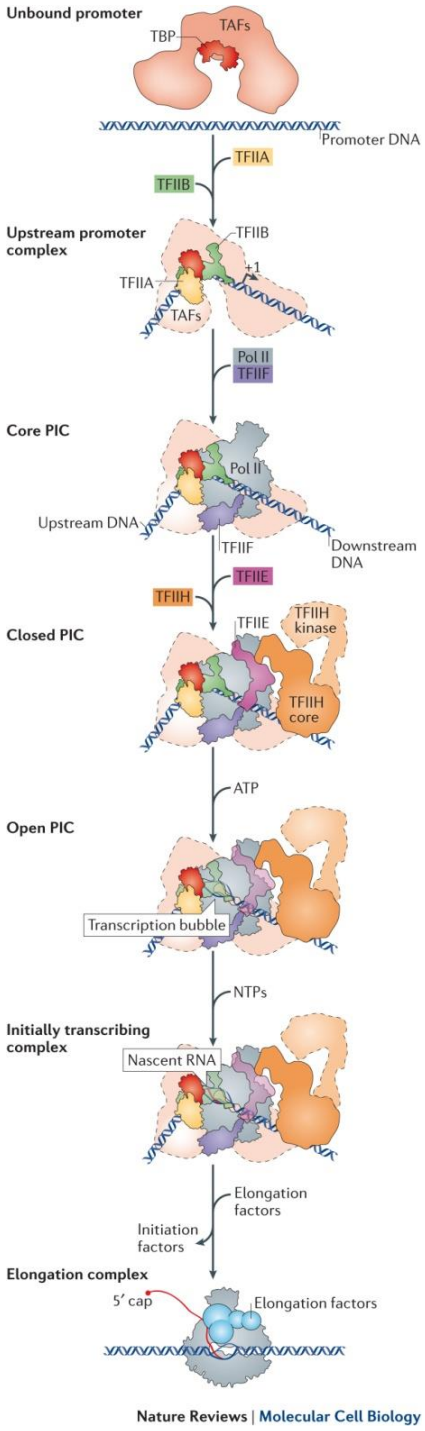



**Figure 5: Schematic representation of the preinitiation complex multi-step assembly model at the promoter for RNA Pol II recruitment.** According to the canonical model of PIC assembly, TFIID binds DNA first, notably through TBP which binds the TATA-box and bend DNA. TBP-DNA is stabilized by TFIIB and TFIIA. RNA Pol II is brought to this complex by TFIIF followed by TFIIE and TFIIH, which will melt DNA, initiating the transcription bubble, before productive elongation by RNA Pol II (Sainsbury, Bernecky, and Cramer 2015).

In the canonical model of PIC assembly (**figure 5**) described by Steve Buratowski and colleagues (Buratowski et al. 1989), the GTFs are assembled to the promoter hierarchically. TFIID comes first and recognizes the core promoter sequences. TBP binds the TATA-box, and distorts DNA (Starr et al. 1995), and other TFIID subunits, can also recognize additional elements. TFIID is stabilized on the promoter with TFIIA and TFIIB. RNA Pol II is brought to the PIC with TFIIF. The PIC assembly is finalized with the recruitment of TFIIE and TFIIH. TFIIH is responsible for the transition from the PIC to the open complex through DNA melting.

Nevertheless, an alternative model based from the observation that RNA Pol II purified together with a set of GTFs, TFIID and TFIIA excepted, and other co-activators led to the holoenzyme model [reviewed in (Thomas et al. 2006)]. According to this model, TFIID and TFIIA come first followed by the binding of a pre-assembled RNA Pol II complex with GTFs, remodelers and co-activators. Evidences exist for both assembly pathways but they are still debated in the community [reviewed in (Thomas et al. 2006)].

- **TFIID**

TFIID is the first GTF to be recruited to the promoter and nucleates the PIC assembly. TFIID contains TBP and 13, in metazoan, or 14, in yeast, TBP-associated factors (TAFs). TBP binds TATA-box element but other TAFs are also capable of binding additional core promoter elements. Since TFIID is one of the topic of the thesis, it will be described in more details thereafter.

- **TFIIA**

TFIIA is essential in yeast, where it is composed of two polypeptides encoded by *TOA1* and *TOA2* (Ranish et al. 1991, 1992) whereas in metazoan it is composed of three polypeptides: the α, β and γ subunits (DeJong et al. 1993; Yokomori et al. 1993a). However, the α and β proteins are encoded by a single gene and are cleaved post-translationally or not depending on the cell-type [reviewed in (Høiby et al. 2007)]. TFIIA has been identified as a component of the basal transcription machinery, and was purified as an interacting partner of TFIID (DeJong et al. 1993; Ranish et al. 1991, 1992; Reinberg et al. 1987; Yokomori et al. 1993a). So, it has been proposed that TFIIA stabilizes TBP binding to DNA (Yokomori et al. 1994; Weideman et al. 1997) and controls TBP/ TFIID dimerization, thus accelerating the binding of TBP to DNA (Coleman et al. 1999). The stimulatory effect of TFIIA on transcription comes from its anti-

repressor activity of inhibitors including Mot1/TAF-172, NC2/Dr1, topoisomerase I, and TAF1 (Auble et al. 1993; Chicca et al. 1998; Inostroza et al. 1992; Kokubo et al. 1998; Merino et al. 1993). It has been shown that TFIIA acts also as a coactivator by physically interacting with several factors (Kobayashi et al. 1998; Kraemer et al. 2001; Ozer et al. 1994; Yokomori et al. 1993a) and is thus required for activation of several genes (Kobayashi et al. 1995; Lieberman 1994; Lieberman et al. 1997; Stargell et al. 1995, 2000).

- **TFIIB**

TFIIB is a single polypeptide (Ha et al. 1991; Maldonado et al. 1990; Malik et al. 1993) that was shown to bind the TFIIA-TFIID complex (Maldonado et al. 1990). TFIIB contains three domains: the amino-terminal zinc ribbon domain (the B-ribbon) that contacts the RNA Pol II, the finger domain inserted in the RNA Pol II active center and the carboxy-terminal domain (the B-core, comprising two cyclin folds) interacting with RNA Pol II and TBP (Barberis et al. 1993; Buratowski et al. 1993; Malik et al. 1993). TFIIB can also recognize upstream and downstream elements of the TATA-box of the *AdE4* promoter: the $BRE^{u}$ and $BRE^{d}$ elements (Lagrange et al. 1998; Qureshi et al. 1998).

- **TFIIF**

TFIIF was initially found as an RNA Pol II interacting partner (Sopta et al. 1985) and is formed by the hetero-dimerisation of its two subunits RAP30 and RAP74 proteins (Burton et al. 1988). *In vitro* studies showed that transcription initiation can occur to a certain extent without TFIIE and TFIIH but not without TFIIF, illustrating the critical role of TFIIF (Pan et al. 1994). TFIIF plays a role in PIC formation at several levels. TFIIF was shown to facilitate RNA Pol II recruitment to the TFIIB and D complex and stabilizing the PIC (Flores et al. 1991). It has been also described that TFIIF can induce a topological conformation that stabilizes a TBP-TFIIB-pol II-TFIIF- promoter DNA complex (Hou et al. 2000). Moreover, by interacting directly with TFIIE (Maxon et al. 1994) it mediates the recruitment of both TFIIE and TFIIH [reviewed in (Orphanides et al. 1996)]. In yeast, TFIIF has also been shown to control the start site selection (Ghazy et al. 2004), for which TFIIB is also involved (Fairley et al. 2002). Not only TFIID plays multiple roles in transcription initiation, but it also facilitates and enhances the transition between initiation and elongation (Cheng et al. 2007a; Cojocaru et al. 2008; Renner et al. 2001; Schweikhard et al. 2014; Újvári et al. 2011).

- **TFIIE**

TFIIE is recruited with TFIIH to the PIC and is responsible together with TFIIH for promoter melting (Holstege et al. 1996). It is a hetero tetramer composed of two subunits, α and β (Ohkuma et al. 1994; Peterson et al. 1991; Sumimoto et al. 1991). TFIIE binds to TFIIH, TFIIB, promoter DNA, RNA Pol II and help recruiting TFIIH (Flores et al. 1989; Forget et al. 2004; Maxon et al. 1994; Watanabe et al. 2003). Furthermore, TFIIE has been shown to stimulate the ATPase, CTD kinase and DNA helicase activities of TFIIH which helps the formation of an initiation-competent Pol II complex (Ohkuma et al. 1994, 1995; Serizawa et al. 1994)

- **TFIIH**

The last factor to be recruited to the PIC TFIIH has been discovered as an indispensable factor for transcription initiation *in vitro,* and was purified from rat liver and HeLa cells and was originally called general transcription factor-δ or BTF2 (Conaway et al. 1989; Gerard et al. 1991). Interestingly, TFIIH has been also shown to be required for RNA Pol I transcription [reviewed in (Compe et al. 2016)]. TFIIH is a multi-subunit complex that consists of 10 subunits [reviewed in (Compe et al. 2016)]. The complex is organized into two sub-complexes: the core complex and the cyclin-dependent kinase (CDK)-activating kinase (CAK) complex. The core complex contains xeroderma pigmentosum group B complementing protein (XPB), p62, p52, p44, p34 and p8 and the CAK complex contains CDK7, cyclin H and MAT1 [reviewed in (Compe et al. 2016)]. TFIIH displays three ATP-dependent activities with the subunits XPD (catalyzing a 3' --> 5' DNA helicase activity), XPB (catalyzing a 5' --> 3' DNA helicase activity) and CDK7 (catalyzing a kinase activity). From the reconstitution of TFIIH, the role of the XPB DNA helicase was elucidated (Tirode et al. 1999). During transcription initiation, XPB catalyzes the formation of the open complex in a ATP-dependent manner before the synthesis of the first phosphodiester bond of nascent transcripts (Tirode et al. 1999). Recently, it was shown that only the ATPase activity of XPB is required for transcription initiation suggesting that no helicase activity is required for transcription initiation (Alekseev et al. 2017). Moreover, TFIIH controls the elongation efficiency by preventing the premature arrest of RNA Pol II activity at promoter-proximal sites (Dvir et al. 1997; Moreland et al. 1999; Yan et al. 1999). The CDK7 subunit phosphorylates the RNA Pol II CTD at serine 5 and 7, and thus plays a role in coupling transcription and RNA processing [reviewed in (Compe et al. 2016)]. In addition,

CDK7 phosphorylates other substrates such as nuclear receptors [reviewed in (Compe et al. 2016)]. TFIIH is also a key player in DNA repair within the nucleotide excision repair pathway [reviewed in (Compe et al. 2012)].

### c. Structural aspects of the PIC assembly model

Recent structural analyses brought new insights into the assembly of the PIC. The assembly of the PIC results in a closed promoter complex (CC) where DNA is loaded into RNA Pol II (**figure 5**). Through DNA melting mediated by TFIIH, the CC transitions into an open promoter complex (OC) and is finally converted into an initial transcribing complex (ITC) (Cheung et al. 2011). So, in presence of nucleoside triphosphates, RNA Pol II initiates transcription with the first phosphodiester bond of RNA (Cheung et al. 2011; Liu et al. 2011). After several abortive initiation cycles where short RNAs are released, the ITC is finally transformed into an elongation complex (EC) with promoter escape of RNA Pol II. The structural aspects of transcription initiation have been well described in yeast and human [reviewed in (He et al. 2016; Louder et al. 2016; Plaschka et al. 2016; Schilbach et al. 2017)].

## 3. Promoter proximal pausing

Elongating RNA Pol II has been shown to enter a transient promoter proximal pausing mode 30-60nt downstream of the TSS. This pausing phenomenon has been mainly described in *Drosophila,* initially for heat shock coding genes (Rougvie et al. 1990) and is proposed as a regulatory mechanism that can establish a permissive chromatin, as well as integrating external stimuli and coordinate gene synchronicity [reviewed in (Adelman et al. 2012; Mayer et al. 2017)]. For instance, in *Drosophila,* RNA Pol II pausing was shown to control the timing of activation of the gene *snail*, which controls the coordinated invagination of mesodermal cells during gastrulation (Lagha et al. 2013). In mammals, accumulation of engaged polymerases at the 5' region associated with transcript fragments was first described in mature erythrocytes, that are not supposed to transcribe, and suggested already that elongation could be a limiting step for transcription (Gariglio et al. 1981). In line with that, paused RNA Pol II has been observed at promoters of *c-fos* or *c-myc* (Plet et al. 1995; Strobl et al. 1992), and according recent genome-wide analyses, it is widespread in mammals with up to 30% of human genes which are concerned (Core et al. 2008; Guenther et al. 2007). Paused RNA Pol II are enriched

for serine 5 phosphorylation at the CTD and the DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF) have been shown to block transcription elongation (Aiyar et al. 2004; Luecke et al. 2005; Yamaguchi et al. 2013). In addition, the core promoter composition can also play a role in the recruitment of the pausing factors, such as the GAGA motif, the DPE, the 'pause button', and the TATA box (Amir-Zilberstein et al. 2007; Chen et al. 2013; Gaertner et al. 2012; Hendrix et al. 2008). RNA Pol II resumes transcription after recruitment of P-TEFb which phosphorylates the serine 2 of the RNA Pol II CTD, DSIF and NELF. Phosphorylated NELF dissociates from the pausing site and RNA Pol II enters productive elongation with DSIF [reviewed in (Liu et al. 2015)].

## 4. Transcription reinitiation

The transition between initiation to elongation has been shown to disrupt the PIC, with only TFIID remaining immobilized on the *in vitro* template, where some of them of the GTF were recycled into the EC (TBP, TFIIB, TFIIF, TFIIE, and TFIIH were detected together) (Zawel et al. 1995). While transcription initiation requires a full PIC assembly at the promoter, reinitiation can occur only from a subset of factors (Yudkovsky et al. 2000). In this study, Steve Hahn and colleagues, described the existence of a complex made of the remaining TFIID, TFIIA, TFIIH, TFIIE and Mediator. This scaffold intermediate could reinitiate transcription upon addition of ATP and dependent on TFIIH (Yudkovsky et al. 2000). Recently, the laboratory of Steve Buratowski showed that TAFs were also capable of binding 30 bp downstream from the TATA-box, and were able to direct transcription reinitiation *in vitro*, independently of an activator (Joo et al. 2017). So, these two studies provide some clues for reinitiation mechanisms that differ but remain largely not well understood.

## 5. Control of transcription activation

### a. Distal cis-regulatory elements

Distal elements, such as upstream activator sequence (UAS) in yeast and enhancers in metazoans, are found upstream or downstream of the promoter, or even in the intron, from hundreds of bases and up to many mega bases control gene activation [reviewed in (Blackwood et al. 1998)]. They are bound by transcription factors (detailed thereafter) and thus activate gene

expression. Communication between enhancers and promoters is thought to be achieved via DNA looping [reviewed in (Levine et al. 2014)]. Function of enhancer elements has been well studied in a developmental context, where they control the spatio-temporal gene expression. For instance, in *Drosophila* the expression of the segmentation gene *even-skipped* in seven stripes results from the activity of five distinct enhancers [reviewed in (Levine et al. 2014)]. Other elements such as silencers, which have a repressing effect on gene expression, and insulators, which block activating or repressing effect from enhancer and silencer elements, also exist in combination with enhancers to modulate gene expression [reviewed in (Narlikar et al. 2009)].

b. Transcription factors

Gene activation or repression depends on the binding of proteins to DNA distinct to the GTFs, which are named transcription factors (TFs). They usually bind to 6–12 bp long degenerate DNA sequences located in enhancers [reviewed in (Spitz et al. 2012)]. In the context of development and cellular differentiation, certain TFs, called pioneer TFs, have the ability to bind inaccessible regions of DNA, leading to the recruitment of histone-modifying proteins and nucleosome remodeling [reviewed in (Iwafuchi-Doi et al. 2016)].

c. Co-activators

i. Mediator

Transcriptional activation requires RNA Pol II and the GTFs which are necessary and sufficient for directing accurate transcription *in vitro*. It has been proposed that Mediator, a large multi-subunit complex with modular organization, is as important as GTFs for transcription genome wide in both yeast and metazoan (Holstege et al. 1998; Ito et al. 2000, 2002; Soutourina et al. 2011; Tudor et al. 1999; Westerling et al. 2007). The Mediator complex was initially purified in yeast (Kim et al. 1994) and was described in metazoan under several acronyms: ARC (activator recruited cofactor), DRIP (VDR-interacting proteins), SMCC, SREBP, TRAP (thyroid hormone receptor associated proteins) or PC2 [reviewed in (Boyer et al. 1999; Fondell et al. 1996; Jiang et al. 1998; Rachez et al. 1999; Malik et al. 2005; Näär et al. 1999; Ryu et al. 1999)]. This data showed that Mediator complex is conserved from yeast to

human, comprising 25 subunits in budding yeast and up to 30 subunits in humans, and a unified nomenclature was established (Boube et al. 2002; Bourbon et al. 2004).

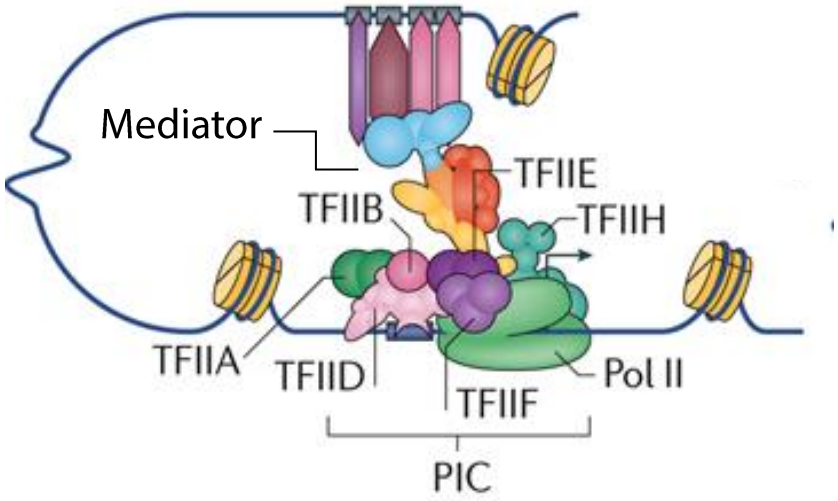



**Figure 6: Schematic representation of the role of Mediator in transcription initiation.** Mediator is recruited to enhancers where it interacts with transcription factors, and mediate their action by contacting the basal transcription machinery at the promoter, adapted from (Soutourina 2017).

Mediator is recruited to enhancers to mediate contacts between the transcription and the basal transcription machinery (**figure 6**) [reviewed in (Soutourina 2017)]. It was shown that Mediator contacts directly the RNA Pol II CTD and induces the phosphorylation of the CTD by TFIIH that allows the release of RNA Pol II for elongation (Myers et al. 1998; Thompson et al. 1995). Mediator thus integrates both positive and negative signals for transcriptional regulation.

### ii. Chromatin remodeling complexes

- ATP-dependent complexes

Since DNA is wrapped around a complex of histones, it is inaccessible to transcription factors, with the exception of pioneer factors mentioned before. Thus, remodeling complexes are necessary to expose DNA regulatory sequences to transcriptional activators by regulating nucleosome positioning and spacing [reviewed in (Vignali et al. 2000)]. All remodeling complexes contain subunits with an ATPase activity that allow them to modify the nucleosome structure. The main ATP-remodeling complexes are the imitation switch (ISWI),

chromodomain helicase DNA-binding (CHD), switch/sucrose non-fermentable (SWI/SNF) and INO80 [reviewed in (Vignali et al. 2000)].

- Histone modifiers

Another group of co-activators are able to add several post-translational modifications on aminoacid sequences of histone N-terminal tails, that influence the chromatin state (non-permissive *versus* permissive) and participate to the recruitment of other transcriptional factors. Among those coactivators, the Spt-Ada-Gcn5-Acetyl transferase (SAGA) complex acetylates histones H3 and deubiquitylates histones H2B, and will be detailed thereafter.

# III. TFIID, a General Transcription Factor

## 1. Discovery of TBP

TFIID was identified from HeLa nuclear extracts as an essential component for RNA Pol II transcription from both cellular and viral promoter templates (Flores et al. 1990). However, due to TFIID hetereogenity in mammals making its purification challenging, it is in yeast that the first characterization of TFIID was performed (Buratowski et al. 1988; Cavallini et al. 1988; Hernandez 1993). Steve Buratowski *et al* and Bruno Cavallini *et al* identified in yeast a factor capable to substitute for the mammalian TFIID in a reconstituted mammalian RNA Pol II transcription system with a TATA-box binding property (Buratowski et al. 1988; Cavallini et al. 1988). Isolation and purification of this factor in yeast revealed a protein of 25-28kDa (Cavallini et al. 1988; Hahn et al. 1989; Horikoshi et al. 1989; Schmidt et al. 1989). So, TBP was the first TFIID subunit to be characterized and was cloned from yeast (Cavallini et al. 1988; Hahn et al. 1989; Horikoshi et al. 1989; Schmidt et al. 1989), paving the way for the cloning and characterization of TBP homologs in *Drosophila* (Hoey et al. 1990; Muhich et al. 1990) and in human (Hoffman et al. 1990; Kao et al. 1990; Peterson et al. 1990a). Moreover, TBP-like proteins have also been found in several organisms, but will be detailed thereafter [reviewed in (Müller et al. 2010)].

## 2. TBP-Associated-Factors form a large TFIID multi-subunit complex

Experiments from cell-free reconstituted transcription systems showed that TBP is necessary and sufficient for basal transcription level *in vitro*, but not for transcription factor stimulated transcription (Dynlacht et al. 1991; Peterson et al. 1990b; Pugh et al. 1990; Kao et al. 1990; Reese et al. 1994). Reconstituted transcription system containing *Drosophila* basal factors and TFIID, but not TBP, were for instance necessary for activated transcription of a developmentally regulated transcription factor (Dynlacht et al. 1991). In this study, TBP sedimented within a complex of 350 kDa, a molecular weight much larger than the single TBP. TBP immunoprecipitation from the TFIID fraction further revealed the association of TBP with several polypeptides, named TBP-associated factors (TAFs). Similar experiments in human cells also showed that TBP was associated with TAFs (Tanese et al. 1991). Then, many additional TAFs have been cloned from yeast, *Drosophila* and human cell lines (Chiang et al. 1995; Goodrich et al. 1993; Hisatake et al. 1995; Hoffmann et al. 1996; Hoey et al. 1993; Jacq et al. 1994; Klemm et al. 1995; Kokubo et al. 1994; Mengus et al. 1995, 1997; Moqtaderi et al. 1996b; Poon et al. 1995; Ruppert et al. 1993; Tanese et al. 1996; Verrijzer et al. 1994; Weinzierl et al. 1993; Yokomori et al. 1993b). So, TFIID forms a large multi-subunit complex, comprised of TBP and 13 TAFs in metazoan or 14 TAFs in yeast. TAFs have been ranged according their molecular weight in human cells and are now designated as TAF proteins according to a unified nomenclature (**Table 2**) (Tora 2002). In addition of those TAFs, many paralogs have also been identified in several organisms, but will be detailed later, [reviewed in (Müller et al. 2010)].

**Table 2: Unified nomenclature for the TFIID subunits including orthologs and paralogs.**

Nomenclature for the TFIID subunits as published by (Tora 2002).

| **New name** | ***H. sapiens*** | ***D. melanogaster*** | ***Caenorhabditis elegans*** | | ***S.cerevisiae*** | ***S. pombe*** |
|---|---|---|---|---|---|---|
| | | | Previous name | New name | | |
| **TAF1** | $TAF_{II}250$ | $TAF_{II}230$ | taf-1 (W04A8.7) | *taf-1* | Taf145/130 | $TAF_{II}11$ |
| **TAF2** | $TAF_{II}150$ | $TAF_{II}150$ | taf-2 (Y37F11B.4 | *taf-2* | Taf150 or TSM1 | (T38673) |
| **TAF3** | $TAF_{II}140$ | $TAF_{II}155$ or BIP2 | (C11G6.1) | *taf-3* | Taf147 | |
| **TAF4** | $TAF_{II}130/135$ | $TAF_{II}110$ | taf-5 (R119.6) | *taf-4* | Taf48 or MPT1 | (T50183) |
| **TAF4b** | $TAF_{II}105$ | | | | | |
| **TAF5** | $TAF_{II}100$ | $TAF_{II}80$ | taf-4 (F30F8.8) | *taf-5* | Taf90 | $TAF_{II}72$ |
| **TAF5b** | | | | | | $TAF_{II}73$ |
| **TAF5L** | PAF65β | Cannonball | | | | |
| **TAF6** | $TAF_{II}80$ | $TAF_{II}60$ | taf-3.1 (W09B6.2) | *taf-6.1* | Taf60 | (CAA20756) |
| **TAF6L** | PAF65α | (AAF52013) | taf-3.2 (Y37E11AL.8) | *taf-6.2* | | |
| **TAF7** | $TAF_{II}55$ | (AAF54162) | taf-8.1 (F54F7.1) | *taf-7.1* | Taf67 | $TAF_{II}62$/PTR6 |
| **TAF7L** | TAF2Q | | taf-8.2 (Y111B2A.16)<br>(ZK1320.12) | *taf-7.2* | | |
| **TAF8** | (BAB71460) | Prodos | (ZK1320.12) | *taf-8* | Taf65 | (T40895) |
| **TAF9** | $TAF_{II}32/31$ | $TAF_{II}40$ | taf-10 (T12D8.7) | *taf-9* | Taf17 | (S62536) |
| **TAF9L** | $TAF_{II}31L$ (AAG09711) | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **TAF10** | $TAF_{II}30$ | $TAF_{II}24$ | taf-11 (K03B4.3) | *taf-10* | Taf25 | (T39928) |
| **TAF10b** | | $TAF_{II}16$ | | | | |
| **TAF11** | $TAF_{II}28$ | $TAF_{II}30\beta$ | taf-7.1 (F48D6.1) | *taf-11* | Taf40 | (CAA93543) |
| **TAF11L** | | | taf-7.2 (K10D3.3) | *taf-11.1* | | |
| **TAF12** | $TAF_{II}20/15$ | $TAF_{II}30\alpha$ | taf-9 (Y56A4.3) | *taf-12* | Taf61/68 | (T37702) |
| **TAF13** | $TAF_{II}18$ | (AAF53875) | taf-6 (C14A4.10) | *taf-13* | Taf19 or FUN81 | (CAA19300) |
| **TAF14** | | | | | Taf30 | |
| **TAF15** | $TAF_{II}68$ | | | | | |
| | | | | | | |
| **BTAF1** | $TAF_{II}170$/ TAF-172 | Hel89B | (F15D4.1) | *btaf-1* | Mot1 | (T40642) |

3. Distinct TFIID complexes exist

TBP was found in two different TFIID complexes with two different molecular weights, one of 300 kDa and one of >700 kDa, from phosphocellulose chromatographic analyses on HeLa nuclear extracts (Timmers et al. 1991). Most of TBP was actually found in the 300 kDa complex (called B-TFIID) (Timmers et al. 1991). In this complex, TBP was associated with a 170 kDa protein, called BTAF1 (Timmers et al. 1992). BTAF1 belongs to the family of SNF2-like ATPases (Chicca et al. 1998; van der Knaap et al. 1997), and also exists in *S. cerevisiae* as Mot1p that associates with TBP as well (Poon et al. 1994), and under 89B helicase in *Drosophila* (Goldman-Levi et al. 1996). However, the B-TFIID complex did not mediate SP1 or GAL4-AH transcriptional response and only weakly for the major-late transcription factor (Timmers et al. 1991). Note also that TBP was not detected free in HeLa cells (Timmers et al. 1991), and we also know now that TBP is also part of the SL1 and TFIIIB complexes for RNA Pol I and III transcription machineries respectively [reviewed in (Thomas et al. 2006)].

While looking for the factors responsible of mediating the transcription response of transcriptional co-activators, two distinct TFIID complexes from HeLa cell extracts were detected (Brou et al. 1993). Transcription stimulation of the three chimeric recombinant activators GAL-TEF-1, GAL-ER(EF) and GAL-VP16 activators required the fraction obtained from phosphocellulose chromatography (PC) with 0.5-1.0M KCl elution, while the PC0.3 (0.1-0.3M KCl elution) fraction was only able to mediate the response of GAL-TEF-1. TBP immunoprecipitation revealed that TFIID was present, but with a different TAFs composition between the two fractions (Brou et al. 1993).

Furthermore, in the PC1.0 fraction, two TFIID sub-populations were detected (Jacq et al. 1994). Double immunoprecipitation with first an antibody against TBP pulled down all the distinct TFIID complexes previously described (Brou et al. 1993), and the TAF10 immunoprecipitation, against the TBP-immunoprecipitation elution, pulled down a TAF10-containing TFIID separated from another TFIID complex that remained in the supernatant (Jacq et al. 1994). In both complexes, TBP together with TAF1, TAF5, TAF6, TAF7 and TAF11 were present. TAFII125 (probably TAF4b), TAFII37 (probably TAF8), TAF12 and TAF13 were detected in TAF10-containing TFIID while almost not found in TAF10-containing TFIID (Jacq et al. 1994). Those studies have been limited so far by the antibodies available, so the composition of these different complexes, or sub-complexes, might not be complete.

Last but not least, TAFs were also found in multi-subunit complexes without TBP associated with other proteins such as the TBP-free $TAF_{II}$-containing complex (TFTC) (Wieczorek et al. 1998), SAGA, STAGA and SLIK complexes that will be detailed thereafter (Grant et al. 1998; Martinez et al. 2001).

## 4. Architecture of TFIID and structural model for the assembly

TAFs contain several conserved structural domains (**figure 7**). The histone fold domain is one such conserved domain. Several TAF-TAF interactions have been identified such as TAF3-10, TAF4-12, TAF6-9, TAF8-10 and TAF11-13 (Birck et al. 1998; Gangloff et al. 2000; Wertent et al. 2002; Xie et al. 1996). Since TAF10 does not contain a Nuclear Localizing Sequence (NLS) it cannot enter the nucleus alone (Soutoglou et al. 2005). So, TAF10 requires TAF8 or TAF3 that have been identified as interacting partners *in vitro* (Soutoglou et al. 2005). Although TAF3 is not described in the current model of assembly of TFIID for the

incorporation of TAF10 (Bieniossek et al. 2013), it has been observed that in the absence of TAF8, TAF10 could still be incorporated into a TFIID-like complex in fibroblasts (El-Saafin et al. 2018).

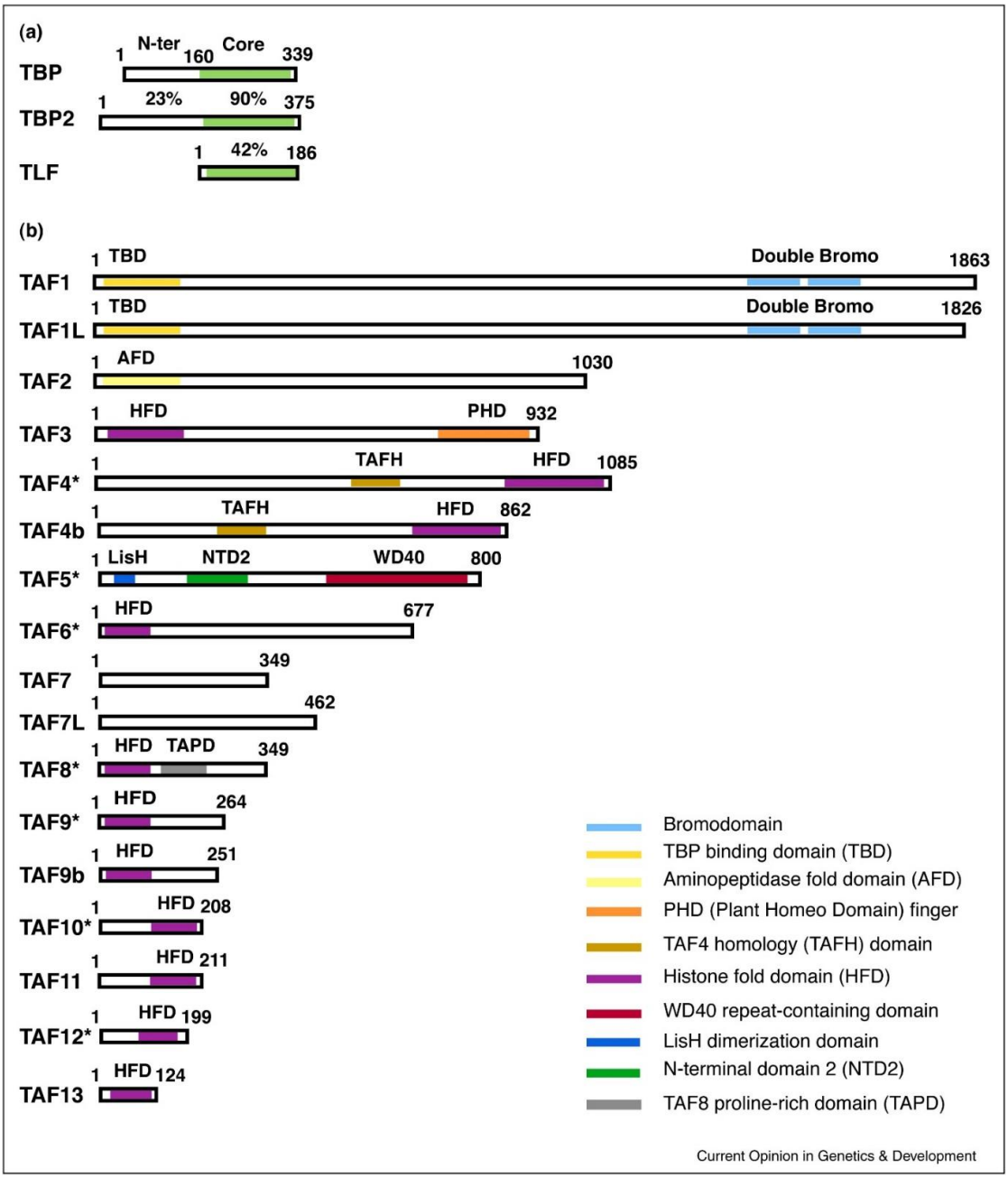



**Figure 7: Schematic representation of the TFIID subunits.** (a) TBP and TBP-like factors organization and comparison of the identity between the N-terminal (N-ter) and core DNA binding domains (%). (b) Human TAFs organization (except TAF5L and TAF6L). The numbers represent the amino acid position (Müller et al. 2010).

Analyses of TFIID in yeast revealed that several TAFs are present in two copies and provided the first stoichiometry of the complex (Sanders et al. 2002). More recently, the TFIID complex assembly has been described as a stepwise model leading to three sub-complexes: the core complex (containing 5 TAFs), the 8TAF complex and the holo-TFIID based on single-particle cryo-electron microscopy analyses of the recombinant human TFIID complex (**figure 8**) (Bieniossek et al. 2013; Trowitzsch et al. 2015b).

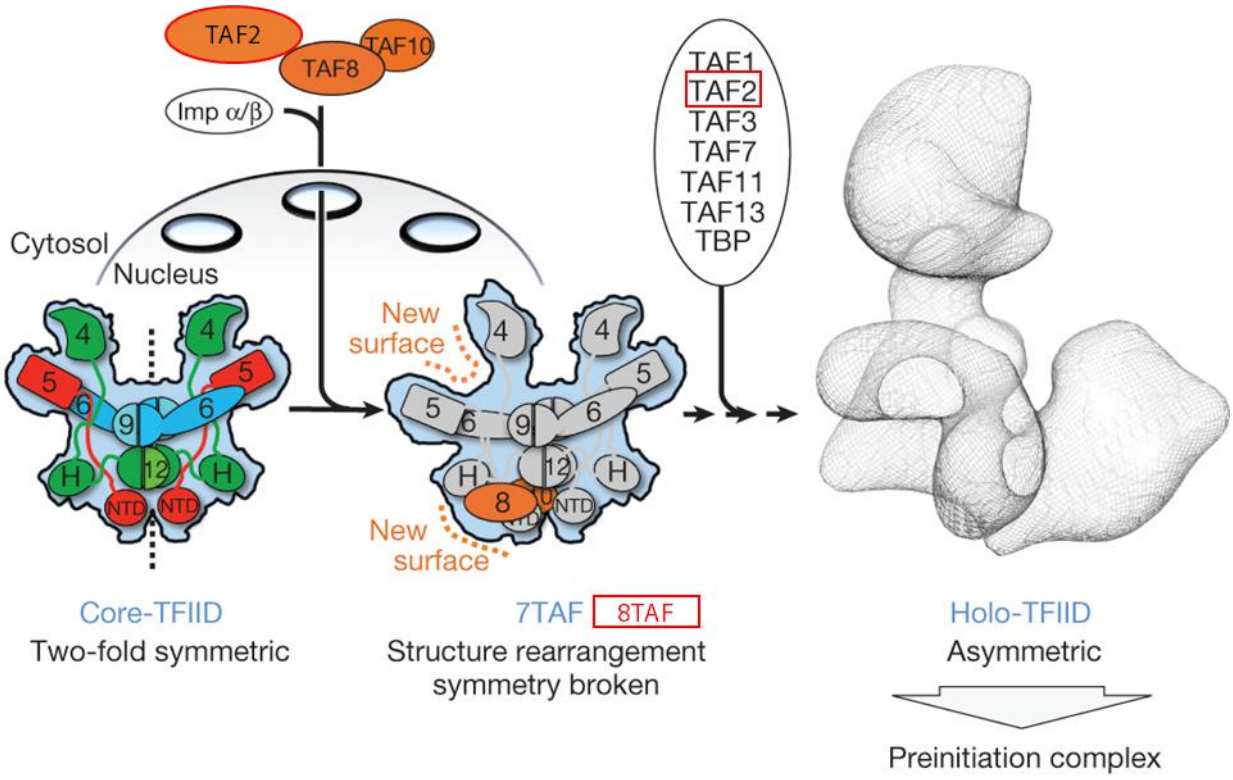



**Figure 8: Schematic representation of the TFIID model of assembly.** The core TFIID is composed of two copies of TAF4, TAF5, TAF6, TAF9 and TAF12. The symmetry of the core TFIID is broken by the import of TAF2-TAF8-TAF10. TAF8 interact with TAF10 via the HFD and with TAF2 via its C-terminal domain. The 8TAF complex is completed with the addition of single copies of TAF1, TAF3, TAF7, TAF11, TAF13 and TBP**.** Adapted from (Bieniossek et al. 2013) and updated based on (Trowitzsch et al. 2015), modifications are shown in red.

As proposed in *Drosophila* (Wright et al. 2006), the core complex is composed of two copies of TAF4, TAF5, TAF6, TAF9 and TAF12. TAF10 forms a prebuilding block in the cytoplasm of human cells with TAF2 and TAF8. This complex associates with importin α, through the TAF2 NLS, to translocate into the nucleus. These three TAFs break the symmetry and leads to a structural rearrangement resulting in the 8TAF complex which provides a platform for final incorporation of single copies of TAF1, TAF3, TAF7, TAF11 and TBP.

Structural analyses by cryo-electron microscopy and single-particle image analysis revealed that the human TFIID has a horseshoe shape, containing three lobes A, B and C that surround a central cavity (Andel et al. 1999; Brand et al. 1999a; Grob et al. 2006). However, TFIID displays a structural flexibility because the complex switches from a canonical to rearranged state that interacts with promoter DNA in a TFIIA-dependent manner (Louder et al. 2016). DNAse I footprinting analyses revealed that TFIID occupies a promoter region spanning from -40 to +35bp (Zhou et al. 1992). Since TBP binds the TATA-box promoter region and that the holo-TFIID was also required for transcription from TATA-less promoters (Pugh et al. 1991; Zhou et al. 1992), TAFs have been shown to also recognize additional core promoter elements. It has been suggested that TAF1 and TAF2 bind to the Inr (Verrijzer et al. 1994), TAF6 and TAF9 bind the DPE and TAF1 can bind the DCE and the MTE (Juven-Gershon et al. 2006; Lee et al. 2005a). TFIID is able to interact also with post-translationally modified nucleosomes. TAF3 can bind H3K4me3 (van Nuland et al. 2013; Vermeulen et al. 2007) and TAF1, that contains two bromodomains, recognize acetylated histones (Jacobson et al. 2000). Such interactions are thought to stabilize TFIID to the promoter.

### 5. TAFs are transcriptional co-activators

TAFs are considered as co-activators necessary for activating transcription. From *in vitro* analyses, several TAFs have been identified as interacting partners for TFs. TAF4 was the first TFIID subunit for which such interaction has been shown for Sp1 and CREB activators (Asahara et al. 2001; Gill et al. 1994; Hoey et al. 1993; Rojo-Niersbach et al. 1999; Saluja et al. 1998). Similarly, using transcriptional activation by nuclear receptors as a model, it was demonstrated that TAFs mediate selective transcription activation (May et al. 1996). For instance, TAF4 interacts with the activation function 2-containing region (AF2) of the receptors of retinoic acid, thyroid hormone, and vitamin D3 (Mengus et al. 1997). Another example relies

on TAF10 which interacts with the human estrogen receptor but not with VP16 (Jacq et al. 1994). Moreover, those *in vitro* data were supported by functional *in vivo* analyses. In yeast, while the role of TAFs in transcription activation was debated (Moqtaderi et al. 1996a), TAFs were shown to be required for the expression of a subset of genes (Walker et al. 1997; Wang et al. 1994) and also interact with several activation domains of TFs (Hall et al. 2002; Klein et al. 2003). As an example, the ribosomal protein gene expression requires the interaction between Rap1p and TFIID (Garbett et al. 2007).

In metazoan, TAFs were also shown to be clearly required for mediating transcription activation. For instance, in *Drosophila*, the transcription factors Dorsal and Prodos interact with certain TAFs to activate specific set of genes (Hernández-Hernández et al. 2001; Zhou et al. 1998). In mouse models, TAF10 has been shown to interact with GATA1 during erythropoiesis (Papadopoulos et al. 2015), TAF4-TAF12 heterodimer is important for HNF4A binding in liver (Alpern et al. 2014) and also for MYB in acute myeloid leukemia (Xu et al. 2018). Moreover, the TAF paralog TAF7L interacts with PPARγ to control spermiogenesis and lineage specification during adipogenesis (Zhou et al. 2013a, 2014). Interestingly, their conditional deletion phenocopies the deletion of their interacting co-activators. This data suggests that TFIID represents a selective interacting platform for specific transcription factors during development and differentiation.

## 6. Enzymatic activities of TAFs

In somewhat contradictory studies it has been suggested that the largest TFIID subunit, TAF1, displays HAT, kinase and ubiquitinylation activities. *In vitro*, TAF1 was shown to acetylate H3 and H4 as well GTFs such as TFIIE and TFIIF (Imhof et al. 1997; Mizzen et al. 1996). Similarly, TAF1 kinase activity also acts on GTFs, since TAF1 phosphorylates both serine and threonine residues of TFIIFα and TFIIA *in vitro* (Dikstein et al. 1996; O'Brien et al. 1998; Solow et al. 2001). HAT and kinase activities together have been shown to be required for cell cycle progression in a hamster cell line, and interestingly TAF1 HAT regulates only a specific subset of genes related to cell cycle (Dunphy et al. 2000; Hilton et al. 2005; Sekiguchi et al. 1996; Toshiro et al. 1994; Wang et al. 1994). TAF1 was also shown to mono-ubiquitinylate PAX3, which is involved in myogenic differentiation, and thus regulate PAX3 protein level (Boutet et al. 2010).

# IV. The SAGA co-activator complex, a TAF-containing complex

## 1. Identification and characterization of the SAGA complex and SAGA-related complexes in yeast

The SAGA complex has been initially characterized in yeast, as a GCN5-containing complex, associated to SPT and ADA proteins, so called Spt-Ada-Gcn5-Acetyl transferase (Grant et al. 1997), and was shown to be evolutionary conserved, though several differences that will be detailed thereafter, from yeast to human (**Table 3**) [reviewed in (Helmlinger et al. 2017)].

**Table 3: Composition of orthologous SAGA complexes.** Adapted from (Helmlinger *et al.*, 2002).

| | ***S.s cerevisiae*** | ***S. pombe*** | ***D.melanogaster*** | ***H.sapiens/M. musculus*** |
|---|---|---|---|---|
| **HAT module** | Gcn5 | Gcn5 | KAT2 (GCN5) | KAT2A |
| | | | | KAT2B |
| | Ada2 | Ada2 | Ada2b | TADA2b |
| | Ngg1 (Ada3) | Ngg1 (Ada3) | Ada3 | TADA3 |
| | Sgf29 | Sgf29 | Sgf29 | SGF29 |
| **DUB module** | Ubp8 | Ubp8 | dNonstop | USP22 (UBP22) |
| | Sgf11 | Sgf11 | dSgf11 | ATXN7L3 |
| | Sgf73 | Sgf73 | dATXN7 | ATXN7/ATXN7L1/L2 |
| | Sus1 | Sus1 | dE(y)2 | ENY2 |
| **Core structural module** | Taf5 | Taf5 | WDA | TAF5L |
| | Taf6 | Taf6 | SAF6 | TAF6L |
| | Taf9 | Taf9 | TAF9 | TAF9/TAF9b |
| | Taf10 | Taf10 | TAF10b | TAF10 |
| | Taf12 | Taf12 | TAF12 | TAF12 |
| | Spt7 | Spt7 | Spt7 | SUPT7L (STAF65G) |
| | Hifi1 (Ada1) | Hifi1 (Ada1) | Ada1 | TADA1 |
| | Spt20 | Spt20 | Spt20 | SUPT20H |

| | | | | |
|---|---|---|---|---|
| **TBP binding** | Spt3 | Spt3 | Spt3 | SUPT3H |
| | Spt8 | Spt8 | - | - |
| **TF-binding module** | Tra1 | Tra1 | Nipped-A (Tra1) | TRRAP |
| **Splicing module** | - | | SF3B3 | SF3B3 |
| | - | | SF3B5 | SFB3B5 |

Originally, the histone acetyl transferase GCN5 has been identified as a yeast homolog of the histone acetyltransferase A p55 of the ciliated protozoan *Tetrahymena thermophila* (Brownell et al. 1995, 1996). However, yeast and human GCN5 alone, can only acetylate free histones and not nucleosomal histones *in vitro* (Grant et al. 1997; Yang et al. 1996). The substrates for GCN5 correspond to multiple histone lysines *in vitro*, primarily lysine 14 of histone H3 (H3K14), but also H3K9, H3K18, H3K23, H3K27, H3K36, and additional lysines found in histones H4 and H2B (Brownell et al. 1996; Grant et al. 1997; Kuo et al. 1996; Suka et al. 2001). Chromatographic purification of a yeast fraction containing a HAT activity led to the identification of a 0.8 MDa and a 1.8 MDa native complexes both containing GCN5 and ADA2 (Grant et al. 1997). The 1.8MDa contained Spt3, Spt7, Spt20/Ada5 and corresponds to SAGA (Grant et al. 1997).

The lower molecular weight complex corresponds to the ADA complex. Both ADA and SAGA share common subunits, Gcn5, Ada2 and Ada3, but the ADA complex contains two distinct subunits. One of them is Ahc1 (Eberharter et al. 1999), which is necessary for the integrity of the ADA complex only but not for SAGA, showing that ADA represents an additional HAT containing complex in yeast. The other is Ahc2 which was found later (Lee et al. 2011).

Another complex similar to SAGA has been purified by chromatography with a HAT activity that was named SLIK for SAGA-like complex or SALSA for "SAGA altered, Spt8 absent", and capable of regulating transcription (Pray-Grant et al. 2002; Sterner et al. 2002). This complex contains Tra1, Spt3, Spt7, Spt20/Ada5, Ada1, Ada2, Ada3, Gcn5, Taf5, Taf6, Taf9 and Taf12 (Pray-Grant et al. 2002). However, Spt8 is missing due to the proteolytic cleavage of the C-terminal domain of Spt7, which is required for the interaction with Spt8

(Pray-Grant et al. 2002; Sterner et al. 2002; Wu et al. 2002). Rtg2 was found to be uniquely present in the SLIK complex and required for its assembly (Pray-Grant et al. 2002). Mutations in both SAGA and SLIK resulted in synthetic lethality but single mutations in either SAGA or SLIK were less severe and displayed differences. This data suggests that SAGA and SLIK have overlapping activities but also specific activities (Pray-Grant et al. 2002).

## 2. The SAGA complex is conserved in metazoan

### a. Conservation and divergence of the SAGA subunits

In metazoan, most of the SAGA subunits are conserved but several differences are observed (**Table 3**) [reviewed in (Helmlinger et al. 2017; Spedale et al. 2012)]. PCAF, a GCN5 homologue that has a C-terminal GCN5-related region (Yang et al. 1996), was found to be associated with TAF proteins and the metazoan counterparts of the yADA2, yADA3, and ySPT3 (Ogryzko et al. 1998), and is also found within SAGA. Taf5 and Taf6 from yeast which are present in both TFIID and SAGA, are replaced by TAF5L and TAF6L, as a result of duplication events, and are specific of SAGA in metazoan, except in flies where they exist but are not found in SAGA (replaced by WDA and SAF6 respectively) (Guelman et al. 2006a; Ogryzko et al. 1998; Weake et al. 2009).

In higher eukaryotes, several SAGA-like complexes have been identified: the PCAF complex (Ogryzko et al. 1998), STAGA (Martinez et al. 2001) and TFTC (Brand et al. 1999b; Wieczorek et al. 1998).

#### i. PCAF complex

Purification of flagged PCAF from nuclear extracts of HeLa cells revealed that it was associated with about 20 proteins within a complex, containing a nucleosomal histone activity (Ogryzko et al. 1998). The proteins associated were TRRAP, ADA2B, ADA3, SPT3, TAF5L, TAF6L, TAF9, TAF10 and TAF12 (Ogryzko et al. 1998). It was shown that PCAF preferentially acetylates H3 and weakly H4.

#### ii. STAGA

SUPT3 immuno-precipitation from nuclear extracts of HeLa cells pulled down GCN5 and TAF9, but TAF9 immuno-precipitation pulled down TFIID and SUPT3. These results showed that SUPT3 was not part of TFIID, but was part of another complex, together with GCN5 as well as with TAF9 (Martinez et al. 1998, 2001). This complex was thus called STAGA for SPT3-$TAF_{II}$31-GCN5-L acetyltransferase. In addition, mass spectrometry analysis of STAGA, showed that it also contains TRRAP, ADA3, SFGF29, TAF5L, TAF6L, ADA1 and SPT7L (Martinez et al. 2001). The spliceosome-associated protein 130 (SAP130), which is a component of the splicing factor SF3b, also associates transiently to STAGA (Brand et al. 2001; Martinez et al. 2001).

iii. TFTC

TAF10-immunoprecipitation followed by TBP-immunoprecipitation revealed the existence of a TAF-containing complex, that was not containing TBP and was called TFTC (TATA-binding Protein-free TAF-containing Complex), with a HAT activity (Wieczorek et al. 1998). It was described as containing several TFIID subunits: TAF2, TAF4, TAF5, TAF6, TAF7, TAF9, TAF10, TAF12 together with TRRAP, GCN5, ADA3, SPT3 (Brand et al. 1999b). So, the "TFTC complex" represented a mixture of partial TFIID assemblies and the SAGA complex (Demény et al. 2007).

Altogether, those different approaches revealed the existence of GCN5/PCAF-containing-complexes, without TBP, with similar composition that correspond to the metazoan ortholog of the unique SAGA complex.

b. ATAC

A distinct HAT complex, called ATAC, has been identified in metazoan. ATAC differs from SAGA with the presence of ADA2A instead of ADA2B in *Drosophila* and which is well conserved (Barlev et al. 2003; Guelman et al. 2006a, 2009; Kusch et al. 2003; Muratoglu et al. 2003; Nagy et al. 2010; Wang et al. 2008). The complex composition of ATAC has been elucidated by MudPIT (multidimensional protein identification technology) and revealed the presence of Gcn5, Ada3, HCF (host cell factor) and Atac2 among others (Guelman et al. 2006b). In *Drosophila*, the Atac2 subunit containing a putative acetyltransferase domain was suggested to acetylate H4K16 (Suganuma et al. 2008, 2010). In mammals, knock-down of

*Atac2* showed that ATAC was necessary for overall levels of H3K9ac, H4K5ac, H4K12ac, and H4K16ac in mammals (Guelman et al. 2009).

## 3. Modular organization of SAGA

SAGA is a multi-subunit complex that is organized into several modules (**figure 9**) [reviewed in (Helmlinger et al. 2017)]. Each module associates tightly to the complex, with the exception of the DUB modules variants that will be described thereafter, and participate to the enzymatic activities of SAGA. The first insight into SAGA structure was obtained by electron microscopy (EM) from TFTC purification from human cells with a 3.5-nanometer resolution (Brand et al. 1999a). SAGA was described as a five domains structure, including four domains similar to TFIID (Brand et al. 1999a). Another work in *S. cerevisiae*, showed that the yeast SAGA complex was similar in size and in structure to the human TFTC complex, and which mapped 9 of the 19 known SAGA subunits using single EM reconstruction (Wu et al. 2004). The EM analyses have been complemented by the combinatorial approach of gene deletions of non-essential SAGA subunits that gave a macromolecular model consisting of all 21 SAGA/ADA subunits (Lee et al. 2011). The most recent architecture of the SAGA complex was proposed by Steve Hahn's laboratory (Han et al. 2014). Protein-protein interactions within the complex were determined by crosslinking and mass spectrometry using TAP-tag Spt7 (Tandem affinity purification) for purifying SAGA with bound recombinant TBP, since it did not copurify with SAGA (Han et al. 2014). Moreover, recent single-particle EM analysis, highlighted the flexibility of the SAGA complex, which is able to adopt different conformational changes, as it has been already evidenced for TFIID (Setiaputra et al. 2015).

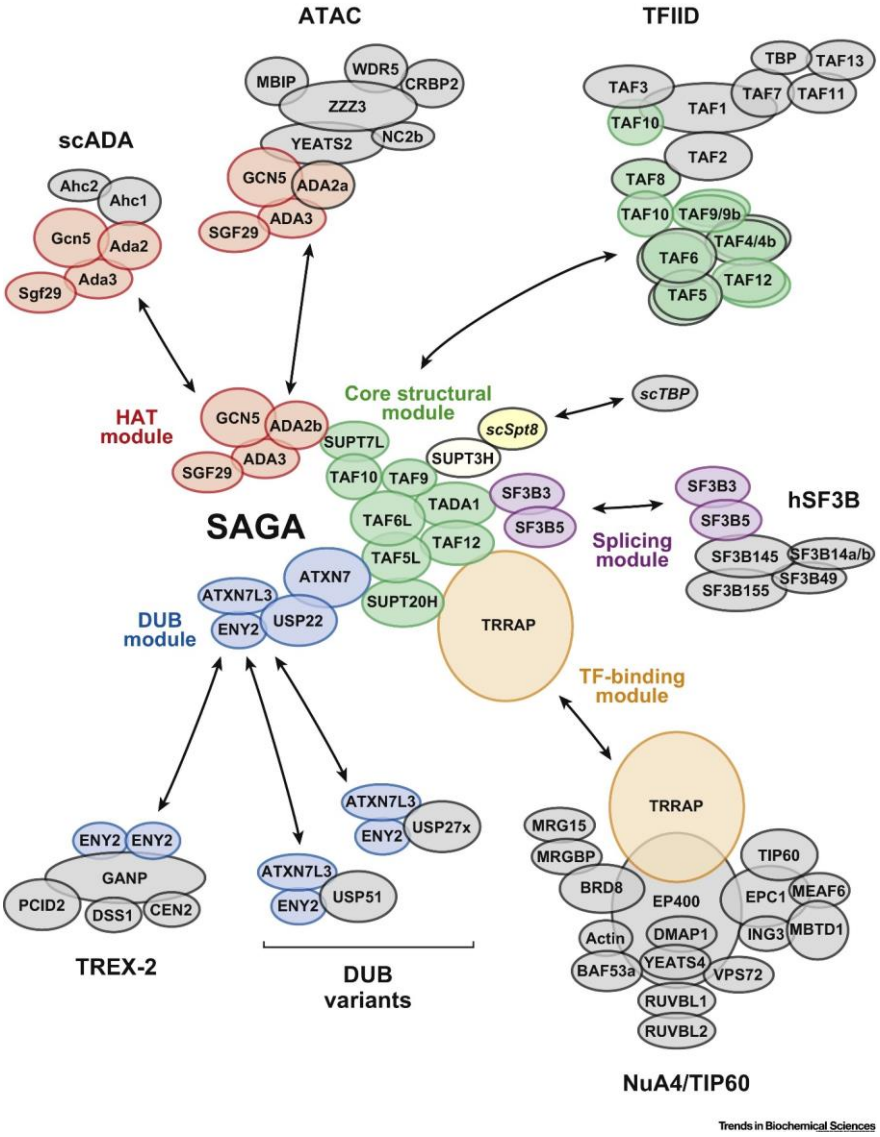



**Figure 9: Schematic representation of the modular organization of SAGA in relation with the other transcriptional complexes.** Human names are indicated, or yeast when mentioned (*sc*). The TAF core structural module (in green) is shared between TFIID and SAGA, note that TAF10 is incorporated into SAGA through SUPT7L and with TAF8 (according the structural model) in TFIID. Other proteins constitute the core module in addition of the TAFs. The HAT module (in red) shares GCN5, ADA3 and SGF29 with ATAC and with the yeast ADA complex. The DUB module (in blue) shares subunits with the TREX-2 complex but can also be associated with other USP enzymes to form DUB variants. The biggest protein TRRAP (in yellow) is also shared with nuA4/TIP60. SUPT3H, plus Spt8 in yeast, are the subunits in contact with TBP. The splicing module, (in purple) is shared with the human SF3B complex, which associates with the U2 snRNP to initiate splicing (Helmlinger and Tora 2017).

a. Core structural module

TFIID and SAGA share several common Tafs: Taf5, Taf6, Taf9, Taf10 and Taf12 in yeast (TAF5L and TAF6L instead of Taf5 and Taf6 respectively in mammals) (**figure 9**). It was suggested that two pairs of each histone fold containing pairs Taf6-Taf9 and Ada1-Taf12, together with two copies of Taf5 are also in SAGA (Gangloff et al. 2000; Han et al. 2014). Crosslinks have been detected between Taf12 and Ada1 as well as between Taf10 and Spt7 (Han et al. 2014), which is the partner for Taf10 to be incorporated into SAGA (Soutoglou et al. 2005). So, SAGA contains a TFIID-like core that serves as a connector for the other modules. Taf12 has been shown to be required for the integrity of the complex and for the nucleosomal HAT function (Grant et al. 1998). The analysis of the role of the Tafs in the architecture of SAGA has been complicated by the fact that they are essential in yeast (Green 2000). Deletion of *ADA1*, *SPT20* or *SPT7* resulted in the most severe phenotypes *in vivo*, due to the disruption of the complex (Grant et al. 1997; Sterner et al. 1999).

b. SPT module

As described previously, Spt7, Spt20 and Ada1 are important for the integrity of the SAGA complex, and can be considered as part of a larger module, comprising the Taf module and those subunits in the "core structural module", which is located in the center of the complex (**figure 9**) (Helmlinger et al. 2017; Lee et al. 2011; Setiaputra et al. 2015). Genetic and biochemical analyses in yeast demonstrated a physical interaction between Spt3 and Tbp (Eisenmann et al. 1992). Other findings interrogated the role of Spt3 for recruiting Tbp into SAGA, since Tbp was still found even in the deletion strain *spt3Δ* (Sterner et al. 1999). According this same study, Spt8 was required for SAGA-Tbp interaction, confirming other studies showing its implication for TBP binding (Eisenmann et al. 1992; Sermwittayawong et al. 2006; Sterner et al. 1999). Nevertheless, both Spt3 and Spt8 have been shown to be necessary for Tbp deposition at promoters of active genes (Baptista et al. 2017; Bhaumik et al. 2001; Dudley et al. 1999; Laprade et al. 2007; Larschan et al. 2001; Mohibullah et al. 2008).

c. Tra1/TRRAP

Tra1 has been demonstrated to mediate SAGA recruitment through direct interactions with activators (Brown et al. 2001; Helmlinger et al. 2011), as well with its mammalian

counterpart TRRAP which interacts with MYC and E2F1 (Helmlinger et al. 2017). *Tra1* is essential in *S. cerevisiae* and its role has not been completely elucidated. It has been proposed that the essentiality is mainly related to the NuA4 complex (Helmlinger et al. 2011), of which Tra1 is also a subunit in *Saccharomyces cerevisiae*. Concerning the location of Tra1 within the SAGA complex, divergent observations have been made. In some studies, Tra1 is described as close to Spt20 (Han et al. 2014; Setiaputra et al. 2015) while in others, Tra1 is positioned at opposite ends of SAGA (Wu et al. 2004). More recently, the cryo-EM structure of the entire SAGA complex in which Tra1 was resolved with an average resolution of 5.7 Å was described (Sharov et al. 2017). Based on this structure, Tra1 represents one entire lobe (lobe A) connected to the second lobe (lobe B), containing the rest of the SAGA complex (Sharov et al. 2017).

d. DUB module

The DUB module contains in yeast the catalytic subunit Ubp8 together with Sgf11, Sus1, and Sgf73 (**figure 9**) (Ingvarsdottir et al. 2005; Köhler et al. 2006, 2010, Lee et al. 2005b, 2009; Samara et al. 2010). The subunit Ubp8 is a de-ubiquitinylation enzyme, and its association within the SAGA complex was shown to be required for its activity (Lee et al. 2005b). This observation was already suggested by the *SPT20* depletion, resulting in a partial disruption of the SAGA complex, that led to a decrease of the H2B ubiquitinylation (H2Bub1) level (Henry et al. 2003). However, the decrease of H2Bub1 was not to the same extent than *Ubp8Δ* alone, suggesting that Ubp8 might belong to other complexes other than SAGA (Henry et al. 2003). Ubp8 and Sgf11, whose absence also causes a decrease of H2Bub1 levels, were shown to function within a module, since each one is reciprocally required for their association within SAGA (Lee et al. 2005b). Another evidence that supported the modular organization of the complex, was that neither the deletion of *UBP8* or *SGF11* altered the HAT activity of SAGA (Lee et al. 2004; Powell et al. 2004). The anchoring of the DUB module to the rest of the complex was demonstrated by the purification of the DUB module only when *SGF73* was absent (Lee et al. 2009).

In human cells, the situation is more complex, the ubiquitin protease USP22 (the yeast orthologue Ubp8) and ATXN7L3 (ySgf11), ENY2 (ySus1) and ATXN7 (ySgf73), forms a DUB module (**figure 9**) (Ingvarsdottir et al. 2005; Lee et al. 2005b; Powell et al. 2004; Rodríguez-Navarro et al. 2004; Zhang et al. 2008; Zhao et al. 2008). However, the ATXN7 protein family contains ATXN7, ATXN7L1 and ATXN7L2 that are all three expressed in

mammalian cells, are mutually exclusive and can form three different DUB modules (Helmlinger et al. 2017; Vermeulen et al. 2010). Moreover, two additional USP22-related deubiquitinating enzymes have been identified. USP51 and USP27X were described to function independently from SAGA, interacting with ATXN7L3 and ENY2 and competing with USP22 (Atanassov et al. 2016). The role of those independent DUB modules and the SAGA complex for de-ubiquitinylation remains to be elucidated.

e. HAT module

The HAT module of SAGA contains in yeast Gcn5 together with Ada2, Ada3, and Sgf29 (**figure 9**) (Balasubramanian et al. 2002; Horiuchi et al. 1997; Lee et al. 2011; Saleh et al. 1997). The deletion of *ADA2* in yeast resulted in the specific loss of the HAT module, demonstrating the anchoring role of Ada2 for the HAT module to SAGA (Lee et al. 2011). In metazoan, the HAT module also contains the yeast homologues of GCN5 together with Ada2b, Ada3 and Sgf29 (Gamper et al. 2009; Riss et al. 2015). In addition, in human, PCAF can also be part of the HAT module, but in an exclusive manner with GCN5 (Nagy et al. 2007; Krebs et al. 2011). SAGA through its HAT module preferentially modifies histone H3 on Lys9 (H3K9) and to a lesser extent Lys14 (H3K14) [reviewed in (Lee et al. 2007)]. Interestingly, the environment within the HAT module regulates the catalytic activity of GCN5 as demonstrated *in vitro*. The GCN5 HAT activity is stronger when incorporated into the HAT module of SAGA than in ATAC (Riss et al. 2015).

# V. TFIID and SAGA are required for transcription of nearly all genes in *S. cerevisiae*

1. TFIID and SAGA control the expression of a large fraction of the genome

TFIID and SAGA share compositional and architectural similarities as described previously. They have also been described as major actors in gene expression regulation in yeast. Initially, the essentiality of Tafs in transcription was debated (Moqtaderi et al. 1996a) but it was demonstrated that *TAF6, TAF9, TAF10* and *TAF12* were essential for transcription (Michel et al. 1998; Natarajan et al. 1998; Sanders et al. 1999). The role of TFIID and SAGA in transcription was then revisited by Richard Young's laboratory, with the use of high-density oligonucleotide arrays in temperature-sensitive mutants for *TAF5, TAF6, TAF9, TAF10* and

*TAF12*, that allowed a genome-wide analysis (Lee et al. 2000). Individually, the TFIID and SAGA shared Tafs control a limited fraction of the genome but when looking at the sum of these gene fractions, they actually control 70% of the genome (Lee et al. 2000). When considering, TFIID- and SAGA-specific subunit mutants, a lower number of genes was affected, about 30% for TFIID-specific subunits and about 12% for SAGA-specific subunits (Lee et al. 2000). Among the SAGA-specific subunits, Spt20 controls the expression of a largest fraction of genes compared to Gcn5 and Spt3, confirming previous results related to the loss of the complex integrity (Grant et al. 1997; Sterner et al. 1999). Since both TFIID and SAGA contain a HAT activity (Imhof et al. 1997; Mizzen et al. 1996), at least demonstrated *in vitro* for TFIID with Taf1, their contribution to gene expression was also analyzed. Taf1 was shown to control 27% of the genome while only 4% depends on Gcn5 (Lee et al. 2000). Double mutation for *TAF1* and *GCN5* showed that transcription of 25% of the genes decreased by two-fold or more (Lee et al. 2000). Further analyses of the role of TFIID and SAGA in *S. cerevisiae* based on steady-state mRNA levels were performed with the mutation of *TAF1*, *GCN5* and *SPT3*, alone or in combination, *TAF1* and *GCN5* or *TAF1* and *SPT3* (Huisinga et al. 2004). Taf1 and Gcn5 were shown to control the expression of 84% and 60% of the genome respectively, while Spt3 was shown to control only 11% of the genome (Lee et al. 2000). It is interesting to note that the previous study observed 24% for Taf1 and 3% for Spt3, indicating major differences in terms of genome dependency (Lee et al. 2000). Double mutation with *TAF1* and *GCN5* gave the same result than with RNA Pol II mutant *rpb1-1*, which was more important that Gcn5 alone or the catalytic Gcn5 mutant. Interestingly, double mutation of *TAF1* and *SPT3* nearly abolished transcription (Huisinga et al. 2004).

## 2. Classification of genes as TFIID-dominated or SAGA-dominated

Clustering analyses of genes affected in those mutants led the authors to the classification of genes into two categories: TFIID- and SAGA-dominated genes (Huisinga et al. 2004). TFIID-dominated genes, representing 90% of the genome, correspond to non-regulated genes contrary to the SAGA-dominated genes, representing 10% of the genome, that correspond to stress/inducible regulated genes (Basehoar et al. 2004). This classification correlated with *in vivo* formaldehyde crosslinking and chromatin immunoprecipitation (ChIP) assays performed in the literature (Huisinga et al. 2004). According to this classification, TFIID-dominated genes correspond to TATA-less promoters while SAGA-dominated

correspond to TATA-containing promoters (Huisinga et al. 2004). It might look surprising that TFIID would regulate predominantly TATA-less promoters, given the essential requirement of TBP for transcription. However, it has been shown that TBP can recognize other sequences, but they are not necessary for transcription of *RPS5*, suggesting other mechanisms for transcription of those genes (Kamenova et al. 2014).

## 3. New approaches reveal a global role for TFIID and SAGA

### a. TFIID

A conditional depletion strategy of several Taf subunits revisited recently the role of TFIID (Warfield et al. 2017). This approach was based on the plant-specific F box protein OsTIR1, which degrades the degron auxin repressor protein IAA7 fused to the C-terminus upon addition of the auxin 3-IAA (Warfield et al. 2017). Depletion of those subunits resulted in a decrease of RNA Pol II occupancy genome-wide, analyzed by native RNA Pol II ChIP-seq, from two- to four-fold, irrespective of the promoter configuration (TATA-containing or TATA-less) (Warfield et al. 2017). This decrease occurred also for the genes that were previously described as insensitive to *TAF1* inactivation (Bhaumik et al. 2001; Kuras et al. 2000).

In this study, the role of TFIID in transcription was investigated by analyzing the newly-synthesized mRNAs. Compared to steady-state mRNA levels analyses that were used previously, metabolic labeling of mRNA with 4-thiouracil (4TU), a nucleotide analog, allows the measurement of mRNA synthesis and thus a direct appreciation of transcription. mRNA synthesis and degradation (or mRNA decay) rates were determined mathematically by comparative dynamic transcriptome analysis (cDTA) (Sun et al. 2012). This approach confirmed RNA Pol II ChIP results, TFIID subunits depletion induced a global effect on gene expression. However, the magnitude of RNA Pol II transcription decrease varies, and genes with the smallest decrease compared to the global average was shown to be enriched for genes that were defined as Taf-depleted genes in previous studies (Warfield et al. 2017). Both mRNA synthesis and mRNA decay were decreased, however, not to the same extent (median decrease of 2.7-fold and 1.8-fold respectively) for all genes. It is interesting to note that a stronger decrease in mRNA synthesis rate was observed for TFIID-dominated genes. An even broader effect was described with *TAF5* depletion, suggesting that TFIID and SAGA act synergistically for transcription regulation (Warfield et al. 2017).

b. SAGA

Analyses of the distribution of histone H3K9 acetylation and H2B ubiquitinylation revealed that SAGA is acting genome-wide in yeast and mammals (Bonnet et al. 2014). Interestingly, de-ubiquitinylation of H2Bub1 was observed on the transcribed region of all expressed genes. Analysis of newly-transcribed mRNAs revealed a new perspective of SAGA transcriptional regulation. Depletion of Spt20 or Spt7 in yeast, a three to ten-fold decrease was observed for newly-transcribed mRNAs, together with a decrease of RNA Pol II peaks observed by ChIP-seq, independently of TFIID- or SAGA-dominated genes that was considered (Bonnet et al. 2014). On the contrary, steady-state mRNA levels remained unchanged for several genes and only a moderate decrease of less than two-fold was observed for some other genes. Those results confirmed other data obtained from a previous study that performed transcriptome analyses from several SAGA subunit mutants (Lenstra et al. 2011). In this study it was demonstrated that the role of SAGA in transcription was beyond the 10% of the genome described previously. This discrepancy can be explained by several factors. So far, functional studies of the contribution of co-activator complexes have mainly relied on the analysis of their binding sites relying on antibody-based techniques, such as ChIP-seq, and on steady-state mRNA analyses. As a consequence, detection of the binding of dynamic factors to chromatin was partial and the transcriptional role of those factors was biased by compensatory mechanisms in mRNA decay. Many technical limitations, such as antibody specificity that might differ between several experiments can explain differences obtained between several studies.

The general role of SAGA has been comforted by another study from our laboratory in collaboration with Steve Hahn's laboratory (Baptista et al. 2017). Instead of using a classical ChIP-seq approach for assessing the presence of SAGA in the genome, chromatin endogenous cleavage coupled with high-throughput sequencing (ChEC-seq) has been used (Baptista et al. 2017). ChEC-seq data for SAGA subunits showed cleavages at an average of 2,700 genes which belong to TFIID-dominated or SAGA-dominated genes (Baptista et al. 2017). SAGA occupancy was observed to be distal, and biased towards UAS, even more pronounced for SAGA-dominated genes (Baptista et al. 2017). Major down-regulation of newly-synthesized mRNAs was observed for RNA Pol II genes, irrespective of their classification as TFIID-dominated or SAGA-dominated genes. The decrease in mRNA synthesis rate was associated

with a decrease in mRNA decay rate and an increase of transcripts half-life. This data confirms a genome-wide action of SAGA for transcription.

Another aspect of this study was the investigation of the contribution of the different modules of SAGA for transcription. Single mutations of the catalytic subunits *GCN5* or *UBP8* gave very different results. Gcn5 depletion that was previously shown to impact only a minor subset of genes (Huisinga et al. 2004; Lee et al. 2000), was shown to decrease globally the levels of newly transcribed mRNAs. However, no decrease was observed for Ubp8, suggesting that increased ubiquitinylation levels does not affect RNA Pol II transcription (Baptista et al. 2017). Nevertheless, double mutation of *GCN5* and *UBP8* revealed a synergistic effect. In the deletion of TBP-binding subunits, *SPT3* and *SPT8*, only *spt3Δ* displayed a decrease in mRNA synthesis (Baptista et al. 2017). The most dramatic decrease was observed when combining the deletion of *SPT3* and *GCN5*. In this case, mRNA synthesis was decreased by 10-fold despite that the integrity of the complex was not impaired (Baptista et al. 2017). This data suggests that TBP binding and the HAT activity are the most essential aspects of SAGA in transcription regulation. Moreover, TBP occupancy was shown to be depleted from a selection of TFIID-dominated and SAGA-dominated genes, arguing for a general role of SAGA genome-wide (Baptista et al. 2017). However, the bias observed by ChEC-seq of SAGA recruitment towards a higher proportion of SAGA-dominated genes, was not really explained by those results. This observation was also made with the Med8-ChEC-seq for Mediator, that was enriched at SAGA-dominated genes (Grünberg et al. 2016).

In conclusion, those studies have shed new light on the role of TFIID and SAGA in transcription, as general factors acting genome-wide (**figure 10**). To recapitulate, from the most recent studies (Baptista et al. 2017; Bonnet et al. 2014; Warfield et al. 2017), it was shown that nearly all active yeast promoters have TFIID and SAGA binding, independently of their classification (TFIID- or SAGA- dominated, and TATA-less or TATA-containing promoters), localized to promoters or UAS respectively. It was also noted several differences between TFIID and SAGA. ChEC-seq analysis of Mediator binding profile underlined a similar profile between SAGA and Mediator, suggesting their cooperativity at UAS (Baptista et al. 2017; Grünberg et al. 2016, 2017), and biased towards SAGA-dominated genes. These differences are still not well understood and might be related to the promoter architecture that remains to be elucidated.

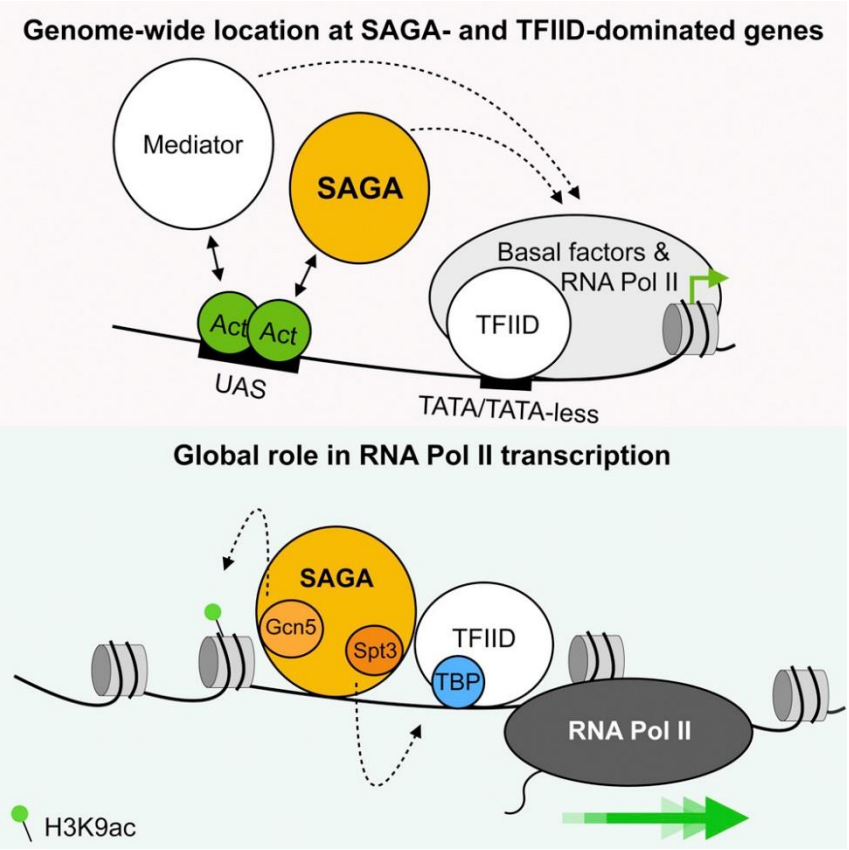



**Figure 10: Schematic model of the general role of TFIID, SAGA and Mediator complexes in *Saccharomyces cerevisiae*.** TFIID binds to both TATA-containing and TATA-less promoters, while Mediator and SAGA are located more distally, at UAS regions of both TFIID- and SAGA-dominated genes. TFIID, SAGA and Mediator act genome-wide and regulate transcription of nearly all genes (Baptista et al. 2017).

## VI. TFIID and SAGA roles during embryonic development in metazoans

The general contribution of TFIID and SAGA to transcription genome-wide has been well described in yeast. Nevertheless, in metazoans little is known about their function in gene expression and regulation. Moreover, their role has been complexified during evolution with the existence of several subunit paralogs. Cellular differentiation and embryonic development have represented a way to investigate their function. During embryonic development, one single cell, the fertilized egg, generates the cellular diversity of the organism from a unique genetic information. One cell gives rise to various cell types which have different and specific

functions. These processes depend on the activation of specific transcriptional programmes and on a tight spatio-temporal regulation of gene expression.

## 1. TFIID and SAGA subunit expression pattern during development

### a. TAF paralogs

Although TFIID has been described in higher eukaryotes as a complex composed of TBP and 13 TAFs, considered as the canonical complex, TAFs paralogs have been identified. TAF4b was the first TAF paralog to be identified, initially in human differentiated B-cells, and was found later to be enriched in ovary and testis in mammals while being expressed at low level in other tissues (Dikstein et al. 1996; Freiman et al. 2001). TAF4b interacts with TBP and other TAF subunits and thus, can be integrated into the canonical TFIID (Freiman et al. 2001). In *Drosophila*, several testis-specific TAFs paralogs exist. The genes *No hitter*, *Cannonball*, *Meiosis I arrest*, *Spermatocyte arrest* and *Ryan express* encode dTAF4, dTAF5, dTAF6, dTAF8 and dTAF12 homologs respectively (Hiller 2004). In vertebrates, TAF1L has been identified as a TAF1 paralog human specific and is specifically expressed in testis germ cells (Wang et al. 2002). TAF7L, the TAF7 paralog, was found in male germ cells (Cheng et al. 2007b; Pointud et al. 2003), but also in adipocytes and white fact tissue (Zhou et al. 2013a, 2013b). TAF9b is a TAF9 homolog, originally found in chicken cells that exists also in human (Cheng et al. 2007b; Pointud et al. 2003). TAF9 and TAF9b are present together in the different cell lines tested, but their ratio might vary, and thus participate to potential alternative TFIID complexes (Cheng et al. 2007b; Pointud et al. 2003). During neuronal differentiation, TAF9b is highly expressed in ES cell differentiated neurons (Herrera et al. 2014). Altogether, this data suggests that the TFIID complex composition can be modulated with the incorporation of some TAF paralogs in certain tissues and potentially regulating tissue specific subset of genes.

### b. TBP paralogs

TBP was initially thought to be universal but TBP-like proteins have been identified form the analyses of expressed sequence tags. They revealed several TBP paralogs known as TBP-related factor (TRF) or TBP-like factor (TLF). TRF1 was the first to be identified, and is specific to insects (Crowley et al. 1993). TRF1 exhibits a high similarity to the TBP core domain that recognizes TATA elements. TRF2 (or TLP, TRP, TLF, and TBPL1) is another TRF that

has been identified and is shared between several organisms such as *Caenorhabditis elegans*, *Drosophila*, mouse, and humans, but not in yeast or *Archaea* (Dantonel et al. 2000; Ohbayashi et al. 1999; Rabenstein et al. 1999). Interestingly, TRF1 is restricted to neural tissue in *Drosophila* (Crowley et al. 1993) but northern blot analysis of human tissues showed that TRF2 is broadly expressed, with an enrichment in the testis (Moore et al. 1999; Rabenstein et al. 1999). Interestingly, TRF2 inactivation leads to development arrest during gastrulation. TBP and TRF2 are proposed to be complementary but differentially regulate transcription. TRF2 function has been analyzed in *Caenorhabditis elegans* and showed that it is required also for early development, its inactivation impairs morphogenesis (Dantonel et al. 2000). Last but not least, TBP2 (also known as TBPL2, TRF3) has been identified only in vertebrates (Bártfai et al. 2004; Gazdag et al. 2007; Jallow et al. 2004; Persengiev et al. 2003). TBP2 shares 95% homology with the core domain of TBP (Bártfai et al. 2004). Its pattern of expression differs between species. In human, TBP2 is detected in several cell lines (Persengiev et al. 2003), in zebrafish it is expressed in the early embryo and enriched in the gonads (Bártfai et al. 2004) while TBP2 is exclusively expressed in oocytes in mouse (Bártfai et al. 2004; Xiao et al. 2006). In mouse TRF2 is required for spermatogenesis (Martianov et al. 2001, 2002a; Zhang et al. 2001) and TBP2 for oogenesis (Gazdag et al. 2007, 2009). TRF1, TRF2 and TRF3/TBP2 interact with TFIIB and TFIIA (Bártfai et al. 2004; Crowley et al. 1993) as well as with other TAFs for TRF1 and TRF2 (Crowley et al. 1993), but TRF2 is not able to bind the TATA-box.

### c. TFIID composition is variable

The TFIID composition is variable, and several evidences highlight such flexibility in different cellular contexts. Comparison of the level of expression of TFIID subunits between cell types revealed that several of them are downregulated during cellular differentiation. In mES cells and embryonic fibroblasts, TAF5 and TAF6 are enriched compared to NIH 3T3 fibroblasts (Pijnappel et al. 2013). During differentiation, several TFIID subunits have been described to be down- or up-regulated. For instance, TAF8 was shown to be selectively up-regulated during differentiation of pre-adipocyte fibroblasts into full adipocytes (Guermah et al. 2003). Retinoic acid induced differentiation of F9 carcinoma cells into primitive or parietal endoderm results in the selective loss of TBP and TAF4 while the protein level of the other TFIID subunits tested remain unchanged (Perletti et al. 2001). Such depletion is concomitant with the depletion of the nuclear receptor RARγ2 which is mediated by the proteasome. Other

studies have proposed that both TBP and TAF4 are reduced in adult hepatocytes (D'Alessio et al. 2011; Freiman et al. 2001), but it was not confirmed for TBP in a more recent *in vivo* study (Alpern et al. 2014). Furthermore, during myogenic differentiation, TAF1, TAF4 and TBP were also reported to be depleted in myotubes (Deato et al. 2007; Malecova et al. 2016; Zhou et al. 2013b). In line with a post-translational regulation occurring during differentiation, TBP can be ubiquitinylated specifically during myotube differentiation (Li et al. 2015). Analysis of the TFIID composition in human embryonic stem cells (hES) revealed that contrary to mouse embryonic stem cells (mES) they contain only 6 out of the 14 canonical TAFs, which become expressed during retinoic acid differentiation (Maston et al. 2012). The fact that human pluripotent cells have a different set of TAFs than mES cells is surprising and may be due to technical issues. Altogether, these results reveal that physiological changes are accompanied with regulation of components of the basal transcription machinery.

#### d. GCN5 and PCAF display different expression pattern

GCN5 and PCAF are both HAT enzymes and they display distinct expression pattern during development. In mouse embryo, *Gcn5* expression is detected as early as E7.5. At E8.5, *Gcn5* is expressed in the whole embryo but not in the allantois and the heart. By E18.5, *Gcn5* expression decreases (Xu et al. 2000). Instead, *Pcaf* is detected at very low level in the embryo, and was detected by *in situ* hybridization from E12.5 (Yamauchi et al. 2000; Xu et al. 2000). Its expression has been reported to be strongly increased in adult tissues (Xu et al. 2000).

### 2. TFIID role during development

#### a. TFIID subunits are required for early embryogenesis

In mouse, *Tbp*, *Taf7*, *Taf8* and *Taf10* null mutation lead to embryonic lethality between E3.5 and E5.5 (Gegonne et al. 2012; Martianov et al. 2002b; Mohan et al. 2003; Voss et al. 2000). At the blastocyst stage, the inner cell mass undergoes massive apoptosis for *Tbp*, *Taf8* and *Taf10* null mutants (Martianov et al. 2002b; Mohan et al. 2003; Voss et al. 2000). However, trophoblast is viable in the absence of TAF10, despite endoreplication was affected (Mohan et al. 2003). Interestingly, transcription was differently affected in the *Tbp* and *Taf10* mutants. While RNA Pol II was affected by the loss of TAF10 in the blastocyst, only RNA Pol I and Pol III transcription was affected in the absence of TBP in the blastocyst (Martianov et al. 2002b;

Mohan et al. 2003). This suggests that the lack of TBP could be compensated for RNA Pol II transcription. Among the other TAFs, *Taf4* null mutation impaired embryogenesis at E9.5, a stage at which the paralog TAF4b cannot compensate anymore for the absence of TAF4 (Langer et al. 2016). No deletion of other TAFs has been reported, but altogether, this data already indicates that TFIID plays an important role in the embryo.

b. Some TFIID subunits are differentially required during development

i. TAF10 and TAF7 are differentially required depending on the cellular context

Although several TFIID subunits are essential for mouse early development, their requirement is, however, relative to the cell-type. The best described example is TAF10. TAF10 is an ubiquitous and scaffold protein required for the survival of both inner cell mass and proliferative F9 carcinoma cells, but not for trophoblasts and retinoic acid differentiated F9 carcinoma cells into primitive endoderm (Mohan et al. 2003; Metzger et al. 1999). Nonetheless, TAF10 is required for cell cycle progression of F9 carcinoma cells and endoreplication of trophoblasts. Interestingly, cellular survival cannot be explained by the maintenance of transcription, since it is impaired in trophoblasts which are viable in the absence of TAF10 (Mohan et al. 2003). In other contexts, TAF10 has been shown to be also differentially required for transcription in *C. elegans* embryo and chicken cells (Chen et al. 2000; Walker et al. 2001). Another example is TAF7 which is required for proliferation of thymocytes prior their differentiation from CD4 CD8 double-negative to double positive thymocytes, but not for the final steps of thymocyte development (Gegonne et al. 2012). The absence of TAF7 prevented the expansion of T-cells after antigenic stimulation, demonstrating the requirement of TAF7 for cellular proliferation.

ii. TAF10 is differentially required depending on the developmental stage

Functional analyses of TAF10 in keratinocytes showed that TAF10 is necessary for establishing the skin at the foetal stage (Indra et al. 2005). Loss of TAF10 affected terminal differentiation of keratinocytes and thus impaired skin barrier function. Expression of several keratinocyte specific genes was also decreased. However, *Taf10* inactivation in adult

keratinocytes did not impair skin integrity, neither their induced proliferation upon UV-irradiation. Total mRNA levels of key genes for epidermal function encoding the Transglutaminase 1, the kruppel-like protein and the tight junction protein Claudin-1 remained unchanged in the mutant animals. Interestingly, the skin retained its ability to regenerate after wounding, suggesting that transcription is still functional (Indra et al. 2005). Conditional inactivation of *Taf10* in both foetal and adult hepatocytes also showed a differential requirement for TAF10 (Tatarakis et al. 2008). Foetal hepatocytes devoid of TAF10 at E15.5, failed to fully differentiate and leads to embryonic lethality. Loss of TAF10 in post-natal hepatocytes resulted in a smaller liver and a dwarf phenotype but animals looked normal. However, they could not survive after P38 (Tatarakis et al. 2008). This phenotype is however milder than the conditional *Taf4* deletion in neonatal hepatocytes, where animals die by P21, confirming that not all TAFs are functionally equivalent and have probably different implications in transcription (Alpern et al. 2014).

#### iii. TAF10 is required for initial gene activation

*Taf10* ablation in foetal hepatocytes led to a major downregulation of hepatic specific genes but to an upregulation of non-hepatic genes, related to hematopoietic or other epithelial cell types that are present in the embryonic liver (Tatarakis et al. 2008). However, in adult hepatocytes, hepatic expressed genes that were already activated during the foetal life were not impaired at P30. Those results suggest that TFIID is actually necessary for initial activation of genes during development. So, this data suggests that embryonic transcription requires a canonical TFIID complex where most of the genes have not been activated yet. But, terminally differentiated cells can support a certain flexibility of TFIID, at least for genes that have been already activated prior TAF10 depletion. Interestingly, it is important to note that genes that were normally silent in the adult, were de-repressed in the absence of TAF10. So, TFIID is necessary for transcription activation but could also play other roles in regulating gene silencing to ensure proper gene expression. This data suggests, that components from the basal transcriptional machinery are differentially required depending on the cellular context.

## 3. SAGA role in development

### a. HAT module

*Gcn5* null mutation leads to several developmental defects at E8.5, characterized by a growth defect by E10 (Xu et al. 2000). Mutant embryos exhibit apoptosis of ectodermal and mesodermal tissues, thus precluding the formation of axial mesoderm, head mesenchyme and paraxial mesoderm (Xu et al. 2000). Such cell death is observed before the morphological defects, suggesting that it is the main cause responsible of the phenotype. Double deletion of *Gcn5* and *p53* did not rescue the lethality phenotype but slightly increased their survival until E11.5, and the development of somites, notochord as well as the brain was initiated (Bu et al. 2007). The point mutation in the catalytic center of GCN5 led to mixed phenotypes among the progeny (Bu et al. 2007). While some $Gcn5^{hat/hat}$ mutant embryos exhibited a similar phenotype than the $Gcn5^{-/-}$, most of them survived even longer than $Gcn5^{-/-}$ or $Gcn5^{-/-};p53^{-/-}$, until E16.5 (Bu et al. 2007). Those mutants displayed apoptosis at E8.5 and exhibited a neural tube closure defect by E8.5, followed by exencephaly at E11.5. However, the mutants for *Pcaf*, a *Gcn5* homolog, are viable, in line with their different expression pattern (Xu et al. 2000; Yamauchi et al. 2000). Interestingly, double mutation for *Pcaf* and *Gcn5* results in earlier lethality (at E8.5) than *Gcn5* single mutation suggesting overlapping functions for those two HAT proteins (Xu et al. 2000; Yamauchi et al. 2000). Furthermore, the expression of several mesodermal specification markers was not affected in $Gcn5^{-/-}$ or $Gcn5^{hat/hat}$ mutant embryos, indicating that transcription is not impaired (Bu et al. 2007; Xu et al. 2000). Deletion of *Gcn5* in mES cells did not induce massive apoptosis and $Gcn5^{-/-}$ mES cells were able to generate the three germ layers when aggregated into embryonic bodies, suggesting that transcription is not affected and permits cellular differentiation (Lin et al. 2007). However, these cells failed to persist by E12.5 in chimeric embryos generated by injection of $Gcn5^{-/-}$ mES cells into wild-type blastocysts since they were less competitive than the wild-type ones (Lin et al. 2007). Conditional deletion of *Gcn5* in neural stem cells resulted in microcephaly, detected from E14, with a 17% mass reduction, and even more pronounced by adulthood, with a severe mass reduction compared to wild-type, due to a proliferation decrease of precursor cells of the developing cortex (Martínez-Cerdeño et al. 2012). Gene expression analyses from neurospheres, established from E12.5 control and mutant embryos, indicated that about 4000 genes were differentially expressed upon *Gcn5* deletion, 17% of them were downregulated also in *N-myc* knockout cells, an important factor controlling neural stem cells in the brain (Martínez-Cerdeño et al. 2012). These down-regulated genes were also acetylated by GCN5, and the authors proposed that GCN5 could be a co-activator for N-MYC in the developing brain. In *Drosophila*, the HAT activity is also

required for viability (Kusch et al. 2003; Muratoglu et al. 2003; Qi et al. 2004; Pankotai et al. 2005; Zsindely et al. 2009). In another study, the HAT activity was shown to be necessary for oogenesis in *Drosophila*, associated with transcription defects for many genes, except for several oocyte specific-genes that were expressed normally (Li et al. 2017).

b. DUB module

Ablation of *Usp22* was shown to be embryonic lethal, the mutants embryos degenerated by E10.5 (Lin et al. 2012). The authors described that the absence of USP22 caused the suppression of p53 activity (Lin et al. 2012). Hypomorphic mutant of *Usp22* displayed a decreased body size and weight compared to their wild-type and heterozygous littermates (Kosinsky et al. 2015). Moreover, USP22 was shown to be required for lineage specification in the small intestine and in the brain, where the cellular composition was changed in the mutants. More stem cells were detected in the small intestine, and impaired cortical differentiation in the brain (Weake et al. 2008). In *Drosophila*, Nonstop, the yeast Ubp8 homologue, and Sgf11 are required for correct axon-targeting during neural development (Weake et al. 2008). Both regulate similar subset of genes that were shown to be distinct than the ones dysregulated by *ada2b*, from the HAT module, suggesting distinct regulation between the de-ubiquitinylation and acetylation activities of SAGA (Weake et al. 2008). Furthermore, the DUB module was shown to be implicated during the cellularization stage in *Drosophila*, and thus to have distinct roles than the HAT module (Li et al. 2017). Indeed, the DUB module was found to regulate a subset of genes (about 40%) in *Ataxin-7* and *non-stop* mutant embryos (Li et al. 2017). Interestingly, in this study the DUB module was found to bind to genes that were not regulated by *Ataxin-7* and *non-stop* mutant embryos (Li et al. 2017). However, the absence of phenotype for the mutants of the DUB used in this study could also be explained by the compensation by other free-DUB for which little is known.

c. SUPT20

From a N-ethyl-N-nitrosourea genetic screen in mouse, defects in gastrulation and neural tube closure were observed with the hypomorphic splicing mutation of *Supt20*, coding for the mouse SUPT20 SAGA subunit, called p38-interacting protein in this study (Zohn et al. 2006). In these mutants, mesoderm cell migration was impaired, explained by a lack of p38 kinase activation, together with a downregulation of the gene encoding E-cadherin required for

Epithelial–Mesenchymal Transition. It is associated with neural tube defects, resulting into exencephaly. Another study using the same hypomorphic mutant reported skeleton defects, with a fusion of the rib and vertebrae of the lower thoracic region (Warrier et al. 2017). The cyclic expression of *Lfng* and *Hes7* was conserved in *Supt20* hypomorph embryos, but *Lfng* was decreased in the caudal domain while *Hes7* was increased in the rostral domain (Warrier et al. 2017). This phenotype suggests a regionalization defect of the somite. In these two studies, *Supt20* function is analyzed independently of the SAGA complex, it is not known whether the activity of the complex is affected and whether it could also contribute to the phenotype.

# VII. A new paradigm to study the role of transcriptional complexes

The developing embryo represents an interesting paradigm for revisiting the mechanisms of transcription initiation. During development, activation of specific transcriptional programmes drive cellular processes that shape the embryo. In this respect, vertebrate segmentation is a morphogenetic event associated with a tight regulation of gene expression, which is controlled by cyclic expression of several genes. Although cyclic gene expression has been described in various biological systems (for example the circadian clock), cyclic expression in the presomitic mesoderm (PSM) represents another interesting situation in the embryo. That is why, we chose to analyze the role of the basal transcription machinery during this developmental event.

## 1. Vertebrate segmentation

Vertebrate segmentation is a morphogenetic process that generates repetitive epithelial structures called somites. Instead of simultaneous generation of segments in the whole body, like in *Drosophila*, vertebrate segmentation proceeds sequentially [reviewed in (Hubaud et al. 2014)]. Metameric segments are generated by pair on each side of the neural tube from the segmentation of the paraxial mesoderm (PM) (**figure 11**).

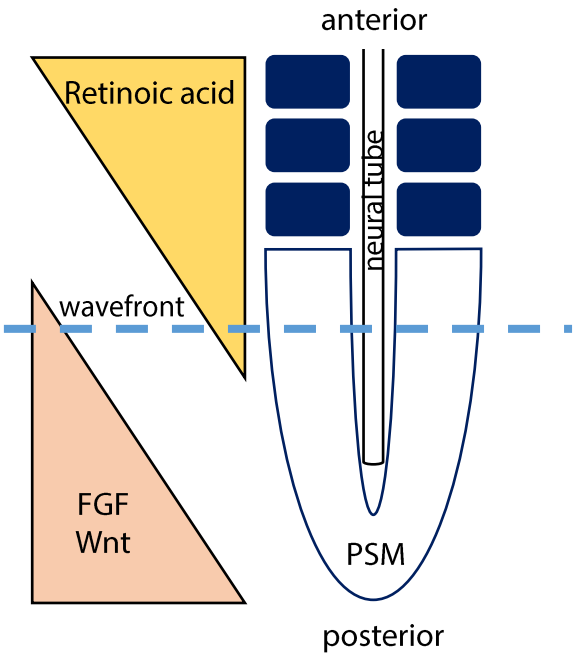



**Figure 11: Schematic representation of the presomitic mesoderm and the signaling gradients**. The PSM tissue flanks both sides of the neural tube. Retinoic acid is an opposite gradient to the posterior-anterior FGF/Wnt signaling gradient in the PSM. The intersection of FGF/Wnt and retinoic acid gradients defines the wavefront. Adapted from (Vincent et al, 2012).

Somites will give rise to the axial skeleton, the dermis of the back, and skeletal muscles of the trunk and the limbs (Christ et al. 1995). The number and the size of those repetitive epithelial segments, is specific of every species. For instance, in the mouse, a pair of somites is regularly formed every two hours until a total of 65 pairs are formed. Somitogenesis starts at E8.0 and ends at E13.0 in mouse, from the unsegmented paraxial mesoderm called PSM. The source of cell progenitors of the PSM originates from the primitive streak and later from the tailbud, and thus represents the motor for elongation (Bénazéraf et al. 2010; Cambray et al. 2002; Tzouanacou et al. 2009). The progenitor cells from the PSM remain undifferentiated until they are instructed to differentiate during the somite formation.

2. The clock and wavefront model

Vertebrate segmentation is a tightly regulated process, controlled in a spatio-temporal manner by the "clock and wavefront model", initially proposed by Cooke and Zeeman (Cooke et al. 1976).

a. The segmentation clock defines the pace

The clock consists in an oscillatory expression of genes belonging to Notch, Wnt and FGF signaling pathways which pace the segmentation. Such oscillation of gene expression makes the PSM a very interesting model for analyzing transcription. Oscillatory expression occurs in every cell, travelling as a posterior to anterior wave along the PSM. The clock controls the cell competency to respond to the signal triggered by the travelling wavefront, which corresponds to a system of signaling gradients. The first evidence of the existence of a cyclic gene in the PSM has been obtained by *in situ* hybridization of *Hairy1* in chicken (Palmeirim et al. 1997). This gene exhibits a periodic pattern, which is cell autonomous, and occurs for every somite formation. Thus, it has been the first evidence of an intrinsic developmental clock that controls a morphogenetic event. A microarray analysis in mouse identified more than 40 cyclic genes of the NOTCH, FGF and Wnt signaling pathways (Dequéant et al. 2006). The oscillatory network is also well conserved among vertebrates but individual cyclic genes differ between species (Krol et al. 2011). Finally, the clock induces the activation of the segmental program with the activation of the genes of the mesoderm posterior (*Mesp1* and *Mesp2*) genes.

b. The wavefront determines the new somites

In the PSM, Wnt and FGF constitute a posterior to anterior signaling gradient and retinoic acid establishes an opposing gradient (**figure 12**) [reviewed in (Aulehla et al. 2010)]. The FGF gradient, and possibly Wnt gradient, is generated in the tailbud by the progressive decay of mRNA from the cellular progenitors that stop transcribing the genes encoding the signaling effectors (i.e. FGF8 for FGF signaling pathway) when entering the PSM (Dubrulle et al. 2004) [reviewed in (Aulehla et al. 2010)]. Retinoic acid is produced in the somites and diffuses in the PSM. This results in an anterior-posterior gradient of retinoic acid gradient, which has been demonstrated by the expression pattern of *Raldh2*, an enzyme involved in retinoic acid synthesis in this region (Niederreither et al. 1997). The regionalization of the retinoic acid activity was shown to be controlled by the expression of retinoic acid degradation enzymes posteriorly, such as CYP26 (Sakai et al. 2001). During PSM segmentation, these

signaling gradients play a major role for defining the new somites. The high concentration of FGF posteriorly maintains PSM progenitors undifferentiated (**figures 11 and 12**). Below a certain threshold of FGF, cells become competent to respond to the signal triggered by the clock (Aulehla et al. 2010; Dubrulle et al. 2001). The region where FGF/WNT and retinoic acid gradients reach a certain threshold is the determination front, and it progressively moves from the anterior to the wavefront posterior end [reviewed in (Hubaud et al. 2014)]. The cells that pass the determination front downregulate *Msgn* and *Tbx6* expression, which are expressed in the progenitors from the posterior PSM. Instead, these cells start expressing instead the *Mesp* genes that control segment boundaries as well as adhesion genes (Saga et al. 1997).

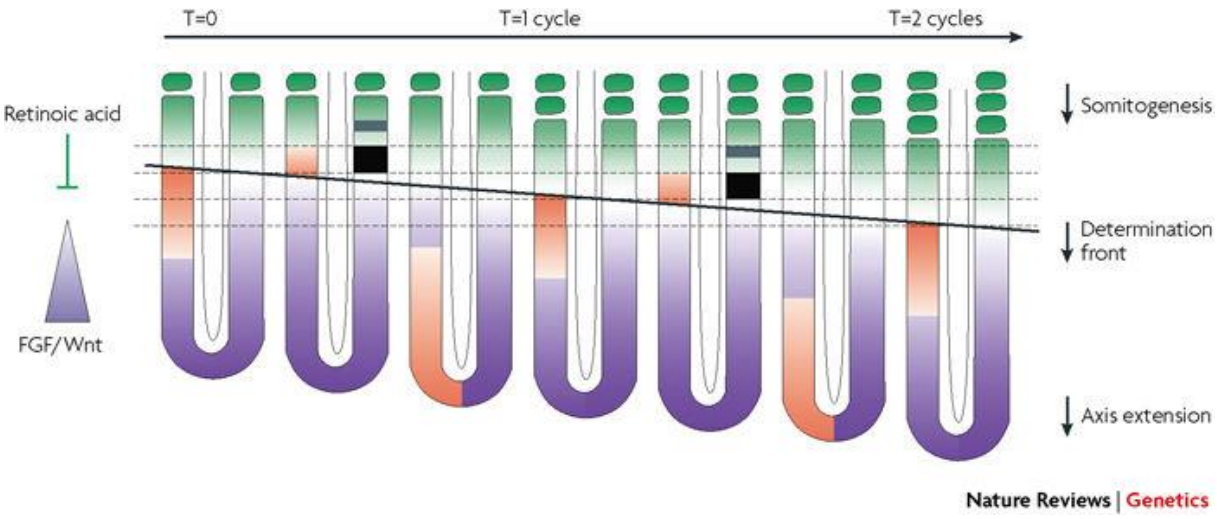



**Figure 12: Vertebrate segmentation of the PSM.** The determination front (wave front) moves posteriorly while the axis elongates and defines the new somites (Dequéant et al. 2008).

3. Gene oscillatory expression in the PSM

a. Negative feedback loops generates expression oscillation

As presented previously, several genes have an oscillatory expression in the PSM. The model for oscillations was described to rely on negative feedback loops (**figure 13**). The *her* genes from the Notch pathway in zebrafish, *her1* and *her7*, or *Hes1* and *Hes7* in mouse, have a cyclic expression based on the repression of their own expression through the action of repressors that they are encoding (Bessho et al. 2001, 2003; Oates et al. 2002). Those proteins

belong to the basic helix–loop–helix protein family, and HES1 protein was shown to bind its own promoter (Sasai et al. 1992; Takebayashi et al. 1994). In order to elucidate the mechanisms controlling *Hes1* oscillation, Hirata *et al.* recapitulated a two-hour *Hes1* oscillation in different cellular systems after a single serum treatment (Hirata et al. 2002). Both mRNA and protein oscillate, with a two-hour period, but with a 15-minute shift. Modulation of the protein stability of HES1 was shown to modulate *Hes1* mRNA levels. Proteasome inhibition, as well as addition of a dominant negative version of HES1, induced a repression of *Hes1* transcription. Instead, translation inhibition led to an increase of *Hes1* mRNA. Similar experiments were performed *in vivo*, in the PSM, and the same observations were made. Whereas proteasome inhibition induced a downregulation of *Hes1*, translation inhibition up-regulated *Hes1* mRNA levels. Furthermore, expression of the 5' region of *Hes1* gene under *Hes1* promoter in the PSM of *Hes1* $^{-/-}$ mice, led to the upregulation of *Hes1* mRNA (Hirata et al. 2002). The same laboratory reported the same behavior for *Hes7* and *Lfng* (Bessho et al. 2003). Transcription of those genes was upregulated in the absence of HES7 protein and conversely downregulated when Hes7 protein was stabilized. Gene oscillations can be visualized *in vivo* with gene reporters. Among them, the *LuVeLu* mouse line expresses a fluorescent reporter (Venus–YFP) under the control of a fragment of the cyclic *Lfng* promoter (Aulehla et al. 2008). The Venus-YFP protein is fused with a modified PEST domain (a peptide sequence rich in proline (P), glutamic acid (E), serine (S), and threonine (T)) and the *Lfng* 3'UTR is added to the construct, increasing the instability of the protein and the transcript respectively.

b. Mathematical modeling of gene oscillation

The oscillatory mechanism was conceptualized by Julian Lewis (Lewis 2003). He proposed a mathematical model in which the period of oscillations depends on the delay between transcription and translation (**figure 13**), as it was already proposed for other biological systems (Winfree 1989).


**Figure 13: Schematic representation of negative feedback loops for gene oscillation.** Gene oscillation is generated by the auto-inhibition of the expression of the gene by its own protein. It requires a time delay which results from transcription of the gene (Tm) and from mRNA translation into a protein (*Tp*). Adapted from (Lewis et al. 2003).

According to this model, the sum of the transcriptional and translational delay corresponds to the oscillation period. In order to generate sustainable oscillations, this model underlines the necessity for short half-lives of mRNA and protein products, together with a high rate production of those molecules. Based on this mathematical model, Julian Lewis predicted a 30-minute oscillation for the *her1/her7* oscillator in zebrafish that corresponded to what was observed *in vivo*. The same mechanism was also shown for transcription of the gene encoding the glycosyl-transferase Lunatic fringe (*Lfng*) in the chicken (Dale et al. 2003). *Lfng* expression is induced by Notch signaling, and the glycosyltransferase, encoded by *Lfng*, in turns modifies the Notch receptor so that it cannot activate transcription anymore until it gets degraded (Dale et al. 2003).

### c. Determination of the parameters controlling gene oscillation

Establishing a time delay between transcription and translation can be achieved at several steps, which are: (1) transcription initiation at the promoter; (2) transcription elongation; (3) splicing; (4) mRNA 3'-end processing and polyadenylation; (5) release of the transcript from the site of transcription; and (6) mRNA export through the nuclear pore.

#### i. Transcription elongation is not a critical parameter for gene oscillation

By using the *her1* and *her7* zebrafish genes as a model, the laboratory of Julian Lewis explored the importance of transcription elongation in the model he proposed earlier (Hanisch et al. 2013; Lewis 2003). Those two genes differ in length (6392 bp for *her1* and 1304 bp for

*her7*) and thus were expected to have a different transcription elongation duration based on an RNA Pol II elongation rate of 1.2 kb/minute (Hanisch et al. 2013). A different transcription elongation time would result in a substantial time period difference and, as a consequence, in a different somite number. To determine the transcription elongation rate, the authors used fluorescent *in situ* hybridization with two probes. The first one was recognizing the initial region of the artificial intron introduced in *her7* and the second was recognizing a region of the intron downstream of the first probe. The offset between the anterior boundaries of staining with probes AI1 and AI2 allowed to determine mathematically the time taken by RNA Pol II to cover the distance separating the regions they recognize. They estimated that RNA Pol II elongation rate was in their conditions 4.8±0.5 kb/minute (Hanisch et al. 2013). This transcription elongation rate is fast compared to measures obtained in other systems, for example it was measured 1.3 kb/minutes in human fibroblasts by BrU-seq (Veloso et al. 2014), but is in the range of the elongations rates determined by live-imaging by others (Darzacq et al. 2007). From this experiment, they predicted theoretically a similar transcriptional delay for both *her1* and *her7*, assuming a splicing delay of five minutes and an export time of four minutes (Hanisch et al. 2013). So, transcription elongation appeared to not be the major determinant for the time delay. Moreover, increasing the length of the gene *Lfng* in mouse with 10kB of intron of human *dystrophin* did not affect neither segmentation nor the clock (Stauber et al. 2012).

### ii. Role of splicing

Nevertheless, the role of splicing could not be excluded since another study showed that intron removal from *Hes7* resulted in a time delay, abolished oscillations and segmentation defects (Takashima et al. 2011). They did not consider the 34% decrease of the protein level compared to wild-type as responsible of the phenotype, since a 80% protein decrease did not dampen oscillations in a previous report (Palmeirim et al. 1997). The *Hes7* promoter-driven *luciferase* gene reporter without introns was expressed 19 minutes earlier than the one with full introns. This suggests that splicing is indeed a critical step in establishing the time delay necessary for generating oscillations (Takashima et al. 2011).

In line with this observation, Hoyle and Ish-Horowicz measured *in vivo* kinetics of mRNA splicing and export (Hoyle et al. 2013). Both *Lfng* and *Hes7* display a transcriptional delay due to splicing, accounting for 9.7 ± 2.8 min and 12.5 ± 6.3 min respectively (Hoyle et

al. 2013). Transcript nuclear export kinetics were shown to be even slower with 16.1 ± 5.7 min for *Lfng* and 17.0 ± 8.3 min for *Hes7*. Together, this data fits with the negative feedback loop model that generates gene expression oscillation, as proposed by Julian Lewis (Lewis 2003). The oscillation is the result of a delay between transcription and translation, and thus, both splicing and mRNA export are two main critical parameters.

### iii. Role of transcript stability

In addition of transcriptional influence on the time delay, the transcript stability has also been studied. *Lfng* and *Hes7* transcripts have been described having different stability in the PSM resulting in a slightly different spatiotemporal distribution when observed by fluorescent *in situ* hybridization (Nitanda et al. 2014). The 3'-UTR of either *Lfng* or *Hes7* was shown to be responsible for the differential transcript stability and generates different distributions of those mRNAs (Nitanda et al. 2014).

Post-translational control of gene oscillations by micro RNAs (miRNAs) targeting the 3'-UTR of cyclic genes has been proposed (Bonev et al. 2012; Riley et al. 2013). The miRNA *mir-125a-5p* was described enriched in the PSM compared to mature somites (Riley et al. 2013). *mir-125a-5p* was shown to bind *Lfng* 3'-UTR, and not for other cyclic genes, thus modulating its half-life (Riley et al. 2013). Morpholino against *mir-125a-5p* stabilized *Lfng* transcript that resulted in an impairment of somite morphology and *Lfng* cylic expression pattern, indicating a role for this miRNA in the control of gene oscillation (Riley et al. 2013). In a different cellular context, miR-9 was shown to bind *Hes1* 3'-UTR in human cells, and to repress its expression in neural cells so modulating its mRNA levels (Bonev et al. 2012). Additional effect also on translation was not excluded by the authors since there was an even stronger decrease of the protein levels. miR9 is important for the oscillation in a certain range because *mir9* over-expression caused damped oscillations in neural progenitor cells (Bonev et al. 2012). However, conditional mutation of *Dicer*, a key factor of miRNA biogenesis, in the PSM by using the *T* (*Brachyury*)-*Cre* mouse line, demonstrated that *Dicer* is dispensable for somite segmentation (Zhang et al. 2011). *Fgf8* and *Lfng* conserved their expression pattern in the *T-Cre;Dicer*$^{flox/flox}$ mutant embryos, and the number of somites was not affected (Zhang et al. 2011). Segmentation was not altered, but anterior PSM exhibited apoptosis as well as in lateral plate mesoderm (LPM), which is also targeted by the *T-Cre* (Perantoni et al. 2005).

Those conclusions are in conflict with the study by Riley *et al.* (Riley et al. 2013). It is also surprising that transcripts that were targeted by miRNAs in the LPM are affected in the *T-Cre;Dicer*$^{flox/flox}$ mutant embryos. One explanation could be that miRNAs have a long half-life, from 28 hours to 211 hours (Gantier et al. 2011), that would still be present during the study by Zhang *et al.* (Zhang et al. 2011). Furthermore, in zebrafish, maternal-zygotic *dicer* mutants did not exhibit segmentation defects (Giraldez et al. 2005).

### 4. The role of co-activators in somitogenesis

Little is known about the role of transcriptional co-activators in somitogenesis. One example was described for the control of somite bilateral symmetry. Retinoic acid is known to be required for maintaining the somite symmetry by maintaining segmentation synchronization between left and right sides (Vermot et al. 2005). A new protein complex composed of WDR5, the deacetylases HDAC1, HDAC2 and RERE/ATROPHIN2 (named WHHERE) has been recently identified and characterized by MudPIT (Vilhais-Neto et al. 2017). This complex co-activates the expression of retinoic acid signaling pathway *in cellulo*, specifically in the presence of retinoic acid. The null mutation of *Hdac1* and the conditional deletion of *T-Cre;Wdr5*$^{fl/+}$ in the mouse embryo disrupts the somite bilateral symmetry, with animals having no somites on the right side, together with a down-regulation of the activation of the *RARE-lacZ* reporter (Retinoic Acid Response Element) (Vilhais-Neto et al. 2017). The WHHERE complex was proposed to activate genes from the retinoic acid signaling by promoting the recruitment of RNA Pol II at those promoters. Inhibition of deacetylase activity reduced the recruitment of RNA Pol II indicating that HDAC1 and HDAC2 play major roles in the activation of retinoic acid signaling pathway. Moreover, the histone methyltransferase EHMT2/G9A and GCN5 also participate to the activation of these genes, underlying the complexity of the signal integration resulting from the interplay between activators and repressors.

## VIII. Goals of the thesis project

According to the classical model of transcription initiation described previously, the basal transcription machinery was thought to be universal. This model has been challenged by evidences showing that the basal transcription machinery is more diverse than initially thought. The basal transcription machinery thus represents an additional layer of gene expression regulation. As detailed previously, in metazoan most of the data about the characterization of TFIID and SAGA composition come from *in cellulo* models and little is known about the role of these transcriptional complexes. TFIID and SAGA are two key players in transcription and several clues indicate that they also play important roles during development. Concerning TFIID, its composition is variable, and the biological significance of this variability is poorly understood.

In particular, TAF10 displays a very specific phenotype: it is differentially required according the cellular context and the developmental stage. It has been proposed that initial activation of genes require the canonical TFIID while transcription maintenance can accommodate partially assembled TFIID complexes. In order to better understand the biological function of those transcriptional complexes *in vivo*, the deletion of *Taf10* has been chosen. Interestingly, *Taf10* was identified from a microarray analysis, among other general transcription factor complexes subunits, to be differentially regulated in conditions for which the main signaling pathways were manipulated in zebrafish (unpublished data). Although this data has not been reproduced in mouse, TAF10 represented an interesting starting point for analyzing the role of general transcription factor, such as TFIID, in a developmental context.

The PSM represents an interesting paradigm for studying embryonic transcription which is very dynamic. At the onset of the project, *Taf10* was conditionally deleted by targeting the mesoderm lineage (PSM and LPM) to analyze the contribution of TFIID and SAGA in vertebrate segmentation. Moreover, the goal of my thesis was to analyze in detail their composition by using classical biochemical approaches. Given the low quantity of material available from the embryo, the project required to establish alternative approaches.

So, the main goals of my thesis were:

(1) To determine the role of TAF10-containing complexes during somitogenesis at E9.5;
(2) To develop alternative approaches in order to use classical biochemical approaches;

(3) To characterize the composition of TFIID and SAGA in the embryo and in different cellular contexts by mass spectrometry;
(4) To characterize the composition of TFIID and SAGA in the PSM and LPM;
(5) To analyze the contribution of TFIID and SAGA in RNA Pol II transcription in mammals.

# MATERIAL & METHODS

1. Mouse lines

Animal experimentation was carried out according to animal welfare regulations and guidelines of the French Ministry of Agriculture (ethical committee C2EA-17 projects 2012-077, 2012-078, 2015050509092048). Tg(T-Cre) (Perantoni et al. 2005), Tg(Hes7-Cre) (Niwa et al. 2007), Tg(Luvelu) (Aulehla et al. 2008), $R26^{CreERT2}$ (Ventura et al. 2007), $R26^{R}$ (Soriano 1999) and $Taf10^{flox}$ (Mohan et al. 2003) lines have already been described. The day of vaginal plug was scored as embryonic day (E) 0.5. Tamoxifen (Sigma) resuspended at 20 mg/ml in 5% ethanol/filtered sunflower seed oil was injected intraperitoneally [150 µl (3 mg) for a 20 g mouse] at E7.5.

2. Cell culture

a. Mouse F9 embryonal carcinoma cells

Wild-type F9 cells and 23i clone of F9 cells *Taf10*$^{-/-}$ knock-out rescued by hTAF10 full length upon doxycycline induction (Metzger et al. 1999) were cultivated on plates coated with 0.1 % gelatin in DMEM medium (4.5g/L glucose) containing 10% foetal calf serum, phenol red and 40 µg/ml gentamicin at 37 °C and 7 % CO2. 23i cells were treated with do 1000 µg/ml doxycycline.

b. Mouse embryonic stem cells

mES cells have been derived from *R26*$^{CreERT2/YFP}$*; Taf10*$^{flox/flox}$*; Tg(Mesogenin-YFP/+)* and *R26*$^{CreERT2/R}$*; Taf10*$^{flox/flox}$*; Tg(Mesogenin-YFP/+)* blastocysts, that express the YFP transgene under control of the *Mesogenin* promoter (Chal et al. 2015). Each clone was thawed on a monolayer of irradiated primary mouse embryonic fibroblasts, and then plated on gelatin for three passages prior use for further experiments. mES cells were cultivated in DMEM 4.5 g/l glucose, 15 % fetal bovine serum, penicillin, streptomycin, 4 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1 % β-mercaptoethanol, 1,500 U/ml leukemia inhibitory factor and 2i inhibitors: CHIR99021 (STEMGENT or STEMCELLS), a glycogen synthase kinase 3 beta inhibitor, and PD0325901 (STEMGENT or STEMCELLS), a phosphorylation and activation of MAPK/ERK inhibitor, at 37 °C and 5 % CO2.

c. T-cell leukemia cell line T29

The cell line was described in (Dumortier et al. 2006). T29 were grown in suspension in RPMI 1640 w/o HEPES, 10% fetal calf serum and 40 µg/ml gentamycine.

3. Cell count and cell death assay

a. Cell count and viability assay

mES cells were trypsinized for 5 minutes and were stained with acridine orange and propidium iodide which stains only the altered membrane of dead cells. Fluorescence is measured and analyzed by the counter Nexcellcom. Cells stained with both acridine orange and propidium iodide fluoresce red due to quenching, so all live nucleated cells fluoresce green and all dead nucleated cells fluoresce red.

A fraction of each F9 cellular suspension was mixed with an equal volume of 0.4 % Trypan blue and loaded onto a slide which was analyzed by the Countess II Automated Cell Counter (ThermoFischer Scientific).

b. Apoptosis assay

Apoptosis was examined using the Cell-APOPercentage™ Apoptosis Assay (Biocolor) following the manufacturer's instructions. Cells were seeded on a 24 well plate with 10,000 cells/cm². As a positive control, apoptosis was induced with 10 mM hydrogen peroxide for 5 hours and 30 minutes in untreated cells. Control and mutant cells were incubated for 30 minutes with 5 % of the dye at 37 °C and 5 % of CO2. The dye selectively stains apoptotic cells which display a translocation of phosphatidylserine from the interior to the exterior surface of the membrane. The supernatant was collected and cells were washed twice with 1 X phosphate buffer saline (PBS) prior to be harvested by trypsinization for 10 minutes. Cells were washed twice with 1 X PBS. After centrifugation at 153 rcf, the pellet was resuspended with the dye release solution. Absorbance was read at 550 nm and normalized to the blank control (with culture medium only).

4. Proliferation assay

Cells were seeded at 22,222 cells/cm² and incubated for 1 h at 37 °C, 5 % CO2 with 10 µM 5-ethynyl-2-deoxyuridine (EdU) from the Click-iT EdU Alexa Fluor 488 Flow Cytometry Assay Kit (ThermoFischer Scientific). Cells were harvested by trypsinization and washed with 1 % BSA in 1 X PBS. Cells were pelleted after 153 rcf centrifugation at room temperature. Cells were fixed with the Clik-it fixative (ThermoFischer Scientific) for 15 min protected from light and were washed with 1 % BSA in 1 X PBS. Cells were permeabilized for 15 min at room temperature with 1 X click-it saponin based permeabilization buffer (ThermoFischer Scientific). The click-it reaction cocktail containing 1 X PBS, CuSO4, the fluorescent dye azide and the reaction buffer additive was added to the cellular suspension for 30 min at room temperature protected from light. Cells were washed with 1 X click-it saponin-based permeabilization buffer and centrifugated at 153 rcf for 5 min. The cellular pellet was resuspended in 50-100 µl of 1 X click-it saponin-based permeabilization buffer with 50µg/ml of propidium iodide prior analysis by flow cytometry with the BD FACSCELESTA (BD Biosciences) with a detection wavelength of 488 nm. A total of 20,000 events were recorded for the analysis.

5. Cellular extracts

a. Acidic extracts

Cells were washed twice with cold 1 X PBS and were harvested into cold 1x 1XPBS on ice. They were centrifugated at 1,467 rcf for 5 min at 4 °C and the pellet was resuspended with five times the packed cell volume of lysis buffer containing 10 mM hepes pH 7.9, 1.5 mM MgCl2, 0.5mM dithiothreitol (DTT) supplemented with protease inhibitor cocktail. 0.2 M HCl was added to the cell lysate and was incubated for 30 min on ice. The acid was neutralized with five times of 1 M Tris pH 7.9. The acidic extracts were used directly for western-blots.

b. Whole Cell Extracts

i. Large scale whole cell extract preparation

5 liters of T29 cells were centrifugated for 20 min at 363 rcf at room temperature. The pellet was resuspended with ten ml of 1 X PBS per ml of packed cell volume. The cellular suspension was centrifugated for 10 min at 344 rcf. The pellet was resuspended with four times the packed

cell volume of a solution containing 10 mM tris pH 7.8, 1 mM EDTA (Ethylene Diamine Tetra Acetic Acid), 0.5 M DTT supplemented with protease inhibitor cocktail (Roche) and incubated for 30 min on ice. The cells were lysed by 20 strokes with a pestle B in a glass dounce tissue grinder. Four times the packed cell volume of a solution containing 50 mM Tris pH 7.8, 10 mM MgCl2, 25 % sucrose, 50 % glycerol, 2 mM DTT supplemented with protease inhibitor cocktail. One time the packed cell volume of 4 M ammonium sulfate was added drop by drop for one hour at 4 °C under agitation. The precipitate mixture was ultra-centrifugated at 49 K for two hours and 30 min at 4 °C. The supernatants were collected and 0.3 g of ammonium sulfate powder per ml of supernatant was added progressively for one hour at 4 °C under agitation. 1 M NaOH was added to the solution and was incubated under agitation for 30 min at 4 °C. The mixture was ultra-centrifugated at 30 K for 30 min at 4 °C. The new packed cell volume of the pellet was measured. The pellet was resuspended with 0.1 time the new packed cell volume with a solution containing 50 mM Tris pH 7.8, 10 % glycerol, 0.5 mM EDTA, 50 mM KCl, 0.5 mM DTT supplemented with protease inhibitor cocktail. The mixture obtained were dialyzed overnight at 4 °C under agitation in the same buffer solution supplemented with freshly prepared phenylmethylsulfonyl fluoride. The mixture was centrifugated at 20,817 rcf at 4 °C for 30 min and the supernatant was collected and stored at −80°C for immuno-precipitations western blots.

ii. Small scale whole cell extract preparation

Cells were washed twice with 1 X PBS and were harvested by scrapping and centrifuged at 20,817 rcf at 4°C. Lysates were obtained after pipette breakdown in 10 % glycerol, 20 mM Hepes pH 7, 0.35 M NaCl (Sodium Chloride), 1.5 mM MgCl2 (Magnesium Chloride), 0.2 mM EDTA, 0.1% Triton X-100 with protease inhibitor cocktail on ice, and followed by three liquid nitrogen freeze-thaw cycles. Lysates were centrifuged at 20,817 rcf for 15 min at 4 °C and the supernatants were used directly for immuno-precipitations or stored at −80°C for western blots. E9.5 mouse embryos (16-20 somites) were used for lysates following the same protocol, but lysis was performed after three times pestle stroke treatment on ice.

c. Nuclear extracts

Cells were washed twice with 1X PBS and were harvested by scrapping and centrifuged at 20,817 rcf at 4°C. The cellular pellet was resuspended with an equal volume of 1x hypotonic

buffer, containing 20 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2 with protease inhibitor cocktail, by pipetting up and down several times and incubated on ice for 15 min. 10 % NP40 was added at $1/20^{th}$ of the pellet resuspension volume to the mixture to lyse the cellular membrane followed by 10 seconds vortex at higher settings. The homogenate was centrifuged for 10 min at 4,460 rcf at 4 °C. The supernatant corresponding to the cytoplasmic fraction was frozen in liquid nitrogen and stored at -80°C. The pellet was resuspended into 50 ul of complete cell extraction buffer that contain 10 mM Tris, pH 7.4, 2 mM Na3VO4, 100 mM NaCl, 1 % Triton X-100, 1 mM EDTA, 10 % glycerol, 1 mM EGTA , 0.1 % SDS, 1 mM NaF, 0.5 % deoxycholate and 20 mM Na4P2O7 with protease inhibitor cocktail. The lysis mixture was incubated 30 min on ice and vortexed every 10 min. Lysates were centrifuged for 30 min at 20,817 rcf at 4°C. The supernatant corresponding to the nuclear fraction was frozen in liquid nitrogen and stored at -80 °C for gel filtration.

6. Bradford protein assay

One microliter of a cellular extract was mixed with the dye reagent (Biorad protein assay, Biorad) diluted five times $1/5^{th}$, the absorbance was read at 595 nm. The concentration of the sample was determined from the standard curve which was calculated from six standard samples with a known concentration of bovine serum albumin protein.

7. Immuno-precipitation

700 ug of whole cell extract from whole embryos or 4 mg of whole cell extracts from mES cells were incubated with Protein A (for polyclonal antibodies) or Protein G (for monoclonal antibodies) Sepharose beads, or Dynabeads coated with antibodies overnight at 4°C. Immunoprecipitated proteins were washed twice for 5 min each with 500 mM KCl buffer [25 mM Tris-HCl (pH 7), 5 mM MgCl2, 10 % glycerol, 0.1 % NP40, 2 mM DTT, 500 mM KCl and protease inhibitor cocktail], then washed twice for 5 min each with 100 mM KCl buffer [25 mM Tris-HCl pH 7, 5 mM MgCl2, 10 % glycerol, 0.1 % NP40, 2 mM DTT, 100 mM KCl and protease inhibitor cocktail] and eluted with 0.1 M glycine pH 2.8 three times for 5 min each. Elution fractions were neutralized with 1.5 M Tris-HCl pH 8.8 and analyzed by western-blot and/or mass spectrometry.

## 8. Western-blot

Protein extracts (15 ug -25 ug) were boiled 10 min in 100 mM Tris pH 6.8, 30 % glycerol, 4 % sodium dodecyl sulfate (SDS), 0.2 % bromo phenol blue and fresh 100 mM DTT resolved on a home-made 10 % acrylamide gel or a precast SDS-polyacrylamide gel 4-12% (Novex) and transferred to nitrocellulose membrane (Protran, Amersham). Samples migrated under 150 V and were blotted on a nitrocellulose membrane (Protran, Amersham) for 65 min at 250 mA at 4°C under gentle agitation. The nitrocellulose membrane was incubated for 20 min in 1 X PBS 3 % of non-fat milk. Primary antibodies were diluted 1/1,000 times and were incubated overnight at 4 °C with the membrane in 1 X PBS and 0.3 % of non-fat milk. Membrane was washed three times for 5 min each with 0.05% Tween 20 in PBS and incubated with the secondary antibody (goat anti mouse or goat anti rabbit) coupled to the horseradish peroxidase was diluted 1/10,000 times and incubated for 50 min at room temperature. Membrane was washed three times for 5 min each with 0.05% Tween 20 in PBS and was incubated for one minute with the HRP substrate ECL (ThermoFisher Scientific). The signal was acquired with the Chemidoc imaging system (Biorad).

## 9. Antibodies

| **Antibody** | **Type** | **Reference** |
|---|---|---|
| Anti-TAF1 | Rabbit polyclonal | 2439, 2440 |
| Anti-TAF2 | Rabbit polyclonal | 3038 (Trowitzsch et al. 2015a) |
| Anti-TAF3 | Mouse monoclonal | 2F5 (Gangloff et al. 2001) |
| Anti-TAF4 | Mouse monoclonal | 32TA2B9 (Perletti et al. 2001) |
| Anti-TAF5 | Mouse monoclonal | 1TA1C2 (Jacq et al. 1994) |
| Anti-TAF6 | Mouse monoclonal | 22TA2A1 (Wieczorek et al. 1998) |
| Anti-TAF7 | Rabbit polyclonal | 3475 (Bardot et al. 2017) |
| Anti-TAF8 | Rabbit polyclonal | 3477, 3478 (Bardot et al. 2017) |

| | | |
|---|---|---|
| Anti-TAF10 | Mouse monoclonal | 6TA2B11 (Mohan et al. 2003) |
| Anti-Nterm-TAF10 | Mouse monoclonal | 23TA1H8 (Wieczorek et al. 1998) |
| Anti-TBP | Mouse monoclonal | 3TF13G3(Brou et al. 1993)s |
| Anti-ATXN7L3 | Mouse monoclonal | 1ATX2D7 (Zhao et al. 2008) |
| Anti-GCN5 | Mouse monoclonal | 5GC2A6 (Orpinell et al. 2010) |
| Anti-SPT20 | Rabbit polyclonal | 3006 (Krebs et al. 2011) |
| Anti-SUPT3H | Rabbit polyclonal | 3118 (Bardot et al. 2017) |
| Anti-TRRAP | Mouse monoclonal | 2TRR1B3 (Helmlinger et al. 2004) |
| anti-Rabbit IgG Peroxydase conjugate | Goat polyclonal | Jackson ImmunoResearch 111-035- 144 |
| anti-Mouse IgG Peroxydase conjugate | Goat polyclonal | Jackson ImmunoResearch 111-036- 071 |

10. Mass spectrometry

Samples were analyzed using an Ultimate 3000 nano-RSLC (ThermoFischer Scientific) coupled in line with an linear trap Quadrupole (LTQ)-Orbitrap ELITE mass spectrometer via a nano-electrospray ionization source (ThermoFischer Scientific). Peptide mixtures were loaded on a C18 Acclaim PepMap100 trap column (75 µm inner diameter × 2 cm, 3 µm, 100 Å; ThermoFischer Scientific) for 3.5 min at 5 µl/min with 2% acetonitrile (ACN), 0.1% formic acid in $H_2O$ and then separated on a C18 Accucore nano-column (75 µm inner diameter × 50 cm, 2.6 µm, 150 Å; ThermoFischer Scientific) with a 120-min linear gradient from 5 % to 50 % buffer B (A: 0.1 % FA in $H_2O$; B: 80 % ACN, 0.08 % FA in $H_2O$) followed with 10 min at 99 % B. The total duration was set to 150 min at a flow rate of 200 nl/min. The temperature was kept constant at 40 °C. Peptides were analyzed by Top 10-CID-HCD (collision induced dissociation and high energy collision dissociation) data-dependent mass spectrometry.

11. Gel filtration

A Superose 6 (10/300) column was equilibrated with 25 mM Tris HCl pH 7,9, 150 mM KCl, 5 mM MgCl2, 5 % Glycérol, 1 mM DTT, 1 tablet of protease inhibitor cocktail. 500 ul of nuclear extracts prepared as described above containing 4 mg of protein were injected in an Akta Avant chromatography device (GE Healthcare) and run at 0.4 ml $min^{-1}$. Protein detection was performed with absorbance at 280 nm and 260nm. 500 ul fractions were collected and analyzed by western blot.

## 12. 4sU metabolic labeling of newly-synthesized mRNA and purification

### a. 4sU labeling

The workflow of the experiment is described in (figure 14). mES cells cultivated on gelatin of 10 cm Petri dish and Schneider cells (S2) cultivated on 10 cm Petri dish were incubated separately for 10 min with 500 uM 4sU (Sigma), prepared at 100 mM in dimethylsulfoxyde, in 2i+LIF (Leukemia Inhibitory Factor) medium. Cells were harvested in 1 ml Trizol (ThermoFischer Scientific) per 10 cm Petri dish from mES and S2 cells, and were mixed together prior total RNA isolation. Samples were frozen and stored at -80°C until extraction.

### b. Total RNA extraction and isolation

For total RNA isolation, chloroform was added at $1/5^{th}$ of the total volume and were vortexed 15 sec before centrifugation at 4°C at 20,817 rcf for 15 min. The aqueous phase was collected in a fresh tube and was centrifugated at 4°C at 20,817 rcf for 15 min. The upper phase was collected in a fresh tube to which was added an equal volume of cold isopropanol and 15 ug of GlycoBlue (ThermoFisher Scientific). Tubes were inverted four times and were incubated at room temperature for 10 min. Then the tubes were centrifugated at 4 °C at 20,817 rcf for 15 min. The supernatant was discarded, and the pellet was washed with 70 % EtOH. The pellet was recovered after centrifugation at 4 °C at 20,817 rcf for 5 min and resuspended into DEPC H20. For better resuspension, the RNA pellet was warmed at 65 °C for 10 min.

### c. Purification of newly-synthesized mRNA

Biotinylation of 4sU-labeled RNA was performed using EZ-Link Biotin-HPDP (Pierce) dissolved in dimethylsulfoxyde at a concentration of 1 mg/mL. 100 ug of total RNA was heated

10 min at 65 °C and immediately chilled on ice for 5 min prior biotinylation. Biotinylation was carried out in 1 ml final containing 100 ul biotinylation buffer [1 M Tris-HCl (pH 7.5), 0.5 mM EDTA], 0.2 mg/mL Biotin-HPDP and 200 ul of additional DMSO for better HPDP-biotin solubility for 3 h at room temperature protected from light. Unbound Biotin-HPDP was removed by chloroform/isoamylalcohol (24:1) extraction. Afterward, a $1/10^{th}$ volume of 5 M NaCl and an equal volume of isopropanol were added and RNA was precipitated at 20,000 rcf for 20 min. The pellet was washed with an equal volume of 75 % ethanol and precipitated again at 20,000 rcf for 10 min. The pellet was resuspended in 100 µl RNAse-free water. After denaturation of RNA samples at 65 °C for 10 min followed by rapid cooling on ice for 5 min, biotinylated RNA was incubated with 100 µl of µMACS streptavidin beads (Miltenyi) with rotation for 90 min at room temperature. µMACS columns (Miltenyi) were equilibrated with 900 ul of washing buffer [100 mM Tris-HCl (pH 7.5), 10 mM EDTA, 1 mM NaCl, 0.1% Tween 20 in DEPC water]. Two replicates of the biotinylated RNA/streptavidin beads samples were loaded per column. Columns were washed five times with increasing volumes, 600 ul, 700 ul, 800 ul, 900 ul and 1 ml of washing buffer. Labeled RNA was eluted by the addition of 100 µl of freshly prepared 100 mM DTT followed by a second elution round 5 min later. RNA was recovered from the washing fractions and eluates using the RNeasy MinElute Spin columns (Qiagen).

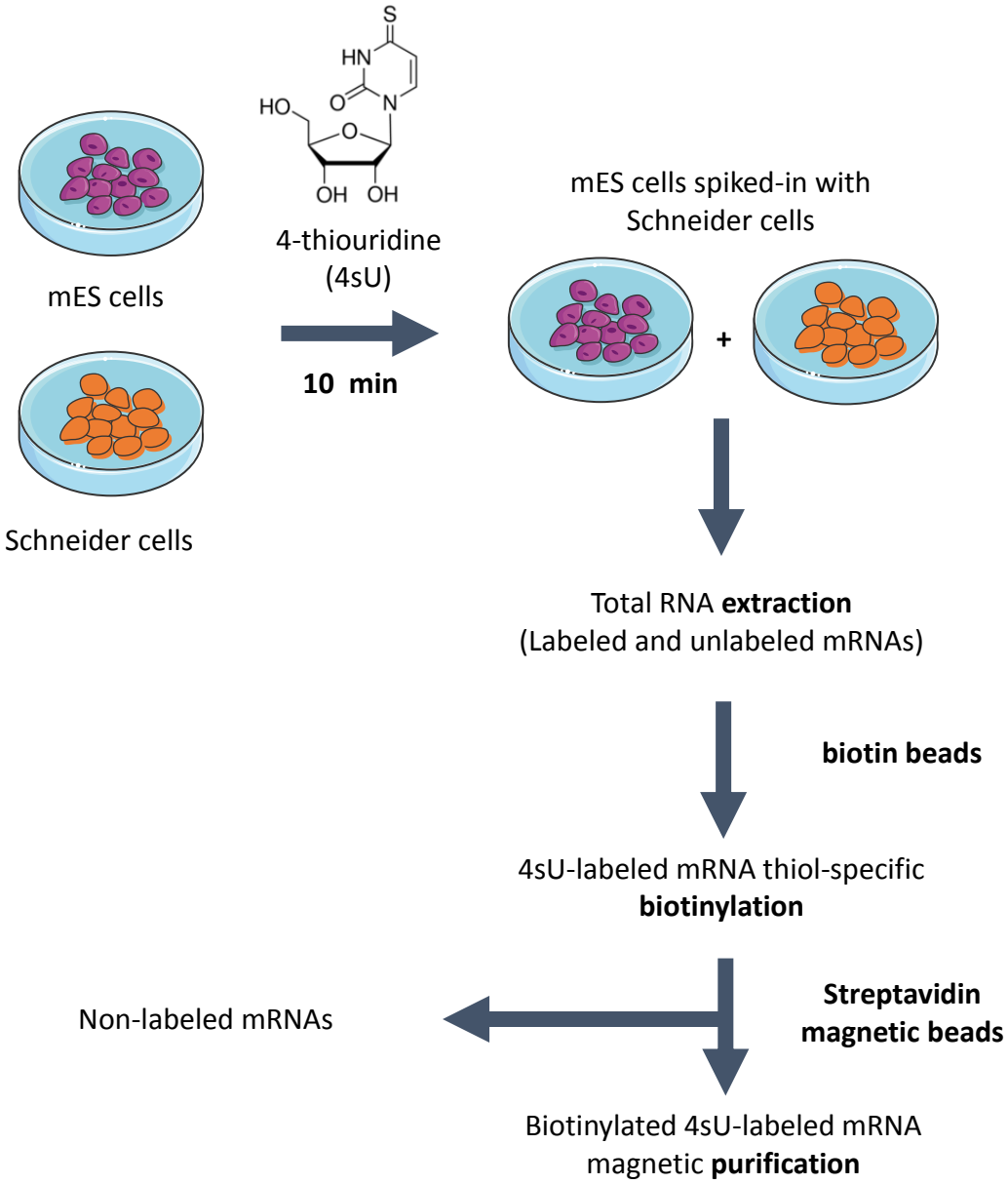



**Figure 14: Workflow of 4sU metabolic labeling of newly-synthesized mRNAs**. mES cells or Schneider cells (S2) are incubated for 10 minutes with 4sU, prior total mRNA extraction. S2 cells serve as a spike-in control, added with a ratio (mES:S2) between 1:1 and 3:1. Labeled mRNAs are first thiol-specifically biotinylated, and then incubated with streptavidin beads which react with biotinylated labeled mRNAs. Unlabeled and labeled mRNAs can be separated when applied to a magnetic column that retain only labeled mRNAs. The labeled mRNAs are eluted with a reducing agent which cleaves the disulfide bonds that link the newly transcribed RNA to the beads.

13. RNA fragmentation

250 µg of total RNA were added with 130µl ultrapure DEPC-H2O (Sigma) in Snap Cap tubes (Covaris) and were sonicated in the Covaris E220 for 80 sec with 1 % duty factor, 100 W, 200 cycles per burst.

14. RT-qPCR

RNAs were reverse transcribed in cDNA by using the Quantitect Reverse Transcription kit (Qiagen). It includes a genomic DNA elimination step. The qPCR reaction was performed with the SybrGreen Mastermix Class I kit (Roche), following the manufacturer guidelines by using 0.5 µM of primers in 384 wells plates and analyzed with the Light Cycler 480 (Roche). The amplification programme consisted in 45 cycles of: denaturation at 95 °C, hybridization at 65 °C and elongation at 72 °C followed by a melting curve from 65 °C to 97 °C and a cooling down of 30 sec at 40 °C.

15. Gene primers

| Gene | Primer sequence (5'-3') |
|---|---|
| *alphaTUB84B* | Forward : GCTTCCTCATCTTCCACTCG |
| | Reverse : GCTTGGACTTCTTGCCGTAG |
| *actin42A* | Forward : GCGTCGGTCAATTCAATCTT |
| | Reverse : AAGCTGCAACCTCTTCGTCA |
| *Ccnb2* | QuantiTect Primer Assay QT00112588 |
| *Ccne2* | QuantiTect Primer Assay QT00129206 |
| *Ccng1* | Forward: AAGCAGCTCAGTCCAACACA |
| | Reverse: CCTTTCAGTCCGCTCCAGAA |
| *Cdk4* | QuantiTect Primer Assay QT00103292 |
| *Cdkn1a* | Forward: CTGAGCGGCCTGAAGATTCC |
| | Reverse: CCAATCTGCGCTTGGAGTGA |
| *Cdkn1b* | Forward: TCAAACGTGAGAGTGTCTAACG |
| | Reverse: CCGGGCCGAAGAGATTTCTG |

| | |
|---|---|
| *Cdkn1c* | Forward: CAGCCTCTCTCGGGGATTC |
| | Reverse: CTCCTGCGCAGTTCTCTTGC |
| *Cdkn2a* | QuantiTect Primer Assay QT00252595 |
| *Cdkn2c* | QuantiTect Primer Assay QT01055005 |
| *Gapdh* | QuantiTect Primer Assay QT01658692 |
| *Gas5* | Forward: GGAAGCTGGATAACAGAGCGA |
| | Reverse: GTATTCCTTGTAATGGGACCAC |
| *Klf4* | Forward: CCAGCAAGTCAGCTTGTGAA |
| | Reverse: GGGCATGTTCAAGTTGGATT |
| *Nanog* | Forward: TTGCTTACAAGGGTCTGCTACT |
| | Reverse: ACTGGTAGAAGAATCAGGGCT |
| *Oct4* | Forward: AGAGGATCACCTTGGGGTACA |
| | Reverse: CGAAGCGACAGATGGTGGTC |
| *Rn45s* | Forward: TGGGTTTTAAGCAGGAGGTG |
| | Reverse: ACGCTTGGTGAATTCTGCTT |
| *Rnu6* | Forward: GCTTCGGCAGCACATATACTAA |
| | Reverse: AAATATGGAACGCTTCACGAAT |
| *Rplp0* | QuantiTect Primer Assay |
| *Taf1d* | Forward: GCGAGATTTCGTCCTTGTCCT |
| | Reverse: CTCCAGCTCTATCCGATGCC |
| *Taf8* | QuantiTect Primer Assay QT00156723 |
| *Tbp* | QuantiTect Primer Assay QT00198443 |

# RESULTS

# IX. Publication: "The TAF10-containing TFIID and SAGA transcriptional complexes are dispensable for early somitogenesis in the mouse embryo" - (Bardot et al. 2017)

Context of the study:

Transcription plays major roles during embryonic development. TFIID is the first GTF to bind the promoter and nucleates the assembly of the PIC for RNA Pol II recruitment. Several evidences have shown that the PIC components might be variable [reviewed in (Levine et al. 2014; Müller et al. 2009)]. In particular, TAF10 is crucial for survival and proliferation of F9 cells, but is not required for their differentiation into primitive endoderm (Metzger et al. 1999). *Taf10* null mutation in mouse leads to embryonic lethality shortly after implantation (Mohan et al., 2003). However, whereas inner cell mass cells die by apoptosis, trophoectodermal cells survive, although RNA Pol II transcription is greatly reduced (Mohan et al., 2003). *Taf10* conditional deletion in skin or liver has shown that TAF10 is required for transcription in the embryo, but not in the adult (Indra et al., 2005; Tatarakis et al., 2008). Contrary to the idea of the general requirement of GTF, altogether, this data indicate that the canonical TFIID requirement depends on the cellular and developmental context. Most of the data concerning the composition of TFIID and SAGA in metazoan has been obtained from cellular models, including cancer cells, which are artificial. Thus, *in vivo* data in metazoans is lacking, and no data has been ever obtained concerning the precise composition of the SAGA complex in the embryo. In order to analyze the composition and the function of the TAF10-containing complexes, TFIID and SAGA, we used the paraxial mesoderm as a new paradigm. Somitogenesis that takes places in this tissue, is a very dynamic morphogenetic event and gene expression is tightly regulated. Since transcription is highly dynamic it represents an interesting model for analyzing embryonic transcription. We used a conditional deletion strategy for inactivating *Taf10* by using the *T-Cre;Taf10*$^{flox/flox}$ mouse line which drives the expression of the *Cre* in the embryo as early as E7.5.

Main goals:

The goal of this study was:

(1) To analyze the composition of TAF10-containing complexes in the embryo;

(2) To analyze the function of TAF10-containing complexes at E9.5 in the PSM;

(3) To assess the importance of TAF10 for gene expression in the PSM at 9.5.

Main results:

We detected all the TAF subunits, except TAF7L, and we showed that TAF10, is essential for the assembly of TFIID and SAGA complexes at E9.5 in the whole embryo. Deletion of *Taf10* in the whole embryo or in the embryonic mesoderm at E7.5 induces a growth arrest of the embryo at E10. Initial paraxial mesoderm differentiation is not prevented in the conditional mutant *T-Cre;Taf10*$^{flox/flox}$ but they do not have a forelimb at E9.5 and lateral plate differentiation is altered. During this period, in the absence of detectable TAF10 protein, steady-state mRNA levels are unchanged in the PSM, except for only a minor subset of genes dysregulated. Among those genes, cyclin inhibitors were found up-regulated. Interestingly, the cyclic gene expression pattern reporter of *Lfng* is conserved in the mutant at E9.5, indicating that transcription is still active. Our data suggests that the TFIID and SAGA complexes are dispensable for early paraxial mesoderm development, arguing against their proposed generic role in transcription.

These results were accepted for publication on 2nd September 2017 in *Development*:

**"The TAF10-Containing TFIID and SAGA Transcriptional Complexes Are Dispensable for Early Somitogenesis in the Mouse Embryo."** Bardot, Paul, Stéphane D Vincent, Marjorie Fournier, Alexis Hubaud, Mathilde Joint, László Tora, and Olivier Pourquié.

© 2017. Published by The Company of Biologists Ltd | Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

RESEARCH ARTICLE

# The TAF10-containing TFIID and SAGA transcriptional complexes are dispensable for early somitogenesis in the mouse embryo

**Paul Bardot[1,2,3,4,§], Stéphane D. Vincent[1,2,3,4,§,**], Marjorie Fournier[1,2,3,4,*,¶], Alexis Hubaud[1,2,3,4,‡,¶], Mathilde Joint[1,2,3,4], László Tora[1,2,3,4] and Olivier Pourquié[1,2,3,4,‡]**

## ABSTRACT

During development, tightly regulated gene expression programs control cell fate and patterning. A key regulatory step in eukaryotic transcription is the assembly of the pre-initiation complex (PIC) at promoters. PIC assembly has mainly been studied *in vitro*, and little is known about its composition during development. *In vitro* data suggest that TFIID is the general transcription factor that nucleates PIC formation at promoters. Here we show that TAF10, a subunit of TFIID and of the transcriptional co-activator SAGA, is required for the assembly of these complexes in the mouse embryo. We performed *Taf10* conditional deletions during mesoderm development and show that *Taf10* loss in the presomitic mesoderm (PSM) does not prevent cyclic gene transcription or PSM segmental patterning, whereas lateral plate differentiation is profoundly altered. During this period, global mRNA levels are unchanged in the PSM, with only a minor subset of genes dysregulated. Together, our data strongly suggest that the TAF10-containing canonical TFIID and SAGA complexes are dispensable for early paraxial mesoderm development, arguing against the generic role in transcription proposed for these fully assembled holo-complexes.

**KEY WORDS: RNA polymerase II, TATA binding protein, Presomitic mesoderm, Paraxial mesoderm, Conditional knockout, Proteomic, Mouse**

## INTRODUCTION

In mouse, the posterior part of the paraxial mesoderm, called presomitic mesoderm (PSM), generates a pair of somites every 2 h and plays crucial roles during vertebrate elongation (Pourquié, 2011). This rhythmic process is under the control of a clock that is characterized by periodic waves of transcription of cyclic genes sweeping from the posterior to the anterior PSM (Hubaud and Pourquié, 2014). In the anterior PSM, the clock signal is converted into a stripe of expression of specific segmentation genes that defines the future somite. This periodic transcription initiation associated with the segmentation clock oscillations in the PSM offers a unique paradigm with which to study transcriptional regulation in development.

During embryogenesis, gene expression is regulated by a combination of extracellular signals triggering intracellular pathways, which converge towards the binding of transcription factors to enhancers and promoters. These interactions lead to the assembly of the transcriptional machinery. In non-plant eukaryotes, three RNA polymerases are able to transcribe the genome, among which RNA polymerase II (Pol II) is responsible for the production of mRNA and some of the non-coding RNAs (Levine et al., 2014 and references therein).

Transcription initiation requires the assembly of the pre-initiation complex (PIC) that allows the correct positioning of Pol II on the promoter and consequent RNA synthesis (Sainsbury et al., 2015). TFIID is the first element of the PIC recruited to active promoters. In its canonical form in higher eukaryotes it is composed of TATA binding protein (TBP) and 13 TBP-associated factors (TAFs) and is involved in the correct positioning of Pol II on the transcription start site. Whereas TBP is also part of Pol I and Pol III transcription complexes, the TFIID-TAFs are specific for Pol II transcription machinery. Among the metazoan TAFs, TAF9, TAF10 and TAF12 are also shared by Spt-Ada-Gcn5-acetyl transferase (SAGA) complex, which is a transcriptional co-activator conserved from yeast to human (Spedale et al., 2012). SAGA exhibits histone acetyltransferase activity at promoters and also deubiquitylates histone H2Bub1 in gene bodies (Bonnet et al., 2014; Wang and Dent, 2014; Weake et al., 2011).

Several structural TAFs, including TAF10, share a histone fold domain (HFD) which is involved in their dimerization with specific partners: TAF10 heterodimerizes with TAF3 or TAF8 within TFIID and with SUPT7L/ST65G within SAGA (Leurent et al., 2002; Soutoglou et al., 2005). Nuclear import of TAF10 is absolutely dependent on heterodimerization with its partners since TAF10 does not have a nuclear localization signal (NLS) (Soutoglou et al., 2005).

TAF10 does not exhibit any enzymatic activity but serves as an interface allowing interaction with other TAFs (Bieniossek et al., 2013; Trowitzsch et al., 2015) or transcription factors, such as the human estrogen receptor α (Jacq et al., 1994) or mouse GATA1 (Papadopoulos et al., 2015). In HeLa cells, only 50% of the TFIID complexes contain TAF10 (Jacq et al., 1994). TFIID complexes lacking TAF10 have also been observed in mouse F9 cells although at much lower level (Mohan et al., 2003), but their functionality is unknown. The structure of TFIID in the absence of TAF10 is unclear. Only partial TFIID subcomplexes, not associated with TBP, were detected in undifferentiated and retinoic acid (RA)-differentiated *Taf10* mutant F9 cells (Mohan et al., 2003), while lack

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch 67400, France. [2]Centre National de la Recherche Scientifique, UMR7104, Illkirch 67400, France. [3]Institut National de la Santé et de la Recherche Médicale, U964, Illkirch 67400, France. [4]Université de Strasbourg, Illkirch 67400, France.
*Present address: Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK. ‡Present address: Brigham and Woman's Hospital, Harvard Medical School, Boston, MA 02115, USA.
§These authors contributed equally to this work
¶These authors contributed equally to this work

**Author for correspondence (vincent@igbmc.fr)

S.D.V., 0000-0003-1638-9615; M.F., 0000-0002-5202-8262; A.H., 0000-0003-1249-4282; M.J., 0000-0002-6497-4328; L.T., 0000-0001-7398-2250

**Received 10 November 2016; Accepted 2 September 2017**

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

of TFIID was observed in *Taf10* mutant liver cells (Tatarakis et al., 2008). SAGA was not investigated in these experiments (Mohan et al., 2003; Tatarakis et al., 2008). Altogether, these data support the idea that TFIID composition can vary, as also suggested by the existence of TAF paralogs and/or tissue-specific TAFs (Goodrich and Tjian, 2010; Müller et al., 2010).

The diversity in TFIID composition may have functional consequences. Whereas TAF10 is crucial for survival and proliferation of F9 cells, it is dispensable for their differentiation into primitive endoderm (Metzger et al., 1999). *Taf10* mutation in mouse leads to embryonic lethality shortly after implantation (Mohan et al., 2003). Interestingly, whereas inner cell mass cells die by apoptosis, trophoectodermal cells survive, although Pol II transcription is greatly reduced (Mohan et al., 2003). *Taf10* conditional deletion in skin or liver has shown that TAF10 is required for transcription in the embryo, but not in the adult (Indra et al., 2005; Tatarakis et al., 2008). Altogether, these data indicate that TAF10 requirement depends on the cellular and developmental context.

In this study, we aimed to closely analyze TAF10 requirement and its role in transcription during mouse development, and to examine the composition of TFIID and SAGA in the absence of TAF10 in embryonic tissues *in vivo*. We performed immunoprecipitations coupled to mass spectrometry analyses on embryonic lysates. We show that, in the mouse embryo, absence of TAF10 severely impairs TFIID and SAGA assembly. In order to gain insight into the functional importance of TAF10 during development, we focused on paraxial mesoderm dynamic differentiation by carrying out a *Taf10* conditional deletion in the mesoderm using the *T-Cre* line (Perantoni, 2005). Although loss of *Taf10* eventually led to growth arrest and cell death at ~E10.5, we identified a time window during which the dynamic transcription of cyclic genes is still maintained in the absence of detectable TAF10 protein. Microarray analysis of mutant PSM revealed that Pol II transcription is not globally affected in this context, although the expression of some genes, such as those encoding cell cycle inhibitors, is upregulated.

## RESULTS

### TAF10 is ubiquitously expressed in the nucleus of embryonic cells at E9.5

*Taf10* is ubiquitously expressed in the mouse embryo at E3.5, E5.5 and E7.5 but with more heterogeneity at E12.5 (Mohan et al., 2003). Whole-mount *in situ* hybridization (WISH) analyses suggest that *Taf10* is also ubiquitously expressed at E8.5 and E9.5 (Fig. S1A,B). TAF10 protein is ubiquitously expressed in the posterior part of the embryo (Fig. S1C, Fig. S2) and no heterogeneity was observed between E9.5 and E10.5. Competition with the peptide used to raise the anti-TAF10 antibody (Mohan et al., 2003) confirms that TAF10 localization is specific, since the TAF10 signal, but not the myogenin signal, is lost under these conditions (Fig. S1D,H). Altogether, these results indicate that TAF10 protein is ubiquitously expressed in nuclei between E8.5 and E10.5.

### Induced ubiquitous deletion of *Taf10* leads to growth arrest at E10, but does not impair transcription at E9.5

In order to analyze the effects of TAF10 absence on development, we performed a tamoxifen-inducible ubiquitous deletion of *Taf10* using the $R26^{CreERT2}$ line (Ventura et al., 2007). This strategy deletes exon 2 of *Taf10*, which encodes part of the HFD (Mohan et al., 2003), and because exon 3 is now out of frame the deletion is expected to produce a truncated protein of 92 amino acids without an HFD (Fig. S3D). Since the HFD is required for heterodimerization and integration of TAF10 into TFIID and SAGA (Leurent et al., 2002; Soutoglou et al., 2005), this potential truncated protein is not supposed to integrate into mature SAGA or TFIID complexes. Tamoxifen was injected intraperitoneally at E7.5 and Cre recombination was followed by the activity of the Cre reporter allele $R26^{R}$ (Soriano, 1999). Complete Cre recombination is observed at E9.5 (Fig. 1A,B). The development of $R26^{CreERT2/+}$;$Taf10^{flox/flox}$ (*R26Cre;Taf10*) mutant embryos was arrested at E9.5, as embryos do not further develop when recovered at E10.5 and E11.5 (Fig. 1D,F). The normal development of $R26^{R/+}$;$Taf10^{flox/flox}$ littermate embryos (Fig. 1C,E) confirmed that tamoxifen injection at E7.5 does not induce secondary defects.

Efficient TAF10 depletion at E9.5 after tamoxifen injection at E7.5 was assessed by western blot (Fig. 1G). At E8.5 TAF10 was still present, albeit at lower levels (Fig. S3E). This observation is in agreement with a previous study in which TAF10 protein was still detected one day after induction of its depletion (Metzger et al., 1999). Since our goal is to assess TFIID and SAGA composition in the absence of TAF10, we performed our analyses at E9.5.

In order to assess transcription initiation *in vivo*, we used the *Luvelu* reporter line (Aulehla et al., 2008) that allows visualization of the dynamic waves of *Lfng* transcription occurring every 2 h in the posterior PSM. This line contains the promoter and 3′-UTR destabilizing sequences of the cyclic gene *Lfng* (Cole et al., 2002; Morales et al., 2002), coupled to the coding sequences of a

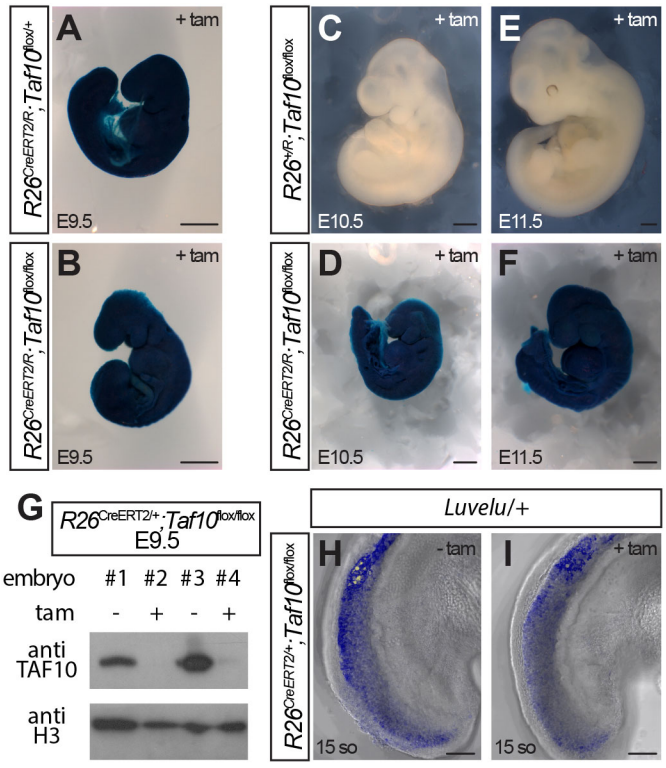



**Fig. 1. Efficient ubiquitous deletion of *Taf10* in E9.5 *R26Cre;Taf10* mutant mouse embryos.** (A-F) Whole-mount X-gal staining of $R26^{CreERT2/R}$;$Taf10^{flox/+}$ control at E9.5 (A), $R26^{+/R}$;$Taf10^{flox/flox}$ control at E10.5 (C) and E11.5 (E), and $R26^{CreERT2/R}$;$Taf10^{flox/flox}$ mutant at E9.5 (B), E10.5 (D) and E11.5 (F) after tamoxifen (tam) treatment at E7.5. (G) Western blot analysis of E9.5 *R26Cre;Taf10* whole embryos, treated (+) or not (–) with tamoxifen at E7.5, with anti-TAF10 or anti-histone H3 antibodies. (H,I) Confocal *z*-stack image projection of E9.25 *R26Cre;Taf10;Luvelu*/+ untreated (H) or tamoxifen-treated (I) embryos. so, somites. Scale bars: 500 µm in A-F; 100 µm in H,I.

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

Venus-PEST fusion. *Luvelu* expression is not affected in the absence of TAF10 at E9.5 (Fig. 1H,I), clearly indicating that transcription initiation still occurs in the *R26Cre;Taf10* mutant embryos, at least in the PSM. Altogether, these results show that, in mutants in which *Taf10* deletion is induced at E7.5, no TAF10 protein is detected in the PSM at E9.5, yet periodic gene transcription in the PSM is not affected.

### Analyses of TFIID and SAGA composition in the absence of TAF10 in the mouse embryo

Next, we set out to analyze TFIID and SAGA composition by mass spectrometry in E9.5 mouse embryos, when no TAF10 protein is detected. To purify these complexes, we collected E9.5 embryos from $R26^{CreERT2/CreERT2};Taf10^{flox/flox} \times Taf10^{flox/flox}$ crosses, treated (mutant) or not (control) with tamoxifen at E7.5. Complete *Taf10* deletion was assessed by PCR (data not shown) and western blot analysis, which confirmed the absence of detectable full-length TAF10 protein (Fig. 2A). Interestingly, in whole cell extracts from mutants, expression of TBP, TAF4A, TAF5 and TAF6 was not affected, whereas expression of TAF8, the main TFIID partner of TAF10, was strongly decreased (Fig. 2A), suggesting that the TAF8-TAF10 interaction is required for the stabilization of TAF8. We then compared TFIID and SAGA composition in the presence or absence of TAF10 by performing immunoprecipitations (IPs) from whole cell extracts of different TFIID and SAGA subunits using anti-TBP or anti-TAF7 antibodies (for TFIID) and with anti-TRRAP or anti-SUPT3 (for SAGA). Composition of the immunoprecipitated complexes was analyzed by mass spectrometry (Table S1). The normalized spectral abundance factor (NSAF) values were calculated for comparison of control and *Taf10* mutant samples (Zybailov et al., 2006).

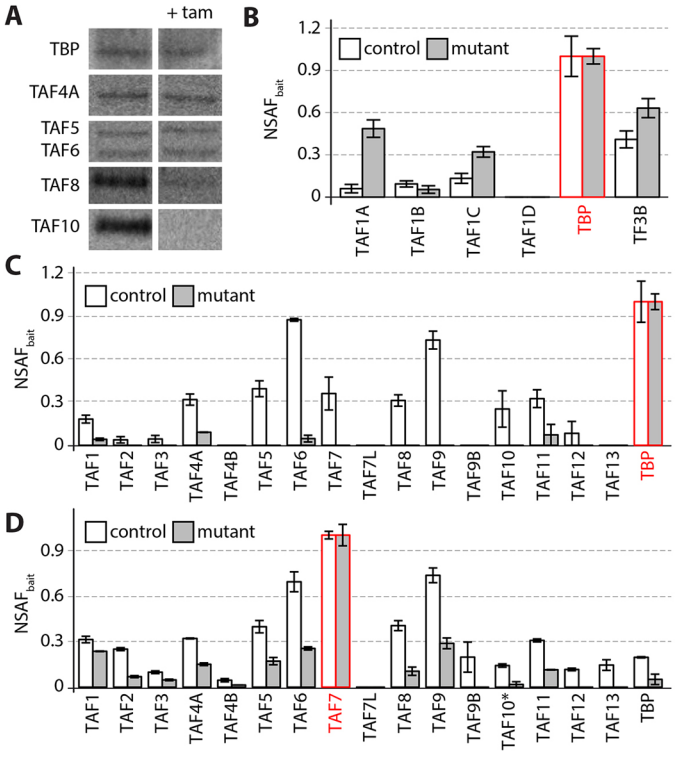



**Fig. 2. TFIID assembly defect in *R26Cre;Taf10* mutant embryos.** (A) Western blot analysis of the expression of TBP, TAF4A, TAF5, TAF6, TAF8 and TAF10 from whole cell extracts of E9.5 *R26Cre;Taf10* control (left, untreated) or mutant (right, treated with tamoxifen at E7.5) embryos. (B) TBP $NSAF_{bait}$ values for SL1 complex subunits (TAF1A, TAF1B, TAF1C, TAF1D and TBP) and TF3B-TBP complex. (C,D) $NSAF_{bait}$ values for TFIID subunits of TBP IP (C) and TAF7 IP (D). Bait proteins are indicated in red. Control and mutant IPs are indicated in white and gray, respectively. TAF10* corresponds to the full-length TAF10 protein. Error bars indicate s.d. $n$=3.

In control embryos, the full-length TAF10 protein is represented by four peptides (Fig. S4A). In mutant embryo samples, no TAF10 peptides were detected in TBP and TRRAP IPs. By contrast, in TAF7 and SUPT3 IPs we detected significant amounts (albeit reduced compared with control) of the TAF10 N-terminal peptide (peptide #1; Fig. S4B,C). The $Taf10^{flox}$ conditional mutation deletes exon 2, resulting in an out-of-frame fusion of exon 1 to exon 3 leading to premature truncation of TAF10 protein. This deletion is thus expected to produce a truncated N-terminal fragment of TAF10 containing peptide #1, but not the other peptides (Fig. S4D). The fact that no TAF10 peptides are detected in TBP and TRRAP IPs suggests that the truncated N-terminal peptide remaining in the mutant cannot participate in fully assembled TFIID or SAGA complexes. In addition, importantly, no TFIID subunits could be immunoprecipitated from murine $R26^{CreERT2/R};Taf10^{flox/flox}$ embryonic stem cells (ESCs), after 4-hydroxytamoxifen treatment, with an antibody that recognizes the N-terminal part of the TAF10 protein (Fig. S3B) and is able to immunoprecipitate the TFIID complex (Fig. S5A,B), showing that the truncated peptide is not part of a fully assembled TFIID complex. No conclusion could be drawn for the SAGA complex since this anti-N-terminal TAF10 antibody did not co-immunoprecipitate any of the mouse SAGA subunits even in control conditions (Fig. S5C). These data are consistent with the fact that the mutant truncated protein does not contain the HFD (Soutoglou et al., 2005). Thus, for further analyses and to score only the full-length protein we took into account peptides #2 to #4, which should be absent from the full-length TAF10 protein after deletion of the genomic sequences (TAF10*; Fig 2D, Fig. 3C, Fig. S4A,D), for TAF7 IPs (Fig. 2D) and SUPT3 IPs (Fig. 3C).

TBP is also part of SL1 and TFIIIB complexes, which are involved in Pol I and Pol III transcription, respectively (Vannini and Cramer, 2012). Importantly, TAF10 absence does not perturb the interaction of TBP with its non-TFIID partners, highlighting the lack of non-specific effects (Fig. 2B). In *Taf10* mutant embryos, we observed an increased interaction between TBP and the larger SL1 subunits TAF1A and TAF1C, suggesting that TBP might be redistributed in Pol I TAF-containing complexes in the absence of TAF10. This is consistent with the observation that there is no free TBP in the cells (Timmers and Sharp, 1991). In control TBP and TAF7 IPs, all the canonical TFIID subunits were detected (Fig. 2C,D). Interestingly, in *Taf10* mutant embryos, TBP IP reveals that TBP is mostly disengaged from TFIID, as only a few TAFs co-immunoprecipitate with TBP and in very low amounts (Fig. 2C). This TFIID dissociation is also observed in the TAF7 IP in the absence of TAF10 (Fig. 2D). Surprisingly, however, owing to the very efficient TAF7 IP (Table S1) we can still detect residual TFIID complexes (Fig. 2D). It is important to note that even if the anti-TAF7 antibody is able to co-immunoprecipitate several TAFs, TAF9B, TAF12 and TAF13 are not detected in the mutant, further supporting the conclusion that TAF10 absence strongly affects TFIID assembly.

In order to assess SAGA composition, we performed IPs against two SAGA subunits: SUPT3 and TRRAP. TRRAP is also a member of the chromatin remodeling complex TIP60/NuA4 (Sapountzi and Côté, 2011). As the interactions between TRRAP and TIP60/NuA4 subunits were not affected (Fig. 3A), we conclude that TAF10 absence does not interfere with the interactions between TRRAP and its non-SAGA partners. In both mutant TRRAP (Fig. 3B) and

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

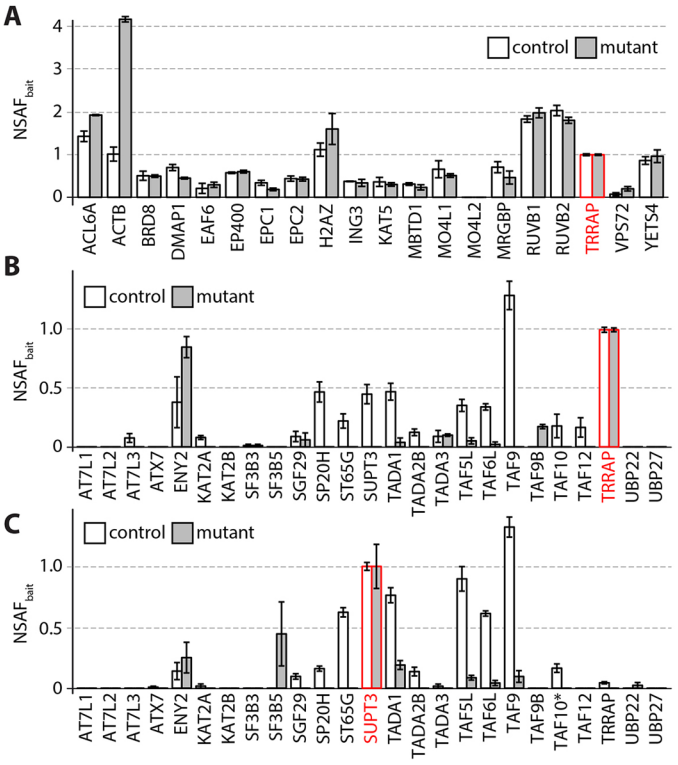



**Fig. 3. SAGA assembly defect in *R26Cre;Taf10* mutant embryos.** (A) $NSAF_{bait}$ values for TIP60/NuA4 complex subunits of TRRAP IP from control or mutant extracts. (B,C) $NSAF_{bait}$ values for SAGA subunits of TRRAP IP (B) and SUPT3 IP (C) from control or mutant extracts. Bait proteins are indicated in red. TAF10* corresponds to the full-length TAF10 protein. Error bars indicate s.d. *n*=3.

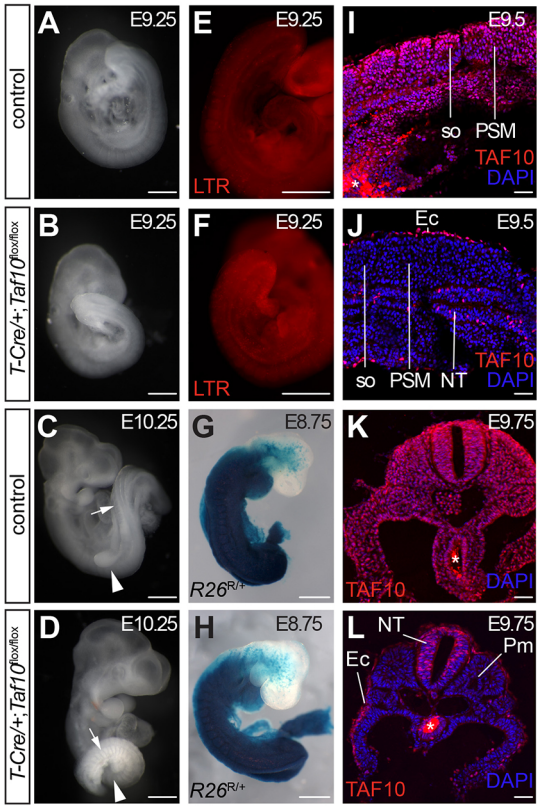



**Fig. 4. Efficient *Taf10* conditional deletion in the paraxial mesoderm.** (A-C) Whole-mount right-lateral view of control (A,C) and *T-Cre;Taf10* mutant (B,D) embryos at E9.25 (A,B) and E10.25 (C,D). Arrowheads indicate the position of the forelimb bud that is absent in the mutant; arrows indicate the somites. (E,F) Cell death assay by LysoTracker Red (LTR) staining of E9.25 control (E) and *T-Cre;Taf10* mutant (F) embryos. (G,H) Whole-mount X-gal staining of E8.75 *T-Cre/+;R26*$^{R/+}$ control (G) and *T-Cre/+;R26*$^{R/+}$*;Taf10*$^{flox/flox}$ mutant (H) embryos showing the efficient early recombination within the paraxial mesoderm. (I-L) DAPI counterstaining of TAF10 immunolocalization on E9.5 sagittal (I,J) and E9.75 transverse (K,L) sections from control (I,K) and *T-Cre;Taf10* mutant (J,L) embryos. Asterisk indicates background due to secondary antibody trapping in the endoderm lumen. Ec, ectoderm; NT, neural tube; Pm, paraxial mesoderm; PSM, presomitic mesoderm; so, somites. Scale bars: 500 µm in A-H; 50 µm in I-L.

SUPT3 (Fig. 3C) IPs we observed a dramatic reduction in the amount of SAGA subunits co-immunoprecipitated, clearly showing a defect in the assembly of SAGA. In contrast to TAF7 IP, we were not able to detect any residual canonical SAGA complexes in the mutant samples in the SUPT3-IP.

Altogether, these results strongly suggest that TAF10 is crucial for the assembly of both TFIID and SAGA in the mouse embryo, since the formation of both complexes is seriously impaired in *R26Cre;Taf10* mutant embryos.

### *Taf10* conditional deletion in the paraxial mesoderm

Our next goal was to analyze the requirement for TAF10 in transcription during development. Somitogenesis is a dynamic developmental process in vertebrate embryos relying on periodic transcriptional waves sweeping from posterior to anterior in the PSM (Hubaud and Pourquié, 2014). As described above, the dynamic expression of the *Luvelu* cyclic reporter is not affected in the PSM of E9.5 *R26Cre;Taf10* mutant embryos (Fig. 1H,I). We carried out a *Taf10* conditional deletion in the PSM using the *T-Cre* line (Perantoni, 2005). This line expresses Cre in the primitive streak under the control of 500 bp of *T* promoter sequence (Clements et al., 1996), leading to efficient recombination in the mesoderm before E7.5, including in paraxial mesoderm progenitors (Perantoni, 2005). *Taf10* conditional deletion is embryonic lethal as no *T-Cre/+;Taf10*$^{flox/flox}$ (*T-Cre;Taf10*) mutants could be recovered at birth (data not shown). At E9.25, control and *T-Cre;Taf10* mutant embryos are very similar, except that some mutant embryos show a curved trunk (Fig. 4A,B). At E10.25, *T-Cre;Taf10* mutant embryos exhibit normal anterior development but show an apparent growth arrest of the trunk region, a helicoidal trunk lacking limb buds (Fig. 4C,D) and a degeneration of the allantois and placenta (data not shown). Whereas at E9.25 mutant and control somites were morphologically similar (Fig. 4A,B), E10.25 mutant somites were much smaller than the controls (Fig. 4C,D). Similar observations were made using the *Hes7-Cre* line (data not shown), which has a similar recombination pattern in the mesoderm (Niwa et al., 2007). LysoTracker Red staining indicates that there is no obvious cell death in the mutants at E9.25 (Fig. 4E,F). Recombination in the mesoderm is efficient, as shown by the profile of activation of the Cre reporter allele *R26*$^{R}$ at E8.75 (Fig. 4G,H). Full-length TAF10 protein expression could no longer be detected in the mesoderm of mutant embryos from as early as E8.5 (Fig. S6, Fig. 4I-L), including the PSM at E9.5 (Fig. 4I,J), whereas it is detected in the ectoderm. TAF10 expression was mosaic in the mutant neural tube (NT), which shares common progenitors with the mesoderm (Gouti et al., 2014; Tzouanacou et al., 2009). Surprisingly, these data show that there is a time window at ~E9.5 when embryonic development is not affected upon TAF10 depletion, except for the absence of limb buds, prior to an apparent growth arrest and decay at E10.5.

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

### Absence of TAF10 in the PSM does not affect somitogenesis at E9.5

To gain more insight into somitogenesis, we compared somite numbers between the different genotypes at E9.5 (Fig. 5A). Although no significant statistical differences could be detected, mutant embryos tended to have half a somite less than the other genotypes. This could be explained by a slowing down of somitogenesis at late E9.5 stage.

We next analyzed the expression of specific PSM markers using WISH. Expression of the posterior PSM marker *Msgn1* (Wittler et al., 2007) (Fig. 5B,C), the segmentation gene *Mesp2* (Saga et al., 1997) (Fig. 5D,E) or the caudal somite marker *Uncx4.1* (Neidhardt et al., 1997) (Fig. 5F,G) was unaffected in the absence of TAF10. WISH of cyclic genes of the Notch [*Lfng* (Forsberg et al., 1998; McGrew et al., 1998) and *Hes7* (Bessho et al., 2003); Fig. 5H,I, Fig. S8A,B], Wnt [*Axin2* (Aulehla et al., 2003); Fig. S7C,D] or FGF [*Snai1* (Dale et al., 2006); Fig. S7E,F] pathways revealed that the different phases of expression could be observed in *T-Cre;Taf10* mutant embryos. Altogether, the rhythmic transcription of the cyclic genes in the absence of TAF10 suggests that active transcription proceeds normally in the PSM of mutant embryos.

### Absence of TAF10 differentially affects mesoderm derivatives

Limb bud outgrowth requires signals such as FGF8 from the apical ectodermal ridge (AER), which controls proliferation of the underlying mesenchyme derived from the lateral plate mesoderm (LPM) (Zeller et al., 2009). On E10.25 transverse sections from control embryos, mesodermal nuclei (including those in the LPM) are regularly shaped (Fig. 6A,C,E). In *T-Cre;Taf10* mutants (Fig. 6B) the paraxial mesoderm nuclei appear normal (Fig. 6D), whereas in the LPM [and in the intermediate mesoderm (data not shown)] we observed massive nuclear fragmentation characterized by the presence of pyknotic nuclei (Fig. 6F). Since we did not observe any difference in the efficiency of TAF10 protein depletion between the paraxial mesoderm and the LPM from as early as E8.5 (Fig. S6), these data indicate that the LPM is more sensitive to *Taf10* loss than the paraxial mesoderm.

We carried out WISH in order to test whether *Taf10* loss differentially affects the expression of specific markers of the different types of mesoderm. Expression of the LPM marker *Hand2* (Fernandez-Teran et al., 2000) is clearly downregulated in the mutants (Fig. 6G,H). Similar observations were made with *Prdm1*, which is expressed in the growing mesenchyme during limb bud outgrowth (Vincent et al., 2005) (data not shown). The absence of *Fgf8* induction in the presumptive AER in E9.5 *T-Cre;Taf10* mutant embryos (Fig. 6K,L) indicates that the LPM defect is early and probably precedes the cell death in this tissue, since no obvious cell death could be detected at E9.25 (Fig. 4F). The cell death observed later on in the LPM is, however, not caused by the lack of *Fgf8* expression as it is also observed at non-limb levels. By contrast, paraxial mesoderm marker analysis shows that *Pax3* expression in the anterior PSM and early somites (Goulding et al., 1991) is normal (Fig. 6I,J). Similarly, *Fgf8* expression domains in the rostral and

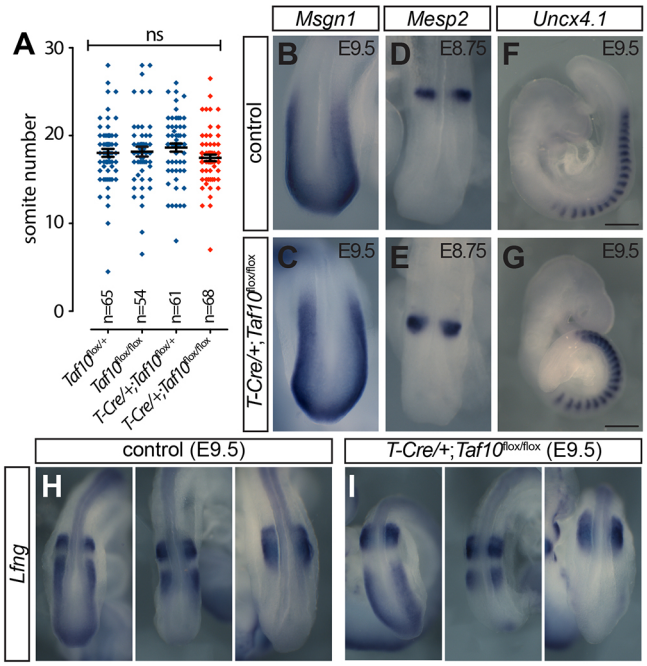



**Fig. 5. Absence of TAF10 in the PSM does not affect segmentation.** (A) Somite number quantification (one-way ANOVA; ns, non significant). Error bars indicate s.e.m. and the middle bar indicates the mean. (B-I) WISH of E9.5 (B,C,F-I) and E8.75 (D,E) control (B,D,F,H) and *T-Cre/+;Taf10*$^{flox/flox}$ mutant (C,E,G,I) embryos for the posterior PSM marker *Msgn1* (B,C), the segmentation gene *Mesp2* (D,E), the caudal somite marker *Uncx4.1* (F,G) and the cyclic gene *Lfng* (H,I). Dorsal tail tip (B-E,H,I) or right-lateral (F,G) views are presented. Scale bars: 100 µm in B-E,H,I; 500 µm in F,G.

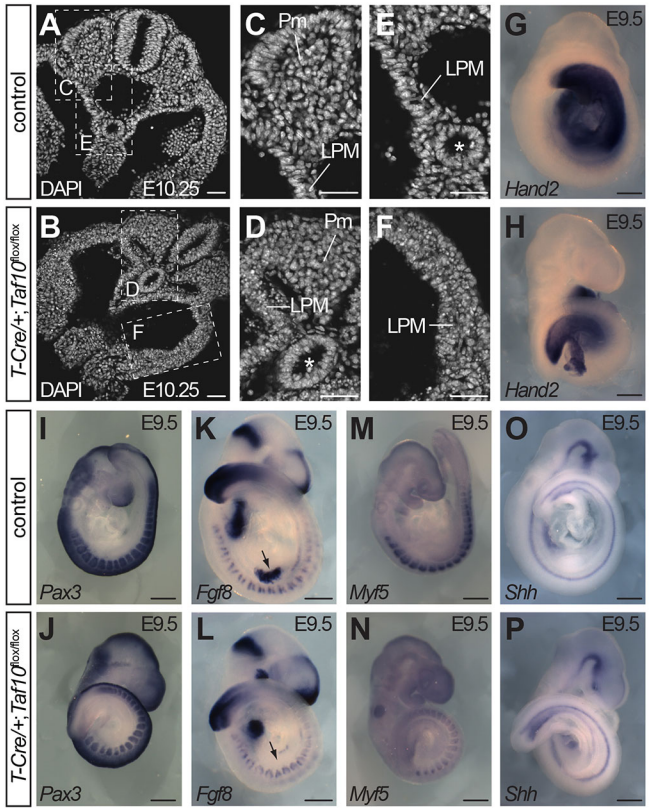



**Fig. 6. Absence of TAF10 differentially affects the different types of mesoderm.** (A-F) DAPI-stained transverse sections of E10.25 control (A, magnified in C,E) and *T-Cre/+;Taf10*$^{flox/flox}$ mutant (B, magnified in D,F) embryos showing nuclear fragmentation in LPM but normal nuclear morphology in the paraxial mesoderm. Asterisks indicate the endoderm. (G-P) WISH of E9.5 control (G,I,K,M,O) and *T-Cre/+;Taf10*$^{flox/flox}$ mutant (H,J, L,N,P) embryos for *Hand2* (G,H), *Pax3* (I,J), *Fgf8* (K,L), *Myf5* (M,N) and *Shh* (O,P). Arrows indicate the apical ectodermal ridge. LPM, lateral plate mesoderm; Pm, paraxial mesoderm. Scale bars: 50 µm in A-F; 500 µm in G-P.

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

caudal lips of the dermomyotome (Crossley and Martin, 1995) are not affected at E9.5 in the mutant paraxial mesoderm (Fig. 6K,L). Expression of *Pax3* in the dermomyotome (Goulding et al., 1991) and of *Myf5* in the myotome (Ott et al., 1991) are however decreased in *T-Cre;Taf10* mutants (Fig. 6I,J,M,N). Defective myotome formation was evidenced by immunolocalization of myogenin or myosin heavy chains at E9.5 and E10.5 (data not shown). Similar observations were made in *Hes7-Cre/+;Taf10*$^{flox/flox}$ mutant embryos (Fig. S8). Expression of *Shh* in the notochord is normal (Echelard et al., 1993), indicating that the axial mesoderm is not obviously affected in *T-Cre;Taf10* mutant embryos (Fig. 6O,P). Altogether, these results indicate different requirements for TAF10 depending on the type of mesoderm. However, we cannot rule out the possibility that the effect seen in the LPM arises secondarily to a defect in the developing paraxial mesoderm.

### Absence of TAF10 does not affect global steady-state mRNA and cyclic transcription in the PSM

Our next goal was to investigate Pol II transcription status in mutant embryos. We first compared steady-state rRNA (Pol I) and mRNA (Pol II) transcript levels by quantifying the absolute expression levels of 18S ribosomal RNA (*Rn18s*) versus classical Pol II housekeeping genes (*Actc1*, *Gapdh* and *Rplp0*) (Fig. 7A). No significant differences between mutant and control samples were detected when comparing the results obtained with three different pairs of *Rn18s* primers (Fig. 7B). The results were similar for *Gapdh* and *Rplp0* (Fig. 7B). Expression of the *Luvelu* reporter (Aulehla et al., 2008) in *T-Cre;Taf10* mutant embryos (Fig. 7C,D) supports the idea that cyclic transcription initiation still occurs in the *T-Cre;Taf10* mutant PSM at E9.5. Altogether, these results indicate that, at ∼E9.5, absence of detectable TAF10 does not affect global steady-state mRNA and PSM-specific cyclic transcription.

### Expression of specific genes is altered in the PSM at E9.5 in the absence of TAF10

We next performed a transcriptome analysis in order to see whether specific genes were affected in the absence of TAF10. We performed microarray analyses from microdissected PSMs of E9.5 (17-19 somites) control and *T-Cre;Taf10* mutant embryos (Fig. 8A). Analysis by scatter plot shows that TAF10 loss has only a very minor impact on gene expression (Fig. 8B). We then performed a statistical analysis using fold change ranking ordered statistics (FCROS) (Dembélé and Kastner, 2014) and found 369 differentially expressed genes (218 downregulated and 151 upregulated) using a fold change cut-off of 1.5 (Fig. 8C, see Table S2). This analysis identified genes related to the cell cycle, TAFs, signaling pathways, and Hox/para-Hox genes (see Table 1). We also observed that some genes previously identified as cyclic genes in the PSM, such as *Egr1*, *Cyr61*, *Dkk1*, *Spry4* and *Rps3a* (Krol et al., 2011), are also differentially expressed in *T-Cre;Taf10* mutant PSMs (Table 1, Fig. S9A). Interestingly, the most highly upregulated gene (4.8-fold) is *Cdkn1a*, which encodes a cyclin-dependent kinase inhibitor involved in G1 arrest (Dulić et al., 1994). We identified *Gas5*, a tumor suppressor gene that encodes two long non-coding RNAs and several small nucleolar RNAs in its introns (Ma et al., 2015), as the most downregulated gene (−2- to −4.9-fold). We confirmed the upregulation of *Cdkn1a*, *Cdkn1c*, *Ccng1* and *Cdkl3* and the downregulation of *Gas5* by RT-qPCR using control and *T-Cre;Taf10* mutant tail tips (Fig. 8D). Upregulation of *Cdkn1a* and *Cdkn1c* could explain the growth arrest that is observed in *T-Cre;Taf10* mutant embryos.

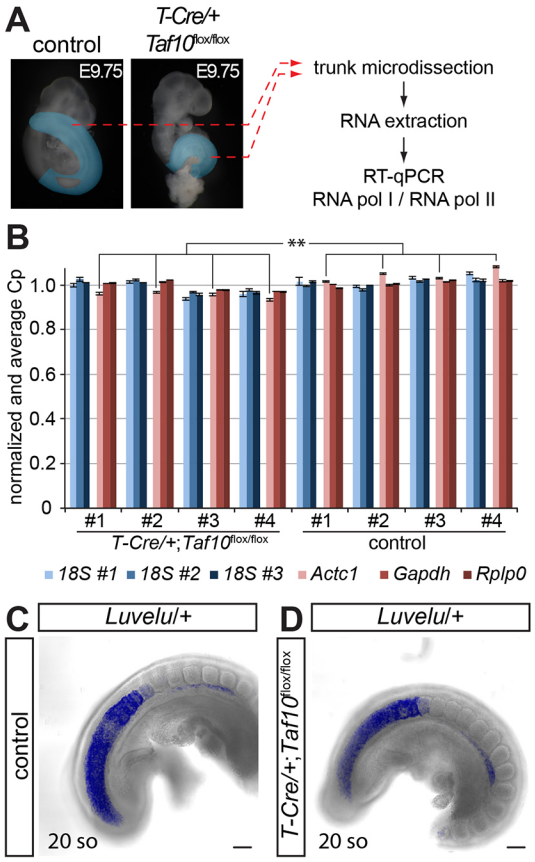



**Fig. 7. Global transcription is not affected in the absence of TAF10 in the paraxial mesoderm.** (A) Comparison between Pol II and Pol I transcription. The trunk axial structures highlighted in blue were dissected from E9.75 control and *T-Cre/+;Taf10*$^{flox/flox}$ mutant embryos and RT-qPCR was performed for Pol I-specific and Pol II-specific housekeeping genes. (B) Comparison of averaged and normalized expression of Pol I-specific (blue) and Pol II-specific (red) markers from control (right side) and mutant (left side) samples. $^{**}P<0.01$ (Aspin-Welch corrected Student's *t*-test). Error bars indicate s.e.m. $n=4$. (C,D) Confocal *z*-stack image projection of E9.5 *Luvelu*/+ control (C) and *T-Cre/+;Taf10*$^{flox/flox}$*;Luvelu*/+ mutant (D) embryos. so, somites. Scale bars: 100 µm.

Some TFIID-TAFs were also upregulated: *Taf5* (1.5-fold), *Taf6* (1.7-fold) and *Taf9b* (1.6-fold) (Table 1, Table S2). We validated these differential expressions by RT-qPCR and found that most of the genes encoding the other TAFs were also upregulated (Fig. S9B). The biological significance of these differences is not clear as no obvious increase in protein levels could be observed for TAF4A, TAF5 and TAF6 (Fig. 2A). *Taf10* expression is downregulated in *T-Cre;Taf10* mutant tail tips, as is that of *Taf8*, which encodes the main partner of TAF10 in TFIID. These data suggest that the decreased level of TAF8 protein observed in *R26Cre;Taf10* lysates (Fig. 2A) could also be related to transcriptional regulation. No differences could be detected for the SAGA-specific TAF5L and TAF6L (Fig. S9C). Altogether, our data show that, in the PSM at E9.5, gene expression controlled by Pol II is not globally affected in the absence of TAF10; however, the lack of TAF10 could induce a change in the steady-state mRNA levels of specific genes.

## DISCUSSION

The composition of TFIID and SAGA complexes in the developing mouse embryo has not yet been described. Here, we analyzed the composition of these complexes in E9.5 mouse embryos in the

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

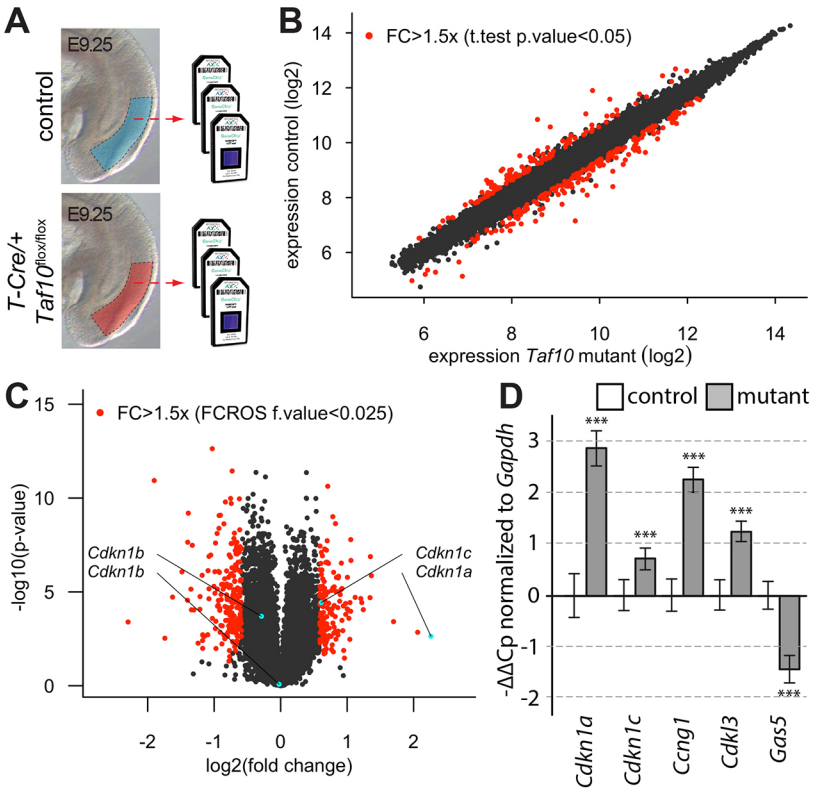



**Fig. 8. A limited specific effect on Pol II transcription in the absence of TAF10 in the PSM.** (A) Strategy used for the microarray analysis from E9.5 microdissected PSM of control (blue) and *T-Cre;Taf10* mutant (red) embryos. (B) Scatter plot comparing gene expression between control and *T-Cre;Taf10* mutant PSM. Red dots correspond to statistically significant differences for a fold change greater than 1.5 after *t*-test. (C) Volcano plot comparing gene expression between control and *T-Cre/+;Taf10*$^{flox/flox}$ mutant PSM after FCROS analysis. Red dots correspond to statistically significant differences for a fold change greater than 1.5. (D) RT-qPCR analysis for cell cycle genes from E9.25 control and *TCre;Taf10* mutant tail tips. $-\Delta\Delta$Cp values are normalized to *Gapdh*. ***$P$<0.001 (Aspin-Welch corrected Student's *t*-test). Error bars indicate s.e.m. $n$=4.

presence and absence of TAF10. We showed that the absence of TAF10 strongly affects TFIID and SAGA formation. *Taf10* deletion during somitogenesis confirmed the requirement of TAF10 during embryonic development in agreement with previous studies (Indra et al., 2005; Mohan et al., 2003; Tatarakis et al., 2008). However, in contrast to these studies, we identified a time window at ∼E9.5 when no obvious somitogenesis defects are detected, despite the absence of detectable full-length TAF10 protein in mutant embryos. In these mutants, transcription is still broadly functional as shown by the lack of any global effect on Pol II transcription.

### TAF10 is required for TFIID and SAGA assembly during development

Our data demonstrate a global decrease in TFIID and SAGA assembly in *Taf10* mutant embryos. In F9 cells, in the absence of TAF10, TFIID is minimally affected by the release of TBP from the complex, while interaction between the different TAFs is maintained (Mohan et al., 2003), whereas in the liver TFIID assembly is completely abrogated (Tatarakis et al., 2008). These differences could be explained either by cell type-specific differences or by a difference in the timing of these analyses following *Taf10* deletion, as Tatarakis et al. (2008) performed their experiments 8-15 days after *Taf10* deletion. The status of SAGA has not previously been investigated in *Taf10* mutant embryos. Our work demonstrates for the first time that not only TFIID, but also SAGA is affected in *Taf10* mutant embryos. Our new data show that the defect in the assembly of canonical TFIID and SAGA is already observed 2 days after the induction of *Taf10* deletion, a timing that coincides with the disappearance of detectable full-length TAF10 protein. On the other hand, we can still detect reduced interactions between TAF7 and several TAFs following *Taf10* deletion suggesting that, as observed in HeLa or F9 cells, there could be some TFIID-like complexes that do not contain TAF10, albeit in reduced levels. Our data exclude the existence of similar TAF10-less SAGA-like complexes in the embryo.

TAF10 depletion is very efficient since no TAF10 proteins can be detected by western blot in the mutant embryo lysates. Analysis of the detected peptides strongly suggests that it is only in the TAF7 IP (TFIID) that potential full-length TAF10 proteins are detected, albeit at very low frequency. This suggests that very low levels of canonical TFIID complexes could still be present at E9.5 in the mutant embryonic lysates. Furthermore, these results, in comparison with the SAGA IPs, suggest that TAF10 is very stable when incorporated into TFIID, probably because of the lower rate of TFIID turnover compared with that of SAGA.

TFIID is built from submodules that assemble in the cytoplasm, at least *in vitro* (Bieniossek et al., 2013; Trowitzsch et al., 2015), and it is likely that such TFIID submodules are immunoprecipitated in our experiments since we performed our analyses using whole cell extracts. The TAF7 paralog TAF7L, which has been associated with germ cells and adipocytes (Zhou et al., 2013a,b), is not present in TFIID IPs, indicating that the majority of TFIID contains TAF7, at least at E9.5. However, other TAF paralogs such as TAF4A and TAF4B, TAF9 and TAF9B, are detected. This potential TFIID diversity could exist inside all the cells or could be cell type specific and could explain the developmental differences observed between LPM and paraxial mesoderm. However, novel methods will be required to characterize the composition of TFIID and SAGA complexes in a cell type-specific manner in the embryo.

### A truncated TAF10 protein can potentially be integrated into TFIID and SAGA submodules

Our strategy conditionally removes exon 2 and theoretically leads to the splicing of exon 1 to exon 3 (Mohan et al., 2003). These exons are not in frame and therefore the 77 amino acids coded by exon 1 are followed by 15 extra amino acids in the mutant

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

**Table 1. Selection of differentially expressed genes in the PSM of E9.5 *T-Cre;Taf10* mutant embryos**

| Description | Gene symbol | Absolute FC | *F*-value |
|---|---|---|---|
| Cell cycle | | | |
| growth arrest specific 5 | *Gas5* | −4.908 | 0.0177 |
| | | −3.736 | 0.0178 |
| | | −2.635 | 0.0179 |
| | | −2.073 | 0.0183 |
| cyclin-dependent kinase inhibitor 1A (P21) | *Cdkn1a* | 4.790 | 0.9820 |
| cyclin-dependent kinase inhibitor 1C (P57) | *Cdkn1c* | 1.525 | 0.9795 |
| cyclin-dependent kinase-like 3 | *Cdkl3* | 1.780 | 0.9811 |
| cyclin G1 | *Ccng1* | 2.006 | 0.9817 |
| RNA pol I-associated complexes | | | |
| TATA box binding protein (Tbp)-associated factor, RNA polymerase I, D | *Taf1d* | −2.317 | 0.0181 |
| | | −2.266 | 0.0181 |
| | | −2.040 | 0.0186 |
| | | −1.790 | 0.0193 |
| | | −1.632 | 0.0204 |
| RNA pol II-associated complexes | | | |
| TAF6 RNA polymerase II, TATA box binding protein (TBP)-associated factor | *Taf6* | 1.724 | 0.9809 |
| TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor | *Taf9b* | 1.591 | 0.9786 |
| TAF5 RNA polymerase II, TATA box binding protein (TBP)-associated factor | *Taf5* | 1.536 | 0.9788 |
| polymerase (RNA) II (DNA directed) polypeptide A | *Polr2a* | 1.505 | 0.9792 |
| Signaling pathways and transcription factors | | | |
| Mix1 homeobox-like 1 (*Xenopus laevis*) | *Mixl1* | 1.566 | 0.9787 |
| T-box 6 | *Tbx6* | 1.547 | 0.9799 |
| E26 avian leukemia oncogene 2, 3′ domain | *Ets2* | −1.538 | 0.0217 |
| fibroblast growth factor 9 | *Fgf9* | −1.550 | 0.0215 |
| ephrin A5 | *Efna5* | −1.628 | 0.0208 |
| dual specificity phosphatase 4 | *Dusp4* | −1.648 | 0.0202 |
| R-spondin 3 homolog (*Xenopus laevis*) | *Rspo3* | −1.662 | 0.0200 |
| cytochrome P450, family 26, subfamily a, polypeptide 1 | *Cyp26a1* | −1.671 | 0.0206 |
| caudal type homeobox 4 | *Cdx4* | −1.519 | 0.0232 |
| homeobox A7 | *Hoxa7* | 1.636 | 0.9806 |
| homeobox B7 | *Hoxb7* | 1.823 | 0.9814 |
| homeobox D1 | *Hoxd1* | 1.971 | 0.9817 |
| homeobox A3 | *Hoxa3* | 2.550 | 0.9820 |
| Cyclic genes | | | |
| early growth response 1 | *Egr1* | 1.610 | 0.9791 |
| cysteine rich protein 61 | *Cyr61* | 1.713 | 0.9810 |
| dickkopf homolog 1 (*Xenopus laevis*) | *Dkk1* | 1.945 | 0.9811 |
| sprouty homolog 4 (*Drosophila*) | *Spry4* | −1.539 | 0.0219 |
| ribosomal protein S3A | *Rps3a* | −1.586 | 0.0209 |

Statistical analysis was performed using FCROS with a cut-off of 1.5 for the fold change (FC). Difference is considered significant for an *F*-value below 0.025 or above 0.975. Where multiple entries appear for the same gene, each corresponds to a different specific probe set.

(Fig. S4D). This mutant protein has the N-terminal unstructured domain of TAF10 but, more importantly, lacks its HFD required for the interaction with TAF3, TAF8 or SUPT7L/ST65G (Soutoglou et al., 2005). HFD-HFD interactions are crucial for nuclear import of TAF10, which does not contain any NLS (Soutoglou et al., 2005). Since no TFIID subunits could be co-immunoprecipitated from whole cell extracts of $R26^{CreERT2/R};Taf10^{flox/flox}$ ESCs, after 4-hydroxytamoxifen treatment, with an antibody that recognizes the N-terminal part of TAF10 (Fig. S5), it is very unlikely that this truncated protein can be incorporated into mature SAGA and TFIID complexes that are functional in the nucleus. However, we cannot rule out the possibility that this truncated protein could be incorporated into rare cytoplasmic submodules containing TAF7 or SUPT3. Nevertheless, because the *Taf10* mutant heterozygotes are indistinguishable from control embryos (Fig. 1A), this also argues against a dominant-negative effect of this peptide.

Another interesting question is the functionality of these potentially partial TFIID and/or SAGA complexes that are fully depleted of TAF10 protein or contain the truncated TAF10. From our data, it is obvious that these different partial complexes cannot fully compensate for the loss of wild-type complexes, but one cannot rule out a partial activity. Future analyses of the difference between the different types of mesoderm could help to elucidate whether such partial non-canonical TFIID and/or SAGA complexes have activities.

### Differential sensitivity to *Taf10* loss in the mesoderm

*Taf10* deletion in the mesoderm or in the whole embryo leads to developmental arrest that could be explained by the upregulation of *Cdkn1a* and *Cdkn1c* expression. Similar observations were made in yeast (Kirschner et al., 2002) and in F9 cells (Metzger et al., 1999) following depletion of TAF10. Surprisingly, we also observed the downregulation of the tumor suppressor *Gas5*, which is associated with increased proliferative and anti-apoptosis effects in cancer cells (Pickard and Williams, 2015). Interestingly, *Cdkn1a* expression is positively controlled by *Gas5* in stomach cancer at the transcript and protein levels (Liu et al., 2015). It is thus possible that TAF10 is required for the correct functioning of the *Gas5* regulatory network during development.

The phenotypes of null mutations in genes encoding TFIID-TAFs, such as *Taf7* (Gegonne et al., 2012) or *Taf8* (Voss et al., 2000), are very similar to that of the *Taf10* mutant (Mohan et al., 2003). In particular, these mutations are embryonic lethal around implantation stage. Moreover, *Taf7* null MEFs stop proliferating, suggesting that the growth arrest observed in our mutants is a direct consequence of the failure to properly build TFIID. We cannot exclude a potential contribution of SAGA loss in our mutants. However, deletion of genes coding for different enzymatic activities of SAGA such as *Kat2a;Kat2b* or *Usp22* are embryonic lethal, but with phenotypes much less severe than that of *Taf10* mutation (Lin et al., 2012; Xu et al., 2000; Yamauchi et al., 2000). Interestingly, axial and paraxial mesoderm formation are affected in *Kat2a;Kat2b* mutants, whereas extraembryonic and cardiac mesoderm formation are not (Xu et al., 2000), strongly suggesting that SAGA could also have different functions in different types of mesoderm.

Another striking observation is that, although no TAF10 protein could be detected as early as E8.5 in the mesoderm of *T-Cre;Taf10* mutant embryos, we observed a difference in sensitivity to *Taf10* loss between the LPM (and the intermediate mesoderm) and the paraxial mesoderm. We observed a very early defect in the LPM, with strong downregulation of specific markers and absence of limb bud outgrowth. The absence of limb buds could be explained by a defect in FGF10 signaling activation in the mesoderm and/or by cell death in the LPM that occurs earlier than in the paraxial mesoderm of *T-Cre;Taf10* mutants. The relative resistance of the mutant paraxial mesoderm to cell death also suggests a difference of sensitivity. A similar observation has been made in F9 cells, where RA-induced differentiation of F9 cells into primitive endoderm rescued the apoptosis of *Taf10* mutant cells (Metzger et al., 1999). This effect was not observed when F9 cells were differentiated into parietal endoderm in the presence of RA and cAMP (Metzger et al., 1999). An interesting possibility is that, being the principal source

Development (2017) 144, 3808-3818 doi:10.1242/dev.146902

of RA (Niederreither et al., 1997), the paraxial mesoderm is protected from cell death in the mutant embryos via an autocrine mechanism. A difference in sensitivity has also been observed in *Taf10* mutant blastocysts, where the inner cell mass dies by apoptosis, whereas trophoblast can be maintained in culture (Mohan et al., 2003). It is interesting to note that trophoblast, primitive and parietal endoderms are extraembryonic structures and are not part of the fully developed embryo. This is the first *in vivo* observation of a difference in sensitivity to the loss of *Taf10* in an embryonic lineage. Since *Taf10* was deleted in paraxial mesoderm and LPM progenitors, we cannot rule out the possibility that the increased sensitivity of the LPM is indirect and mediated by the paraxial mesoderm, although we did not observe any obvious change in gene expression in the PSM at a time when limb bud development is already affected. Nevertheless, a tempting speculation is that TAF10 could serve as an interface of interaction with an LPM-specific transcription factor, as has been described recently for GATA1 during erythropoiesis (Papadopoulos et al., 2015).

## MATERIALS AND METHODS

### Mice

Animal experimentation was carried out according to animal welfare regulations and guidelines of the French Ministry of Agriculture (ethical committee C2EA-17 projects 2012-077, 2012-078, 2015050509092048). All the lines have already been described (supplementary Materials and Methods). The day of vaginal plug was scored as embryonic day (E) 0.5. Tamoxifen (Sigma) resuspended at 20 mg/ml in 5% ethanol/filtered sunflower seed oil was injected intraperitoneally [150 µl (3 mg) for a 20 g mouse] at E7.5.

### Embryos whole cell extracts

E9.5 mouse embryos (16-20 somites) were lysed in 10% glycerol, 20 mM Hepes pH 7, 0.35 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.1% Triton X-100 with protease inhibitor cocktail (PIC, Roche) on ice. Lysates were treated three times with a pestle stroke followed by three liquid nitrogen freeze-thaw cycles. Lysates were centrifuged at 20,817 rcf for 15 min at 4°C and the supernatants were used directly for IPs or stored at −80°C for western blots.

### Immunoprecipitations

Inputs were incubated with Dynabeads coated with antibodies (see supplementary Materials and Methods and Table S3) overnight at 4°C. Immunoprecipitated proteins were washed twice for 5 min each with 500 mM KCl buffer [25 mM Tris-HCl (pH 7), 5 mM $MgCl_2$, 10% glycerol, 0.1% NP40, 2 mM DTT, 500 mM KCl and PIC (Roche)], then washed twice for 5 min each with 100 mM KCl buffer (25 mM Tris-HCl pH 7, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP40, 2 mM DTT, 100 mM KCl and PIC) and eluted with 0.1 M glycine pH 2.8 three times for 5 min each. Elution fractions were neutralized with 1.5 M Tris-HCl pH 8.8.

### Western blots

Immune complexes or 15 µg embryo lysates were boiled for 10 min in 100 mM Tris-HCl pH 6.8, 30% glycerol, 4% SDS, 0.2% Bromophenol Blue, 100 mM DTT, resolved on a precast SDS-polyacrylamide gel 4-12% (Novex) and transferred to nitrocellulose membrane (Protran, Amersham). Membranes were blocked in 3% milk in PBS for 30 min and incubated with primary antibody (Table S3) overnight at 4°C. Membranes were washed three times for 5 min each with 0.05% Tween 20 in PBS. Membranes were incubated with HRP-coupled secondary antibodies (Table S3) for 50 min at room temperature, followed by ECL detection (ThermoFisher Scientific).

### Mass spectrometry analyses

Samples were analyzed using an UltiMate 3000 RSLCnano (Thermo Scientific) coupled in line with a linear trap Quadrupole (LTQ)-Orbitrap ELITE mass spectrometer via a nano-electrospray ionization source (Thermo Scientific). Data were analyzed by calculation of $NSAF_{bait}$ (see supplementary Materials and Methods).

### Section and immunolocalization

Embryos were fixed in 4% paraformaldehyde for 2 h at 4°C, rinsed three times in PBS, equilibrated in 30% sucrose/PBS and embedded in Cryomatrix (Thermo Scientific) in liquid nitrogen vapors. Sections (20 µm) were obtained on a Leica cryostat. Immunolabeling was performed as previously described (Vincent et al., 2014). Sections were counterstained with DAPI (4′,6-diamidino-2-phenylindole dihydrochloride; Molecular Probes) and imaged with an LSM 510 laser-scanning microscope (Carl Zeiss) and 20× Plan APO objective (NA 0.8).

### Luvelu imaging

Freshly dissected embryos were kept in DMEM without Phenol Red (Life Technologies). Luvelu signal was detected using an SP5 TCS confocal microscope (Leica) with a 20× Plan APO objective (NA 0.7).

### Whole-mount *in situ* hybridization (WISH), X-gal and LysoTracker Red staining

WISH was performed as described (Nagy et al., 2002). *Axin2*, *Fgf8*, *Hand2*, *Lfng*, *Msgn1*, *Myf5*, *Shh*, *Snai1* and *Uncx4.1* probes have been described (Aulehla and Johnson, 1999; Aulehla et al., 2008; Crossley and Martin, 1995; Dale et al., 2006; Echelard et al., 1993; Mansouri et al., 1997; Ott et al., 1991; Srivastava et al., 1997; Yoon et al., 2000). A minimum of three embryos were used for the classical markers and a minimum of seven embryos were used for the cyclic genes. X-gal and LysoTracker Red (Molecular Probes) stainings were performed as described (Rocancourt et al., 1990; Vincent et al., 2014).

### RT-qPCR and statistical analysis

Microdissected embryo tail tip or trunk tissue (without limb buds for the controls) was lysed in 500 µl TRIzol (Life Technologies). RNA was extracted according to the manufacturer's recommendations and resuspended in 20 µl (trunk) or 11 µl (tail tips) RNase-free water (Ambion). Reverse transcription was performed using the QuantiTect Reverse Transcription Kit (Qiagen) in 12 µl reaction volume and diluted by adding 75 µl RNase-free water. Quantitative PCRs were performed on a Roche LightCycler II 480 using LightCycler 480 SYBR Green I Master (Roche) in 8 µl reaction volume (0.4 µl cDNA, 0.5 µM primers). Four mutants and four controls with the same somite number were analyzed in triplicate. Statistical analysis and primer sequences are described in the supplementary Materials and Methods and Table S4.

### Microarray and statistical analysis

Posterior PSMs of E9.5 embryos were individually microdissected (Dequéant et al., 2006) and lysed in 200 µl TRIzol, and the yolk sac was used for genotyping. Three PSMs of 17- to 19-somite embryos of the same genotype were pooled for one replicate and analyzed on GeneChip MoGene 1.0 ST arrays (Affymetrix). Data were normalized using RMA (Bioconductor), filtered, and FCROS (Dembélé and Kastner, 2014) was used for the statistical analysis (supplementary Materials and Methods).

#### Acknowledgements

We thank Violaine Alumni and the Biochip and Sequencing Platform (IGBMC) for the microarray experiments; Doulaye Dembele for advice on statistical analysis of the microarrays; Mathilde Decourcelle and the Proteomic Platform (IGBMC) for the Orbitrap analyses; Eli Scheer for skillful advice on immunoprecipitations; Ivanka Kamenova for help with validating the antibodies; Joël Herrmann for validation of the qPCR primers; and Didier Devys, Goncalo Vilhais-Neto and Ziad Al Tanoury for critical reading of the manuscript.

#### Competing interests

The authors declare no competing or financial interests.

#### Author contributions

Conceptualization: S.D.V., L.T., O.P.; Methodology: P.B., S.D.V., M.F., A.H., M.J.; Software: S.D.V., M.J.; Validation: P.B., S.D.V., L.T.; Formal analysis: S.D.V.;

Investigation: P.B., S.D.V., M.F., A.H.; Resources: L.T., O.P.; Data curation: S.D.V., M.J.; Writing - original draft: S.D.V.; Writing - review & editing: S.D.V., M.F., A.H., L.T., O.P.; Visualization: S.D.V.; Supervision: S.D.V., L.T., O.P.; Project administration: S.D.V., L.T., O.P.; Funding acquisition: L.T., O.P.

### Funding

This work was supported by Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Université de Strasbourg, Agence Nationale de Recherche (ANR-13-BSV6-0001-02 COREAC and ANR-13-BSV8-0021-03 DiscoverIID to L.T.) and Investissements d'Avenir (ANR-10-IDEX-0002-02 and ANR-10-LABX-0030-INRT to L.T.). L.T. and O.P. are recipients of European Commission European Research Council Advanced Grants (ERC-2013-340551, Birtoaction to L.T. and ERC-2009-ADG20090506, Bodybuilt to O.P.).

### Data availability

Raw microarray data have been deposited in Gene Expression Omnibus under accession number GSE82186. Raw mass spectrometry data are available via ProteomeXchange under accession number PXD004688.

### Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.146902.supplemental

### References

**Aulehla, A. and Johnson, R. L.** (1999). Dynamic expression of lunatic fringe suggests a link between notch signaling and an autonomous cellular oscillator driving somite segmentation. *Dev. Biol.* **207**, 49-61.

**Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B. and Herrmann, B. G.** (2003). Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev. Cell* **4**, 395-406.

**Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M. B., Deng, C., Taketo, M., Lewandoski, M. and Pourquié, O.** (2008). A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nat. Cell Biol.* **10**, 186-193.

**Bessho, Y., Hirata, H., Masamizu, Y. and Kageyama, R.** (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes Dev.* **17**, 1451-1456.

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

**Bonnet, J., Wang, C.-Y., Baptista, T., Vincent, S. D., Hsiao, W.-C., Stierle, M., Kao, C.-F., Tora, L. and Devys, D.** (2014). The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. *Genes Dev.* **28**, 1999-2012.

**Clements, D., Taylor, H. C., Herrmann, B. G. and Stott, D.** (1996). Distinct regulatory control of the Brachyury gene in axial and non-axial mesoderm suggests separation of mesoderm lineages early in mouse gastrulation. *Mech. Dev.* **56**, 139-149.

**Cole, S. E., Levorse, J. M., Tilghman, S. M. and Vogt, T. F.** (2002). Clock regulatory elements control cyclic expression of Lunatic fringe during somitogenesis. *Dev. Cell* **3**, 75-84.

**Crossley, P. H. and Martin, G. R.** (1995). The mouse Fgf8 gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development* **121**, 439-451.

**Dale, J. K., Malapert, P., Chal, J., Vilhais-Neto, G., Maroto, M., Johnson, T., Jayasinghe, S., Trainor, P., Herrmann, B. and Pourquié, O.** (2006). Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. *Dev. Cell* **10**, 355-366.

**Dembélé, D. and Kastner, P.** (2014). Fold change rank ordering statistics: a new method for detecting differentially expressed genes. *BMC Bioinformatics* **15**, 14.

**Dequéant, M.-L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. and Pourquié, O.** (2006). A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science* **314**, 1595-1598.

**Dulić, V., Kaufmann, W. K., Wilson, S. J., Tisty, T. D., Lees, E., Harper, J. W., Elledge, S. J. and Reed, S. I.** (1994). p53-dependent inhibition of cyclin-dependent kinase activities in human fibroblasts during radiation-induced G1 arrest. *Cell* **76**, 1013-1023.

**Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A. and McMahon, A. P.** (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417-1430.

**Fernandez-Teran, M., Piedra, M. E., Kathiriya, I. S., Srivastava, D., Rodriguez-Rey, J. C. and Ros, M. A.** (2000). Role of dHAND in the anterior-posterior polarization of the limb bud: implications for the Sonic hedgehog pathway. *Development* **127**, 2133-2142.

**Forsberg, H., Crozet, F. and Brown, N. A.** (1998). Waves of mouse Lunatic fringe expression, in four-hour cycles at two-hour intervals, precede somite boundary formation. *Curr. Biol.* **8**, 1027-1030.

**Gegonne, A., Tai, X., Zhang, J., Wu, G., Zhu, J., Yoshimoto, A., Hanson, J., Cultraro, C., Chen, Q.-R., Guinter, T. et al.** (2012). The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells. *Mol. Cell. Biol.* **32**, 1984-1997.

**Goodrich, J. A. and Tjian, R.** (2010). Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nat. Rev. Genet.* **11**, 549-558.

**Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R. and Gruss, P.** (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J.* **10**, 1135-1147.

**Gouti, M., Tsakiridis, A., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V. and Briscoe, J.** (2014). In vitro generation of neuromesodermal progenitors reveals distinct roles for Wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol.* **12**, e1001937.

**Hubaud, A. and Pourquié, O.** (2014). Signalling dynamics in vertebrate segmentation. *Nat. Rev. Mol. Cell Biol.* **15**, 709-721.

**Indra, A. K., Mohan, W. S., Frontini, M., Scheer, E., Messaddeq, N., Metzger, D. and Tora, L.** (2005). TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. *Dev. Biol.* **285**, 28-37.

**Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P. and Tora, L.** (1994). Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* **79**, 107-117.

**Kirschner, D. B., vom Baur, E., Thibault, C., Sanders, S. L., Gangloff, Y.-G., Davidson, I., Weil, P. A. and Tora, L.** (2002). Distinct mutations in yeast TAF(II)25 differentially affect the composition of TFIID and SAGA complexes as well as global gene expression patterns. *Mol. Cell. Biol.* **22**, 3178-3193.

**Krol, A. J., Roellig, D., Dequéant, M.-L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C. and Pourquié, O.** (2011). Evolutionary plasticity of segmentation clock networks. *Development* **138**, 2783-2792.

**Leurent, C., Sanders, S., Ruhlmann, C., Mallouh, V., Weil, P. A., Kirschner, D. B., Tora, L. and Schultz, P.** (2002). Mapping histone fold TAFs within yeast TFIID. *EMBO J.* **21**, 3424-3433.

**Levine, M., Cattoglio, C. and Tjian, R.** (2014). Looping back to leap forward: transcription enters a new era. *Cell* **157**, 13-25.

**Lin, Z., Yang, H., Kong, Q., Li, J., Lee, S.-M., Gao, B., Dong, H., Wei, J., Song, J., Zhang, D. D. et al.** (2012). USP22 antagonizes p53 transcriptional activation by deubiquitinating Sirt1 to suppress cell apoptosis and is required for mouse embryonic development. *Mol. Cell* **46**, 484-494.

**Liu, Y., Zhao, J., Zhang, W., Gan, J., Hu, C., Huang, G. and Zhang, Y.** (2015). lncRNA GAS5 enhances G1 cell cycle arrest via binding to YBX1 to regulate p21 expression in stomach cancer. *Sci. Rep.* **5**, 10159.

**Ma, C., Shi, X., Zhu, Q., Li, Q., Liu, Y., Yao, Y. and Song, Y.** (2015). The growth arrest-specific transcript 5 (GAS5): a pivotal tumor suppressor long noncoding RNA in human cancers. *Tumor Biol.* **37**, 1437-1444.

**Mansouri, A., Yokota, Y., Wehr, R., Copeland, N. G., Jenkins, N. A. and Gruss, P.** (1997). Paired-related murine homeobox gene expressed in the developing sclerotome, kidney, and nervous system. *Dev. Dyn.* **210**, 53-65.

**McGrew, M. J., Dale, J. K., Fraboulet, S. and Pourquié, O.** (1998). The lunatic fringe gene is a target of the molecular clock linked to somite segmentation in avian embryos. *Curr. Biol.* **8**, 979-982.

**Metzger, D., Scheer, E., Soldatov, A. and Tora, L.** (1999). Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. *EMBO J.* **18**, 4823-4834.

**Mohan, W. S., Scheer, E., Wendling, O., Metzger, D. and Tora, L.** (2003). TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol. Cell. Biol.* **23**, 4307-4318.

**Morales, A. V., Yasuda, Y. and Ish-Horowicz, D.** (2002). Periodic Lunatic fringe expression is controlled during segmentation by a cyclic transcriptional enhancer responsive to notch signaling. *Dev. Cell* **3**, 63-74.

**Müller, F., Zaucker, A. and Tora, L.** (2010). Developmental regulation of transcription initiation: more than just changing the actors. *Curr. Opin. Genet. Dev.* **20**, 533-540.

**Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. R.** (2002). *Manipulating the Mouse Embryo*, 3rd edn. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

**Neidhardt, L. M., Kispert, A. and Herrmann, B. G.** (1997). A mouse gene of the paired-related homeobox class expressed in the caudal somite compartment and in the developing vertebral column, kidney and nervous system. *Dev. Genes Evol.* **207**, 330-339.

**Niederreither, K., McCaffery, P., Dräger, U. C., Chambon, P. and Dollé, P.** (1997). Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. *Mech. Dev.* **62**, 67-78.

**Niwa, Y., Masamizu, Y., Liu, T., Nakayama, R., Deng, C.-X. and Kageyama, R.** (2007). The initiation and propagation of Hes7 oscillation are cooperatively regulated by Fgf and notch signaling in the somite segmentation clock. *Dev. Cell* **13**, 298-304.

**Ott, M. O., Bober, E., Lyons, G., Arnold, H. and Buckingham, M.** (1991). Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo. *Development* **111**, 1097-1107.

**Papadopoulos, P., Gutiérrez, L., Demmers, J., Scheer, E., Pourfarzad, F., Papageorgiou, D. N., Karkoulia, E., Strouboulis, J., van de Werken, H. J. G., van der Linden, R. et al.** (2015). TAF10 interacts with the GATA1 transcription factor and controls mouse erythropoiesis. *Mol. Cell. Biol.* **35**, 2103-2118.

**Perantoni, A. O.** (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* **132**, 3859-3871.

**Pickard, M. and Williams, G.** (2015). Molecular and cellular mechanisms of action of tumour suppressor GAS5 lncRNA. *Genes* **6**, 484-499.

**Pourquié, O.** (2011). Vertebrate segmentation: from cyclic gene networks to scoliosis. *Cell* **145**, 650-663.

**Rocancourt, D., Bonnerot, C., Jouin, H., Emerman, M. and Nicolas, J. F.** (1990). Activation of a beta-galactosidase recombinant provirus: application to titration of human immunodeficiency virus (HIV) and HIV-infected cells. *J. Virol.* **64**, 2660-2668.

**Saga, Y., Hata, N., Koseki, H. and Taketo, M. M.** (1997). Mesp2: a novel mouse gene expressed in the presegmented mesoderm and essential for segmentation initiation. *Genes Dev.* **11**, 1827-1839.

**Sainsbury, S., Bernecky, C. and Cramer, P.** (2015). Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129-143.

**Sapountzi, V. and Côté, J.** (2011). MYST-family histone acetyltransferases: beyond chromatin. *Cell. Mol. Life Sci.* **68**, 1147-1156.

**Soriano, P.** (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70-71.

**Soutoglou, E., Demény, M. A., Scheer, E., Fienga, G., Sassone-Corsi, P. and Tora, L.** (2005). The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. *Mol. Cell. Biol.* **25**, 4092-4104.

**Spedale, G., Timmers, H. T. M. and Pijnappel, W. W. M. P.** (2012). ATAC-king the complexity of SAGA during evolution. *Genes Dev.* **26**, 527-541.

**Srivastava, D., Thomas, T., Lin, Q., Kirby, M. L., Brown, D. and Olson, E. N.** (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nat. Genet.* **16**, 154-160.

**Tatarakis, A., Margaritis, T., Martinez-Jimenez, C. P., Kouskouti, A., Mohan, W. S., Haroniti, A., Kafetzopoulos, D., Tora, L. and Talianidis, I.** (2008). Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. *Mol. Cell* **31**, 531-543.

**Timmers, H. T. and Sharp, P. A.** (1991). The mammalian TFIID protein is present in two functionally distinct complexes. *Genes Dev.* **5**, 1946-1956.

**Trowitzsch, S., Viola, C., Scheer, E., Conic, S., Chavant, V., Fournier, M., Papai, G., Ebong, I.-O., Schaffitzel, C., Zou, J. et al.** (2015). Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nat. Commun.* **6**, 6011.

**Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. and Nicolas, J.-F.** (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev. Cell* **17**, 365-376.

**Vannini, A. and Cramer, P.** (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45**, 439-446.

**Ventura, A., Kirsch, D. G., McLaughlin, M. E., Tuveson, D. A., Grimm, J., Lintault, L., Newman, J., Reczek, E. E., Weissleder, R. and Jacks, T.** (2007). Restoration of p53 function leads to tumour regression in vivo. *Nature* **445**, 661-665.

**Vincent, S. D., Dunn, N. R., Sciammas, R., Shapiro-Shalef, M., Davis, M. M., Calame, K., Bikoff, E. K. and Robertson, E. J.** (2005). The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse. *Development* **132**, 1315-1325.

**Vincent, S. D., Mayeuf-Louchart, A., Watanabe, Y., Brzezinski, J. A., Miyagawa-Tomita, S., Kelly, R. G. and Buckingham, M.** (2014). Prdm1 functions in the mesoderm of the second heart field, where it interacts genetically with Tbx1, during outflow tract morphogenesis in the mouse embryo. *Hum. Mol. Genet.* **23**, 5087-5101.

**Voss, A. K., Thomas, T., Petrou, P., Anastassiadis, K., Schöler, H. and Gruss, P.** (2000). Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. *Development* **127**, 5449-5461.

**Wang, L. and Dent, S. Y. R.** (2014). Functions of SAGA in development and disease. *Epigenomics* **6**, 329-339.

**Weake, V. M., Dyer, J. O., Seidel, C., Box, A., Swanson, S. K., Peak, A., Florens, L., Washburn, M. P., Abmayr, S. M. and Workman, J. L.** (2011). Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation. *Genes Dev.* **25**, 1499-1509.

**Wittler, L., Shin, E.-H., Grote, P., Kispert, A., Beckers, A., Gossler, A., Werber, M. and Herrmann, B. G.** (2007). Expression of Msgn1 in the presomitic mesoderm is controlled by synergism of WNT signalling and Tbx6. *EMBO Rep.* **8**, 784-789.

**Xu, W., Edmondson, D. G., Evrard, Y. A., Wakamiya, M., Behringer, R. R. and Roth, S. Y.** (2000). Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat. Genet.* **26**, 229-232.

**Yamauchi, T., Yamauchi, J., Kuwata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K. and Nakatani, Y.** (2000). Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc. Natl. Acad. Sci. USA* **97**, 11303-11306.

**Yoon, J. K., Moon, R. T. and Wold, B.** (2000). The bHLH class protein pMesogenin1 can specify paraxial mesoderm phenotypes. *Dev. Biol.* **222**, 376-391.

**Zeller, R., López-Ríos, J. and Zuniga, A.** (2009). Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nat. Rev. Genet.* **10**, 845-858.

**Zhou, H., Grubisic, I., Zheng, K., He, Y., Wang, P. J., Kaplan, T. and Tjian, R.** (2013a). Taf7l cooperates with Trf2 to regulate spermiogenesis. *Proc. Natl. Acad. Sci. USA* **110**, 16886-16891.

**Zhou, H., Kaplan, T., Li, Y., Grubisic, I., Zhang, Z., Wang, P. J., Eisen, M. B. and Tjian, R.** (2013b). Dual functions of TAF7L in adipocyte differentiation. *Elife* **2**, e00170.

**Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L. and Washburn, M. P.** (2006). Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **5**, 2339-2347.

## Supplementary Figures

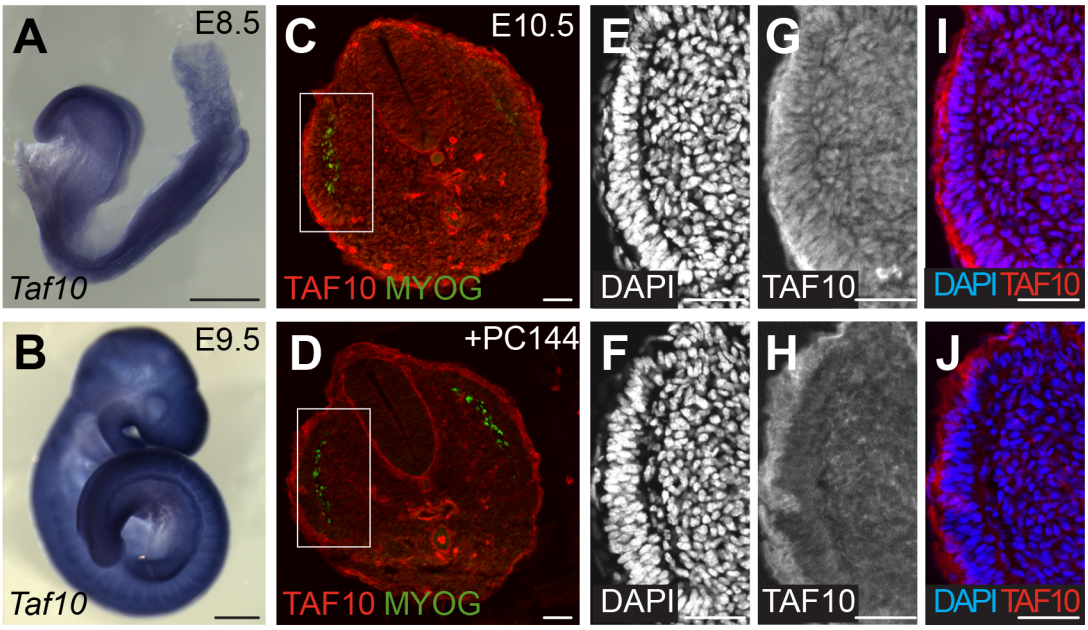



**Fig. S1. TAF10 is expressed in the nuclei of the embryo.** (A-B) Whole-mount *in situ* hybridization of *Taf10* at E8.5 (A) and E9.5 (B). (C-J) Co-immunolocalization of TAF10, Myogenin and DAPI in E10.5 tail transverse sections. (C-D) Colocalization of TAF10 (red) and Myogenin (MYOG, green). (E-J) DAPI (E,F), TAF10 (G,H) and merge (I,J) magnifications corresponding to the boxes indicated in C and D. (D,F,H,J) Competition with the PC144 peptide used to raise the anti-TAF10 antibody. Nuclear signal of TAF10 (D,H,J) is abolished without affecting the Myogenin signal (D). The non nuclear signal that persists after peptide competition is not specific. Scale bars in A-B and C-J represent 500 $\mu$m and 50 $\mu$m, respectively.

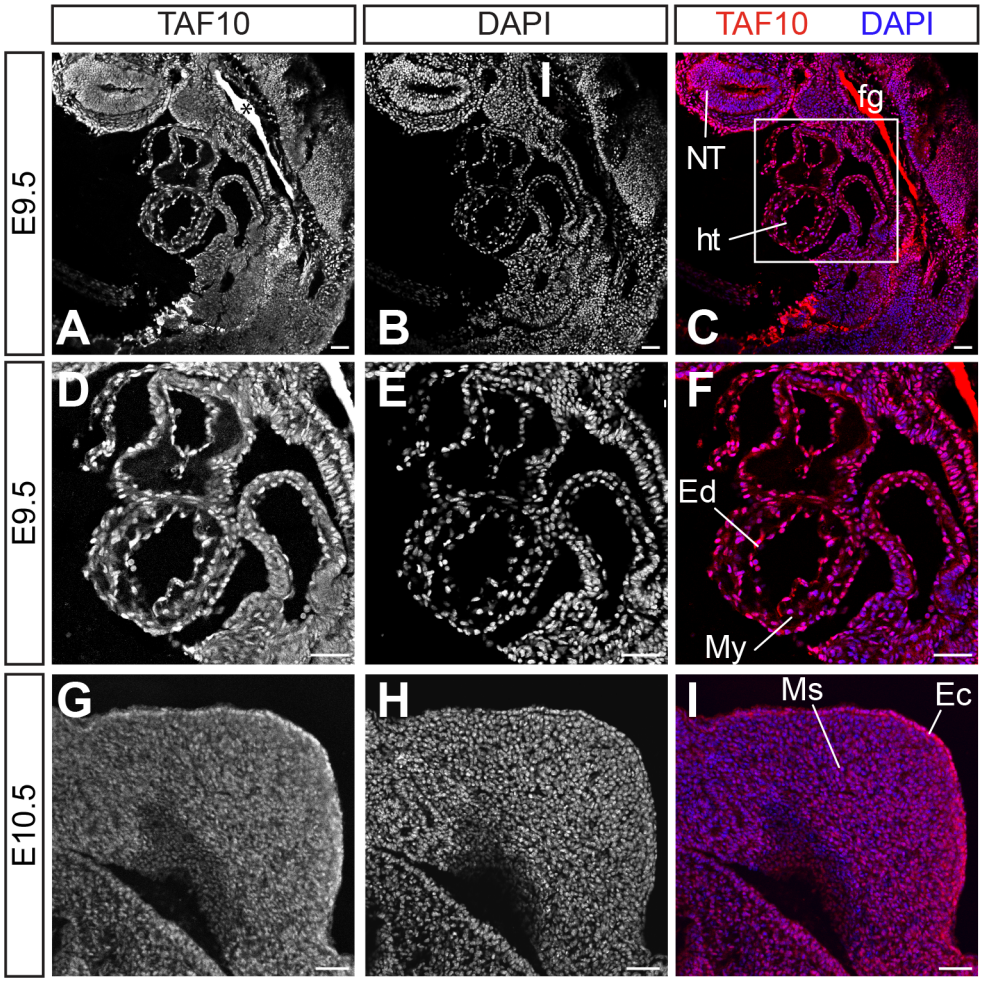



**Fig. S2. TAF10 is ubiquitously expressed in the embryo.** (A-I) Immunolocalization of TAF10 (A,D,G) and DAPI (B,E,H) on wild type embryo sections at E9.5 (A-F) and E10.5 (G-I). (A-C) Sagittal section at the level of the anterior part of the embryo. The asterisk marks the trapping of the secondary antibody in the foregut pocket. (D-F) is a magnification of the region indicated in C, focusing on the heart. (G-I) is a section at the level of the limb bud. NT; neural tube, ht; heart, fg; foregut; Ed; endocardium, My; myocardium, Ms; mesenchyme, Ec; ectoderm. Scale bars represent 50 $\mu$m.

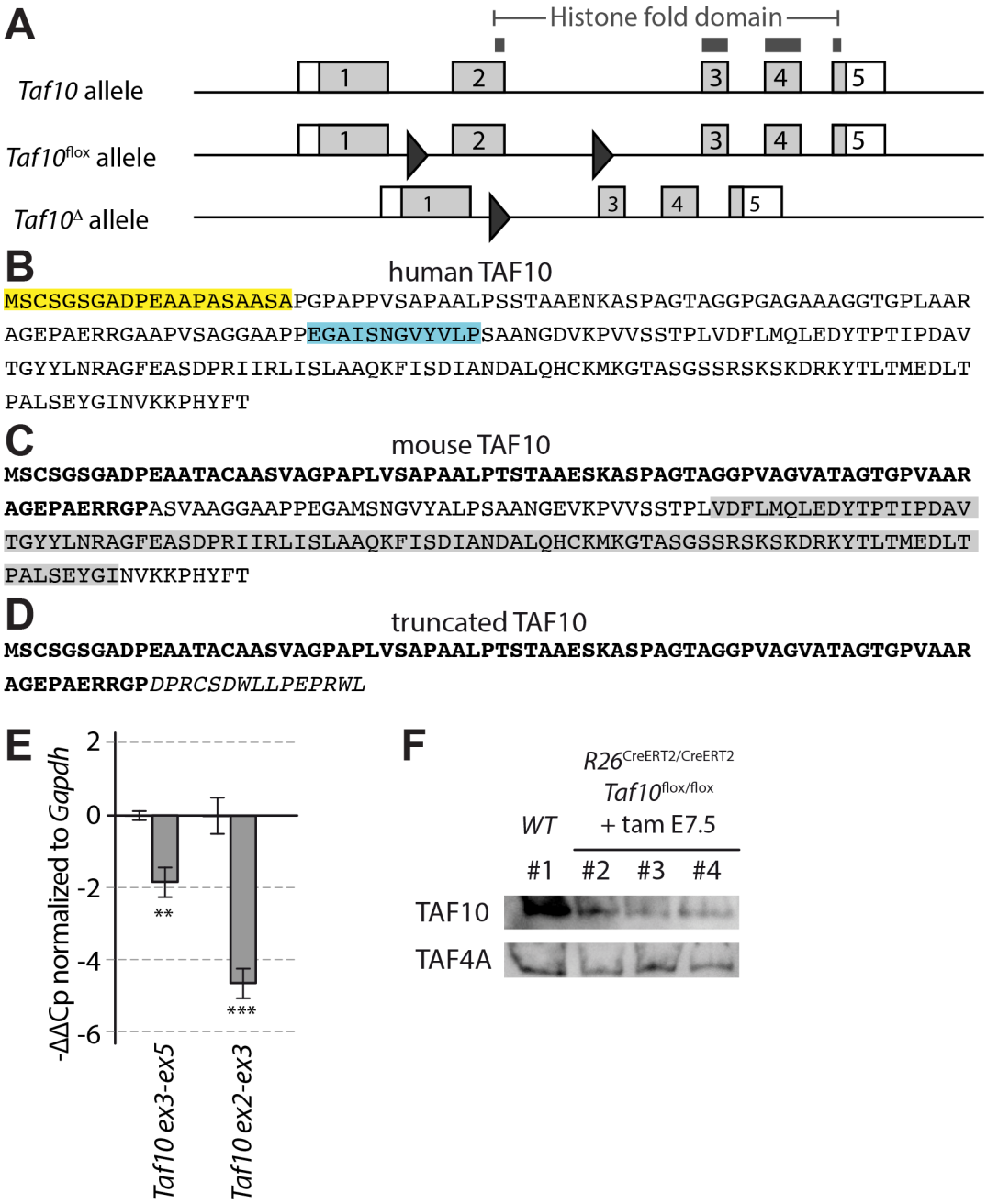



**Fig. S3. Deletion of *Taf10*.** (A) Strategy of the generation of the *Taf10* deletion using the *Taf10*$^{flox}$ allele (Mohan et al., 2003). The exon 2 is deleted upon Cre expression. The coding sequences of the Histone Fold Domain (HFD) is highlighted by dark grey bars. The deleted allele can theoretically produce a truncated protein that does not contain the HFD, coded by exon 1 and but exon 3 that is out of frame. (B) Protein sequence of the human TAF10: the sequence of the peptides used to raise the 23TA1H8 (yellow) and 6TA2B11 (blue) anti-TAF10 antibodies (C,D) Protein sequence of the murine wild type TAF10 (C) and of the truncated protein (D) potentially present after deletion. Coding sequences of exon 1 are indicated in bold characters. The new extra 15 amino-acids encoded by exon 3 are indicated in italics. The HFD is highlighted in grey. (E) RT-qPCR analysis from tail tips of E9.25 control (white) and *TCre*/+;*Taf10*$^{flox/flox}$ mutant (grey) tail tips. *Taf10 ex3-ex5* amplifies a sequence that is shared by the wild type and the deleted transcripts whereas *Taf10 ex2-ex3* amplifies a sequence only present in wild type transcript. -ΔΔCp are normalized to *Gapdh*. **; p-value <0.01, ***; p-value <0.001 (n=4 for *Taf10 ex2-ex3* and n=2 for *Taf10 ex3-ex5*, Aspin Welch corrected Student's t-test). The error bars indicate s.e.m. (F) Anti-TAF10 and anti-TAF4 western blot analysis of whole cell extract from E8.5 *R26*$^{CreERT2/CreERT2}$;*Taf10*$^{flox/flox}$ embryos, induced at E7.5 by tamoxifen injection at E7.5.

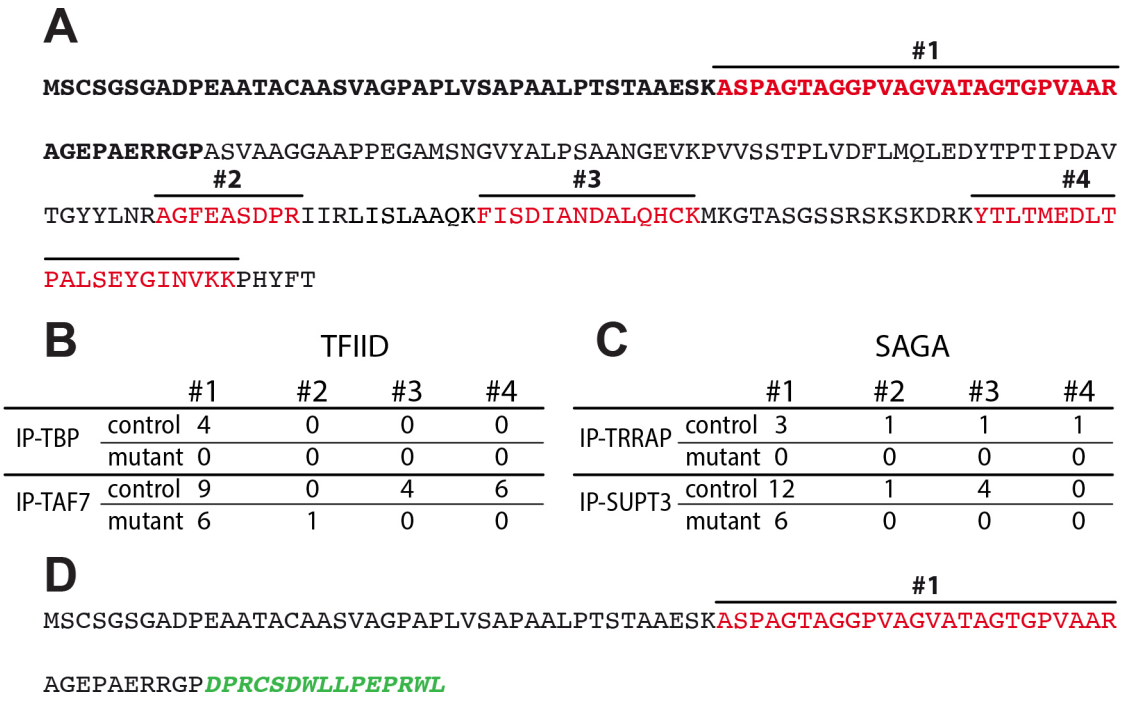



**Fig. S4. Abnormal distribution of the TAF10 MS peptides detected in *R26Cre;Taf10* mutant embryos.** (A) Localization of the MS peptides (#1 to #4) on the sequence of the full-length TAF10 protein. The peptides are indicated in red letters. The coding sequence corresponding to the first exon is highlighted in bold. (B-C) Number of TAF10 detected MS peptides in control and *R26Cre;Taf10* mutant embryos in each TFIID (B) and SAGA (C) IPs. (D) Localization of the MS peptide #1 (red) on the sequence of putative TAF10 truncated protein. The 15 extra amino-acids coded by exon 3 (not in frame) are indicated in green italics.

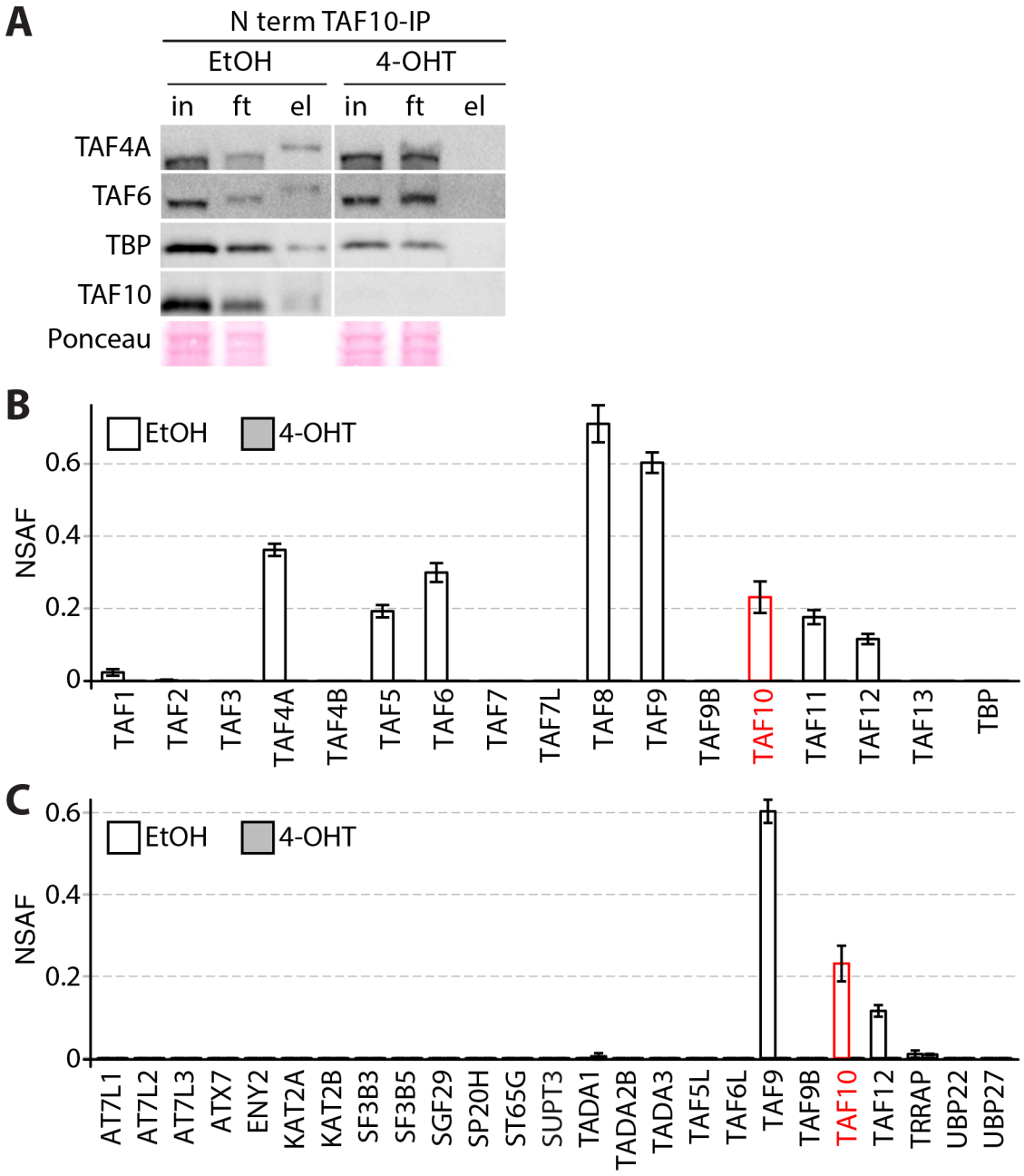



**Fig. S5. The potential truncated TAF10 peptide is not able to form a TFIID complex in *Taf10* mutant ES cells** (A) Western blot analysis of TBP, TAF4A, TAF6, TBP and full length TAF10 protein of input (in), flowthrough (ft) and elution (el) from Nterm-TAF10-IP from control (EtOH) and mutant (4-OHT) *Taf10* conditional mutant ES cells. (B,C) NSAF values for TFIID (B) and SAGA (C) complexes subunits of Nterm-TAF10-IP from control (EtOH) and Taf10 mutant (4-OHT) ES whole cell extracts. The control (EtOH) and mutant (4OHT) conditions are indicated in white and grey, respectively. 4-OHT; 4-hydroxy-tamoxifen.

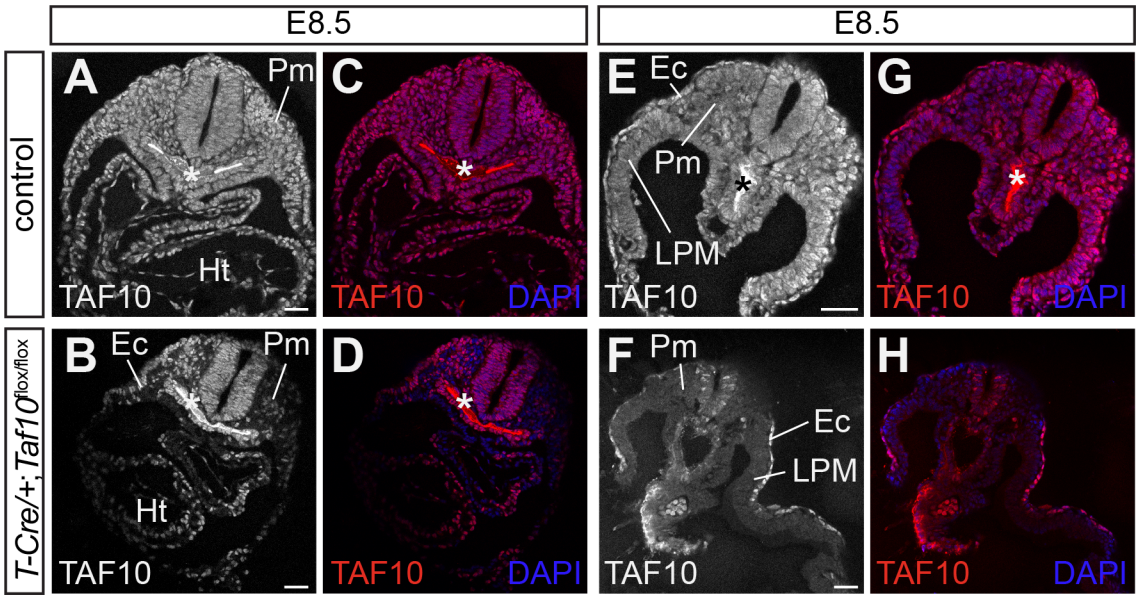



**Fig. S6. TAF10 depletion is already effective in the LPM and in the paraxial mesoderm at E8.5.** (A-H) TAF10 immunocolocalization on E8.5 transverse wild type (10 somites; A,C,E,G) and *T-Cre;Taf10*$^{flox/flox}$ mutant (8 somites; B,D,F,H) embryos, at the heart (A-D) and at the posterior somites (E-H) levels. Pm; paraxial mesoderm, Ec; ectoderm, LPM; lateral plate mesoderm, Ht; heart. The asterisk (*) indicates background due to secondary antibody trapping in the endoderm. Scales bars represent 50 $\mu$m.

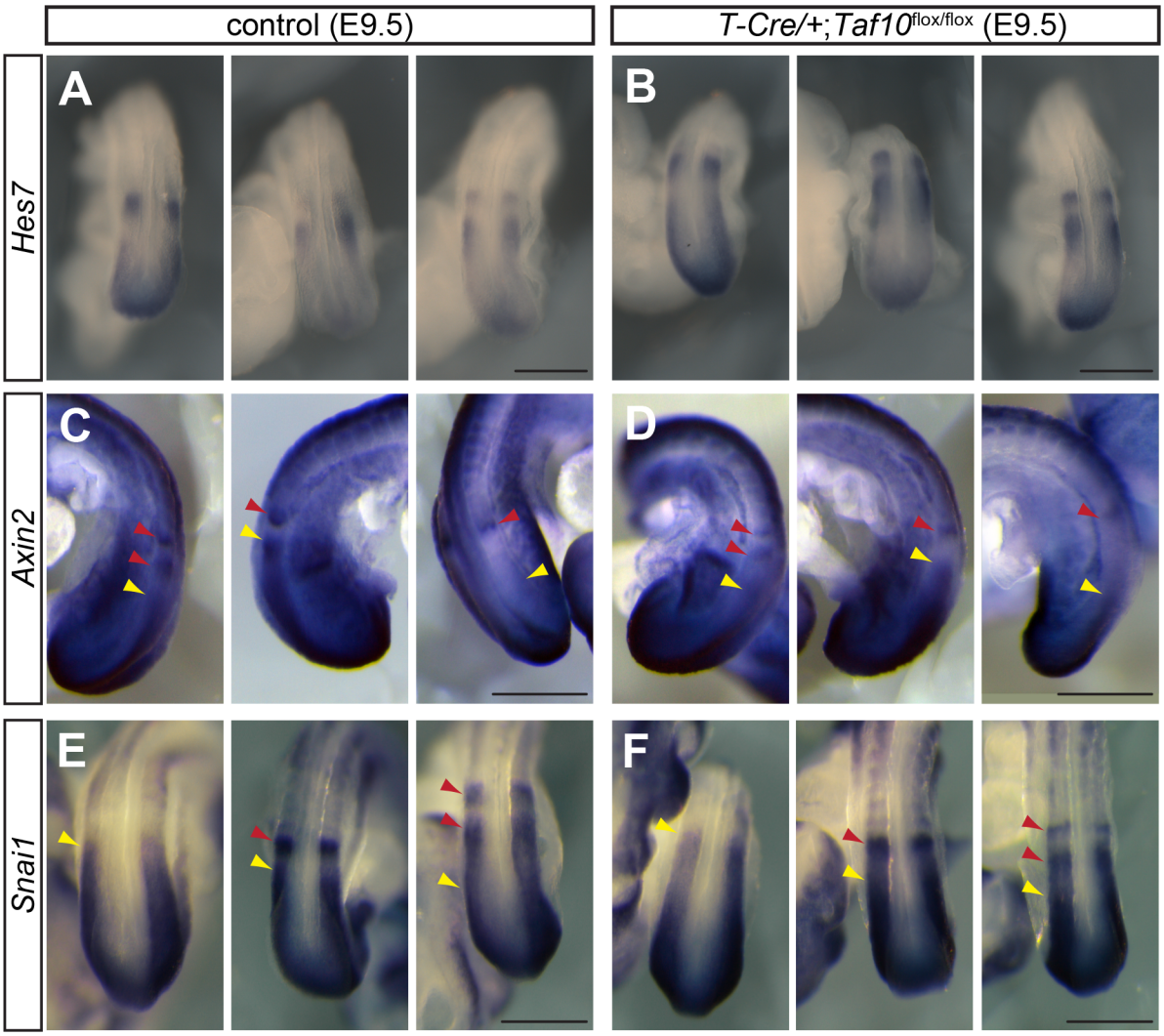



**Fig. S7. Expression of the cyclic genes in the absence of TAF10.** (A-F) Whole-mount *in situ* hybridization of E9.5 control (A,C,E,) and *T-Cre*/+;*Taf10*$^{flox/flox}$ mutant (B,D,F) embryos using *Hes7* (A,B), *Axin2* (C,D) and *Snai1* (E,F). For each probe, 3 different phases expression pattern are displayed, bands are highlighted by red arrows and the anterior limit of the posterior domain yellow arrows in C-F. The absence of TAF10 in the PSM does not affect the cyclic expression of Notch pathway (*Hes7*), Wnt (*Axin2*) or FGF (*Snai1*) pathways. Scale bars represent 500 $\mu$m.

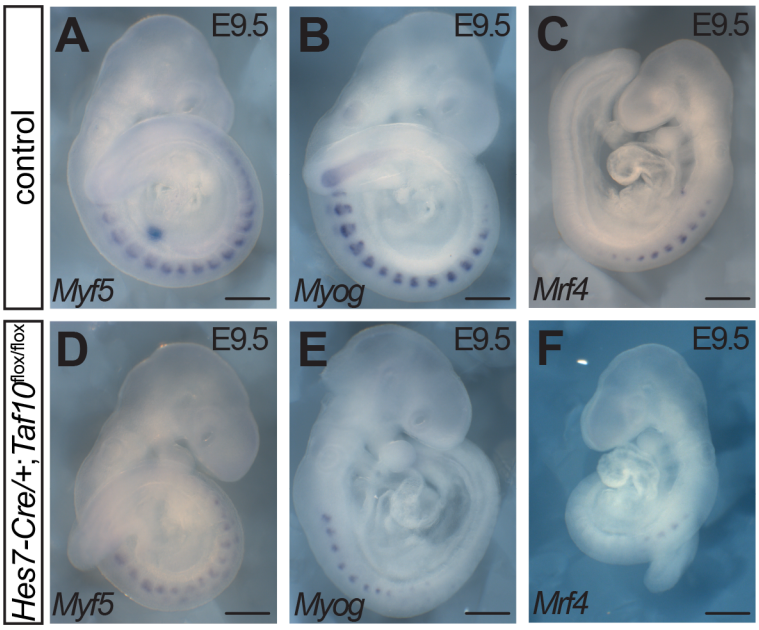



**Fig. S8. Delayed myogenesis in *Hes7-Cre*/+;*Taf10*$^{flox/flox}$ mutant embryos.** (A-F) Whole-mount *in situ* hybridization of E9.5 control (A-C) and *Hes7-Cre*/+;*Taf10*$^{flox/flox}$ mutant (D-F) embryos using *Myf5* (A,D), *Myog* (B,E) and *Mrf4* (C,F) showing decreased expression of these myogenic markers in the absence of TAF10. Scale bars represent 500 $\mu$m.

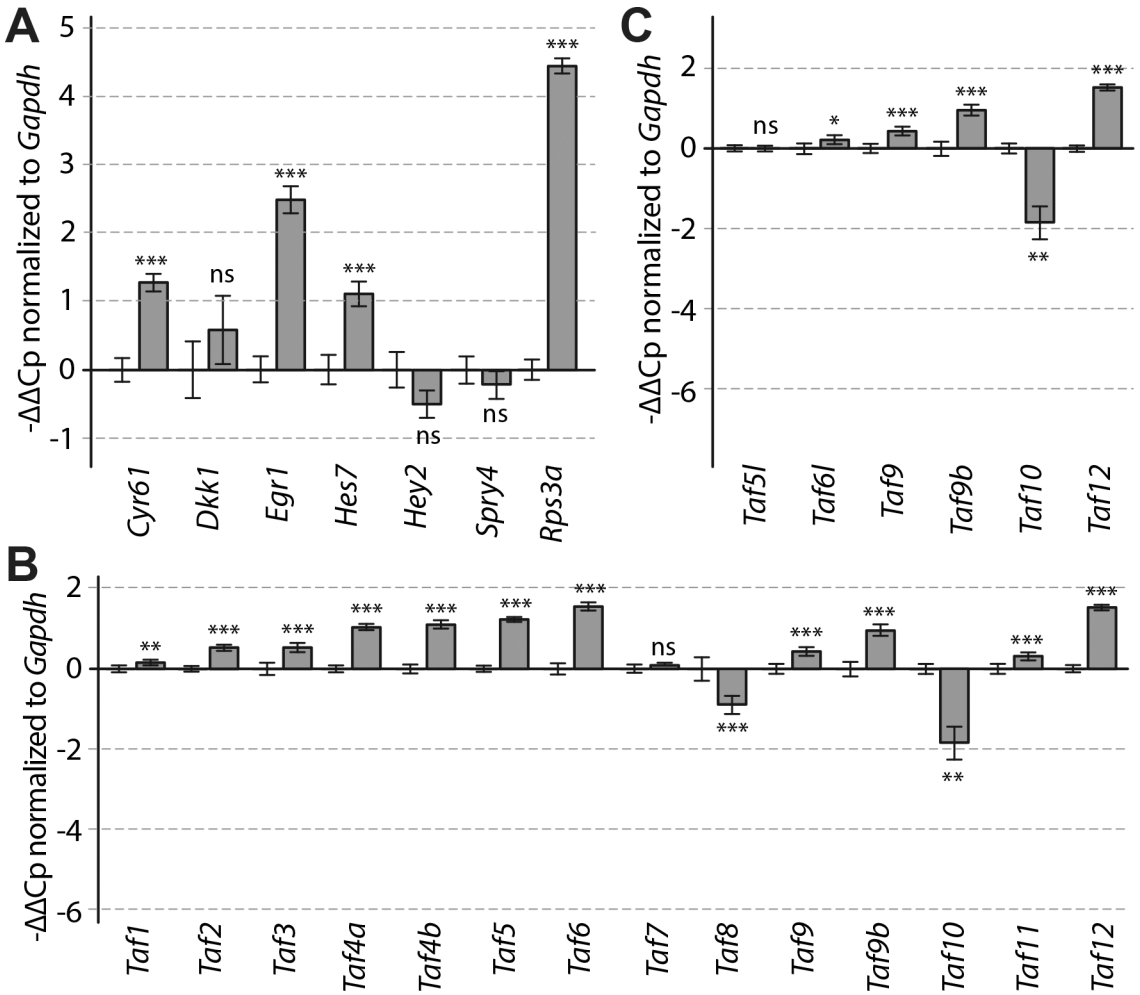



**Fig. S9. Validation of the microarray analysis.** (A-C) RT-qPCR analysis for cyclic genes (A), TFIID-related TAFs (B) and SAGA-related TAFs (C) from tail tips of E9.25 control (white) and *TCre*/+;*Taf10*$^{flox/flox}$ mutant (grey) tail tips. -ΔΔCp are normalized to *Gapdh*. ns; non significant, *; p-value <0.05, **; p-value <0.01, ***; p-value <0.001 (n=4 except for *Taf10* where n=2, Aspin Welch corrected Student's t-test). The error bars indicate s.e.m.

| | **control** | | | |
|---|---|---|---|---|
| | ***IP TBP*** | ***IP TAF7*** | ***IP SUPT3*** | ***IP TRRAP*** |
| PSM* bait | 10/6/7 | 40/37/33 | 19/18/20 | 62/65/71 |
| SAF bait | 0.02421859 | 0.107611082 | 0.046817713 | 0.025712615 |
| total SAF | 3.235144528 | 4.882771593 | 3.73445435 | 7.255567912 |
| total protein detected | 1185 | 1401 | 1213 | 1670 |

| | **mutant** | | | |
|---|---|---|---|---|
| | ***IP TBP*** | ***IP TAF7*** | ***IP SUPT3*** | ***IP TRRAP*** |
| PSM* bait | 7/7/6 | 32/24/29 | 8/15/9 | 67/67/52 |
| SAF bait | 0.021105319 | 0.082964463 | 0.026096693 | 0.02418102 |
| total SAF | 7.205885749 | 3.567467454 | 4.777603295 | 8.991233196 |
| total protein detected | 1525 | 1137 | 1054 | 1613 |

**Table S1. Mass spectrometry results for the different IPs.**
PSM*; peptide spectrum match, SAF; spectral abundance factor

**Table S2. Differentially expressed genes in PSM of E9.5 *T-Cre;Taf10* mutant versus control embryos**

Click here to Download Table S2

| *antibody* | *type* | *reference* |
|---|---|---|
| anti-H3 | rabbit polyclonal | Abcam 1791 |
| anti-TAF10 | mouse monoclonal | 6TA2B11 (Mohan et al., 2003) |
| antiNterm-TAF10 | mouse monoclonal | 23TA1H8 (Wieczorek et al., 1998) |
| anti-TBP | mouse monoclonal | 3TF13G3 (Brou et al., 1993) |
| anti-TAF4 | mouse monoclonal | 32TA2B9 (Perletti et al., 2001) |
| anti-TAF5 | mouse monoclonal | 1TA1C2 (Jacq et al., 1994) |
| anti-TAF6 | mouse monoclonal | 22TA2A1 (Wieczorek et al., 1998) |
| anti-TRRAP | mouse monoclonal | 2TRA1B3 (Nagy et al., 2010) |
| anti-GST | mouse monoclonal | 15TF21D10 (Nagy et al., 2010) |
| anti-Myogenin | rabbit polyclonal | SC-576 Santa Cruz |
| anti-Rabbit IgG, Alexa Fluor® 488 conjugate | goat polyclonal | Molecular Probes A-11008 |
| anti-Mouse IgG Alexa Fluor® 546 conjugate | goat polyclonal | Molecular Probes A-11018 |
| anti-Rabbit IgG Peroxydase conjugate | goat polyclonal | Jackson ImmunoResearch 111-035-144 |
| anti-Mouse IgG Peroxydase conjugate | goat polyclonal | Jackson ImmunoResearch 111-036-071 |

**Table S3. List of antibodies.**

**Table S4. Primer sequences.**

Click here to Download Table S4

## Supplementary Material and Methods

### *Mouse lines*

Tg(*T-Cre*) (Perantoni, 2005), Tg(*Hes7-Cre*) (Niwa et al., 2007), Tg(*Luvelu*) (Aulehla et al., 2008), *R26*$^{CreERT2}$ (Ventura et al., 2007), *R26*$^{R}$ (Soriano, 1999) and *Taf10*$^{flox}$ (Mohan et al., 2003) lines have already been described.

### *Generation of antibodies*

The rabbit polyclonal anti-SUPT3 (3118), anti-TAF7 (3475) and anti-TAF8 (3478) have been generated at the IGBMC antibody facility, with purified proteins. The first 285 amino-acids of human SUPT3 fused to a His tag were produced in BL21DE3 bacteria and purified with Ni-NTA beads (Qiagen). Whole protein cDNAs for human TAF7 and mouse TAF8 were produced via baculovirus in SF9 cells. For TAF8, infected SF9 cell pellet was boiled and resolved on a 10% SDS PAGE gel, then the TAF8 corresponding band was cut, frozen in liquid nitrogen and crushed. Resulting powder was directly injected into rabbits. For SUPT3 and TAF7, the purified proteins were injected into rabbits directly. The resulting sera were then purified using Affigel (Biorad) coupling followed by Poly-Prep columns (Biorad) purification against the TAF7 protein, the first 285 amino-acids SUPT3 or TAF8-TAF10 coupled protein to purify anti-TAF7, anti-SUPT3 and anti-TAF8 antibodies, respectively.

### *Immunoprecipitations* from *R26*$^{CreERT2/R}$;*Taf10*$^{flox/flox}$ embryos

Pooled lysates from control and mutant embryos, respectively, were split in 4. IPs were performed in two series: first, anti-GST, anti-TBP, anti-TAF7 and anti-TRRAP (IP mock, IP TBP #1, IP TAF7 #1 and IP TRRAP #1). For the second series, flow through (FT) was collected after an overnight incubation, and used as inputs for the second IPs of the other complexes with fresh Dynabeads coated with fresh antibodies overnight at 4°C (IP TAF7 #2 from FT GST #1, IP TBP #2 from FT TRRAP #1, IP SUPT3 #2 from FT TBP #1 and IP TRRAP #2 from FT TAF7 #1). IP

SUPT3 #2, IP TRRAP #1, IP TAF7 #2 and IP TBP #1 that yielded the highest number of peptides for the bait were conserved for the study.

**Immunoprecipitations from *R26*$^{CreERT2/R}$;*Taf10*$^{flox/flox}$ mouse ES cells**

*R26*$^{CreERT2/R}$;*Taf10*$^{flox/flox}$ mouse ES cells (mES) were derived from *R26*$^{CreERT2/R}$;*Taf10*$^{flox/flox}$ E3.5 blastocysts (Vincent SD, unpublished data), maintained on gelatin 0.1% in PBS-coated (PAN BIOTECH) feeder-free culture plates at 37 °C in 5% CO2, in a maintenance medium composed of DMEM supplemented with 15% fetal bovine serum (FBS, Millipore), penicillin, streptomycin, 2 mM L-glutamine, 0.1 mM non essential amino acids, 0.1% β- mercaptoethanol, 1,500 U/mL LIF and 2i inhibitors (Ying et al., 2008) (CHIR99021 3$\mu$M and PD0325901 1$\mu$M (Axon medchem)). Mouse ES cells were treated with 0.01% EtOH (control) or 100nM 4-OHT (SIGMA) (mutant) for 4 days.

After 2 PBS washes at 4°C, the cells were then scrapped, collected after 20817 rcf centrifugation for 15 min at 4°C and lysed in 10% glycerol, 20 mM Hepes (pH7), 0.35 M NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, 0.1% Triton X-100 with protease inhibitor cocktail (PIC, Roche) on ice. Lysates were treated 3 times with pestle stroke followed by 3 liquid nitrogen freezing-thaw cycles. Lysates were centrifuged at 20817 rcf for 15 min at 4°C and the supernatants were used directly for IPs or stored at -80°C for western blots.

Four mg inputs were incubated with Protein G Sepharose beads (SIGMA) coated with the anti N terminal TAF10 (23TA1H8) mouse monoclonal antibody (Wieczorek et al., 1998) overnight at 4°C. Immunoprecipitated proteins were washed twice 5 min with 500 mM KCl buffer (25 mM Tris-HCl HCl (pH7), 5 mM MgCl2, 10% glycerol, 0.1% NP40, 2 mM DTT, 100 mM KCl and PIC (Roche)) and eluted with 0.1 M glycine (pH2.8) for 5 min three times. Elution fractions were neutralized with 1.5 M Tris-HCl (pH8.8).

***Mass spectrometry analyses and NSAF calculations***

Samples were TCA precipitated, reduced, alkylated and digested with LysC and Trypsin at 37°C overnight. After C18 desalting, samples were analyzed using an Ultimate 3000 nano-RSLC (Thermo Scientific, San Jose, California) coupled in line with an linear trap Quadrupole (LTQ)-Orbitrap ELITE mass spectrometer via a nano-electrospray ionization source (Thermo Scientific). Peptide mixtures were loaded on a C18 Acclaim PepMap100 trap column (75 µm inner diameter × 2 cm, 3 µm, 100 Å; Thermo Fisher Scientific) for 3.5 min at 5 µl/min with 2% acetonitrile (ACN), 0.1% formic acid in $H_2O$ and then separated on a C18 Accucore nano-column (75 µm inner diameter × 50 cm, 2.6 µm, 150 Å; Thermo Fisher Scientific) with a 240-min linear gradient from 5% to 50% buffer B (A: 0.1% FA in $H_2O$; B: 80% ACN, 0.08% FA in $H_2O$) followed with 10 min at 99% B. The total duration was set to 280 min at a flow rate of 200 nL/min. Peptides were analyzed by high resolution full MS scan (R240K, from 300 to 1650 m/z range) followed by 20 MS/MS events using data-dependent CID (collision induced dissociation) acquisition.

Proteins were identified by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 1.4 software (Thermo Fisher Scientific) a combined Mus musculus database (Swissprot, release 2015_11, 16730 entries) where 5 sequences of protein of interest (TrEMBL entries) were added. Precursor and fragment mass tolerances were set at 7 ppm and 0.5 Da respectively, and up to 2 missed cleavages were allowed. Oxidation (M) was set as variable modification, and carbamidomethylation (C) as fixed modification. Peptides were filtered with a false discovery rate (FDR) and rank 1: FDR at 5 %, rank 1 and proteins were identified with 1 unique peptide.

Normalized spectral abundance factor (NSAF) (Zybailov et al., 2006) normalized to the bait ($NSAF_{bait}$) were obtained as followed (PSM*; peptide spectrum match, SAF; spectral abundance factor, x; protein of interest):

$$SAF_{(x)} = \frac{PSM^{*}_{x(IP)} - PSM^{*}_{x(IPmock)}}{length(x)}$$

$$NSAF_{(x)} = \frac{SAF_{(x)}}{\sum_{i=1}^{n} SAF_{(xi)}} \times 100$$

$$NSAF_{bait(x)} = \frac{NSAF_{(x)}}{NSAF_{(bait)}}$$

For the Nterm-TAF10 IP analyses, only the NSAF was calculated since the bait was not detected in the mutant conditions.

***Microarrays and statistical analysis***

Total RNA was prepared from 3 replicates (control and mutant), following the recommendations of the manufacturer. Biotinylated single strand cDNA targets were prepared, starting from 150 ng of total RNA, using the Ambion WT Expression Kit (Cat # 4411974) and the Affymetrix GeneChip® WT Terminal Labeling Kit (Cat # 900671) according to Affymetrix recommendations. Following fragmentation and end-labeling, 1.9 µg of cDNAs were hybridized for 16 hours at 45ºC on GeneChip® Mouse Gene 1.0 ST arrays (Affymetrix). The chips were washed and stained in the GeneChip® Fluidics Station 450 (Affymetrix) and scanned with the GeneChip® Scanner 3000 7G (Affymetrix) at a resolution of 0,7 $\mu$m. Raw data (.CEL Intensity files) were extracted from the scanned images using the Affymetrix GeneChip® Command Console (AGCC) version 3.2.

Background correction, quantile normalization and summarization by median polish were performed using RMA (Bioconductor package version 2.14 (R version 3.1.0)). Data were filtered automatically by estimating the 100$^{th}$ lowest value of the data series and setting up a background threshold to 3 times this value and by removing manually all the pseudogenes and the expressed sequences. After filtration, 18064 out of 34760 probesets (51.9%) remained. Statistical analysis was performed using the FCROS package version 1.1 (R version 3.1.0) (Dembélé and Kastner,

2014) that calculates a f value. Differences are considered significant for f value below 0.025 or above 0.0975. Scatter plot and vulcano plots were performed using R software version 3.1.0.

***RT-qPCR and statistical analyses***

Unless specified, primers (Table S4) were designed using Primer3 (www.ncbi.nlm.nih.gov/tools/primer-blast) and validated.

To compare RNA polymerases I and II transcriptions, each Cp values were normalised by dividing each Cp to the mean of all Cp (mutants and controls) for one set of primers. Data were analysed using a Student's *t*-test with an Aspin Welch correction.

For the gene expression analyses from tail tips, -ΔΔCp values were calculated first by normalizing each Cp to the mean of the Cps for *Gapdh*, then by subtracting each ΔCp of the different controls from the ΔCp of the sample of interest for a given gene of interest, therefore generating 2x16 -ΔΔCp values for mutants and controls, for one given gene. Data were analysed using a Student's *t*-test with an Aspin Welch correction. Calculations and graphs were obtained using R (3.1.0).

**References**

**Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M. B., Deng, C., Taketo, M., Lewandoski, M. and Pourquié, O.** (2008). A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nat Cell Biol* **10**, 186–193.

**Brou, C., Chaudhary, S., Davidson, I., Lutz, Y., Wu, J., Egly, J. M., Tora, L. and Chambon, P.** (1993). Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO J* **12**, 489–499.

**Dembélé, D. and Kastner, P.** (2014). Fold change rank ordering statistics: a new method for detecting differentially expressed genes. *BMC Bioinformatics* **15**, 14.

**Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P. and Tora, L.** (1994). Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* **79**, 107–117.

**Mohan, W. S., Scheer, E., Wendling, O., Metzger, D. and Tora, L.** (2003). TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol Cell Biol* **23**, 4307–4318.

**Nagy, Z., Riss, A., Fujiyama, S., Krebs, A., Orpinell, M., Jansen, P., Cohen, A., Stunnenberg, H. G., Kato, S. and Tora, L.** (2010). The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cell Mol Life Sci* **67**, 611–628.

**Niwa, Y., Masamizu, Y., Liu, T., Nakayama, R., Deng, C.-X. and Kageyama, R.** (2007). The initiation and propagation of Hes7 oscillation are cooperatively regulated by Fgf and notch signaling in the somite segmentation clock. *Dev Cell* **13**, 298–304.

**Perantoni, A. O.** (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* **132**, 3859–3871.

**Perletti, L., Kopf, E., Carré, L. and Davidson, I.** (2001). Coordinate regulation of RARgamma2, TBP, and TAFII135 by targeted proteolysis during retinoic acid-induced differentiation of F9 embryonal carcinoma cells. *BMC Mol. Biol.* **2**, 4.

**Soriano, P.** (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21**, 70–71.

**Ventura, A., Kirsch, D. G., McLaughlin, M. E., Tuveson, D. A., Grimm, J., Lintault, L., Newman, J., Reczek, E. E., Weissleder, R. and Jacks, T.** (2007). Restoration of p53 function leads to tumour regression in vivo. *Nature* **445**, 661–665.

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

**Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A.** (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523.

**Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L. and Washburn, M. P.** (2006). Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **5**, 2339–2347.

# X. Biochemical characterization of TAF10-containing complexes

We showed that TAF10 is necessary for the assembly of TFIID and SAGA complexes in the whole embryo. Given that the embryo at E9.5 is a mixture of several tissues, we did not obtain any information regarding potential differences in TFIID composition in specific tissues. Despite our efforts, it was not possible to obtain proteomics data of enough good quality directly from the PSM. Therefore, I developed an alternative approach with the use of mES cells. The goal was to assess the composition of TAF10-containing complexes in mES cells, as well as in other cell types to check if there was any difference, in presence and in absence of TAF10.

## 1. Technical optimization of immuno-precipitation for the TAF10-containing complexes

### a. Antibody validation

Prior characterizing the composition and the assembly of the TAF10-containing complexes, I looked for the best antibodies for immuno-precipitation among the available antibodies in the laboratory. Antibodies against almost all the subunits of TFIID and SAGA are available, but most of them were originally raised against the human proteins. Since there was no guarantee that they could also be suited for immuno-precipitations on mouse samples, in collaboration with Ivanka Kamenova, we first checked whether these antibodies could recognize mouse proteins and pulled town the targeted complexes. For that, I prepared large quantities of whole cell extracts from the mouse T-cell leukemia cell line T29 (cell line described in (Dumortier et al. 2006)). These cells grow in suspension and can be amplified at a large scale, making them very convenient to get sufficient quantities of starting material for the optimization experiments. For each immuno-precipitation, I started with 4 mg of whole cell extracts and I checked by western-blot whether I got of the TFIID and/or SAGA subunits. When known subunits of TFIID and/or SAGA were detected by western-blot, samples were sent for analysis by liquid chromatography coupled with a tandem mass spectrometry (LC-MS/MS). Results of this validation are recapitulated in (**Table 4**), and the proteomics data is presented in (**annexes II-VIII**). The best antibodies were chosen based on their ability to pull-down the targeted complex with a high number of peptides for all its subunits in a specific manner. The specificity of the antibody could be assessed by comparing the number of peptides of a given subunit obtained by immuno-precipitation with this antibody and with an antibody raised

against the glutathion S-transferase protein (corresponding to the mock immuno-precipitation, which was expected to give the lowest number of peptides as possible for TFIID and SAGA subunits). Following these criteria, the best antibodies that were chosen were TBP and TAF7 for pulling-down TFIID, and TRRAP and SUPT3H for pulling-down SAGA. TBP and TRRAP were chosen also because they are part of other transcriptional complexes (SL1/TFIIIB complexes and TIP60/NuA4 complexes respectively) which are not targeted by TAF10 protein depletion and could thus serve as internal controls.

**Table 4: Validation for immuno-precipitations of antibodies raised against TFIID and SAGA subunits.**

| **Antibody against** | **Antibody reference** | **Testing material** | **Quantity of material tested** | **Validation** |
|---|---|---|---|---|
| TAF1 | 2439 (fraction #7). 2440 (fraction #2) | Human & mouse | 4 mg | works for immuno-precipitation from human and mouse samples |
| TAF2 | 3038 | Human & mouse | 4 mg | works for immuno-precipitation from human only |
| TAF3 | 2F5 | Mouse | 4 mg | Does not work for immuno-precipitation from mouse samples |
| TAF4 | | Mouse | 4 mg | does not work efficiently for immuno-precipitation from mouse T29 samples |
| TAF7 | 3475 (fraction #2) | Mouse | 4 mg, 1 mg, 0.7 mg | works for immuno-precipitation from mouse samples |
| TAF8 | 3477/3478 | Human & mouse | 4 mg | works for immuno-precipitation from human and mouse samples |

| | | | | |
|---|---|---|---|---|
| TAF10 | 6TA2B11 | Human & mouse | 4 mg | works for immuno-precipitation from human and mouse samples |
| TBP | 3TF13G3 | Mouse | 4 mg, 1 mg, 0.7 mg | works for immuno-precipitation from mouse samples |
| KAT2A (GCN5) | 5GC2A6 | Mouse | 4 mg | works for immuno-precipitation from mouse samples but not completely specific |
| ATXN7L3 | 1ATX2D7 | Mouse | 4 mg | works for immuno-precipitation from mouse samples but not efficient |
| SUPT20H | 3006 | Mouse | 4 mg | does not work efficiently for immuno-precipitation from mouse samples |
| SUPT3H | 3118 (fraction #2) | Mouse | 4 mg, 1 mg, 0.7 mg | works for immuno-precipitation from mouse samples |
| TRRAP | 2TRR1B3 | Mouse | 4 mg, 1 mg, 0.7 mg | works for immuno-precipitation from mouse samples |

b. Starting material reduction

After having validated antibodies suitable for the immuno-precipitation experiments, I optimized the quantity of starting material needed. I started with 4 mg, which was the quantity routinely used in the laboratory for this kind of experiment, and I tested 1 mg, which could be easily obtained with whole cellular extract prepared from mES cells, and 700 ug which corresponded to the maximum quantity I could get from the embryos I collected during one year. For these quantities of starting material, I was still able to detect most of the subunits of TFIID and SAGA. Results are summarized in (**Table 4**).

c. Processing and analysis of the proteomics data

To analyze the proteomics data, I used the data filtered with the false discovery rate of 5% and using the protein identification based on at least one unique peptide. Analyses based on protein identification with at least two unique peptides gave similar results. To make sure that even the smallest subunits of TFIID and SAGA complexes could be detected, the one unique peptide identification was used. From the data obtained, the normalized spectral abundance factor (NSAF) (Zybailov et al. 2006) was calculated for each protein detected in the sample analyzed. NSAF values indicate the proportion of a protein in the sample mixture, and takes into account the sample-sample variation. In addition, it avoids the bias related to the length of the protein, since longer proteins generate more peptides than shorter proteins. Here, the NSAF of every protein was normalized with the NSAF of the bait protein and calculated as followed (PSMx: peptide spectrum match, SAF: spectral abundance factor, IP: immuno-precipitation):

(1) The ΔPSMx of each peptide, calculated with the PSMx obtained from the mock immuno-precipitation against the GST protein, is normalized with the length of the protein (number of amino acids) which gives the SAF value:

$$SAF_{(x)} = \frac{PSM_{x\,(IP)} - PSM_{x\,(IP\ mock)}}{length\ (x)}$$

(2) The NSAF is calculated as a percentage of the sum of all the SAF values:

$$NSAF_{(x)} = \frac{SAF_{(x)}}{\sum_{i=1}^{n} SAF_{(xi)}} \times\ 100$$

(3) Each NSAF is normalized with the NSAF from the bait:

$$NSAF_{bait\ (x)} = \frac{NSAF_x}{NSAF_{(bait)}}$$

Proteomics data was processed following these calculations and was semi-automatized with a script that I designed with the R software (**annexe I**).

## 2. TFIID and SAGA characterization in different cellular contexts

In order to determine the precise composition of TFIID and to assess potential differences between different cell-types, I compared the data obtained from TAF7 immunoprecipitations between thymocytes (T29), pluripotent cells (mES cells) and the embryo (**figure 15 A**). Very similar results were obtained for T29, mES and the embryo (**figure 15 A**), indicating that the global TFIID complex composition is conserved even in different cellular contexts. In thymocytes, most of the TFIID subunits were detected but the stoichiometry of TAF6 and TAF9 was different since they were found in higher amounts (**figure 15 A**). Note that TAF9b was not detected in thymocytes with this immuno-precipitation, but could be detected in TBP immuno-precipitation (**annexe IV**). This data provides a detailed view of the composition of TFIID in different cellular contexts. The canonical TFIID is found, but it undergoes some changes in its composition.

The TRRAP-immunoprecipitation (**figure 15 B and C**) showed that there are differences between T29 cells, mES cells and the embryo for the TIP60/NuA4 complex (**figure 15 B**). In the embryo, ACL6A, ACTB, H2AZ and H2B1F are more represented than in T29 cells and in mES cells (**figure 15 B**). In mES cells, ACTB was not detected but was detected at day 3 but was detected with the TRRAP immuno-precipitation performed at day 5. In T29 cells, VPS72 subunit was not detected (**figure 15 B**) suggesting that there might be slight differences in some transcriptional complexes between different cell types.

Concerning the SAGA complex, a similar composition was found between T29, mES cells and the embryo, except SUPT3 that not found in T29 cells (**figure 15 C**). SUPT3 was not found neither in the other SAGA immuno-precipitation experiments (**annexe VI**), maybe due to its low abundancy. However, SUPT3H immuno-precipitation pulled-down several SAGA subunits (**annexe VII**) indicating that SUPT3 is present but cannot always be properly detected. GCN5/KAT2A is found in T29, mES cells and the embryo, but PCAF/KAT2B is mainly found in mES cells and not in the embryo (**figure 15 C**). The relative stoichiometry of the SAGA complex is conserved in the three systems analyzed here, except TAF9 which is found in larger amounts in the embryo compared to T29 and mES cells which is different than the situation observed in TFIID (**figure 15 A**).

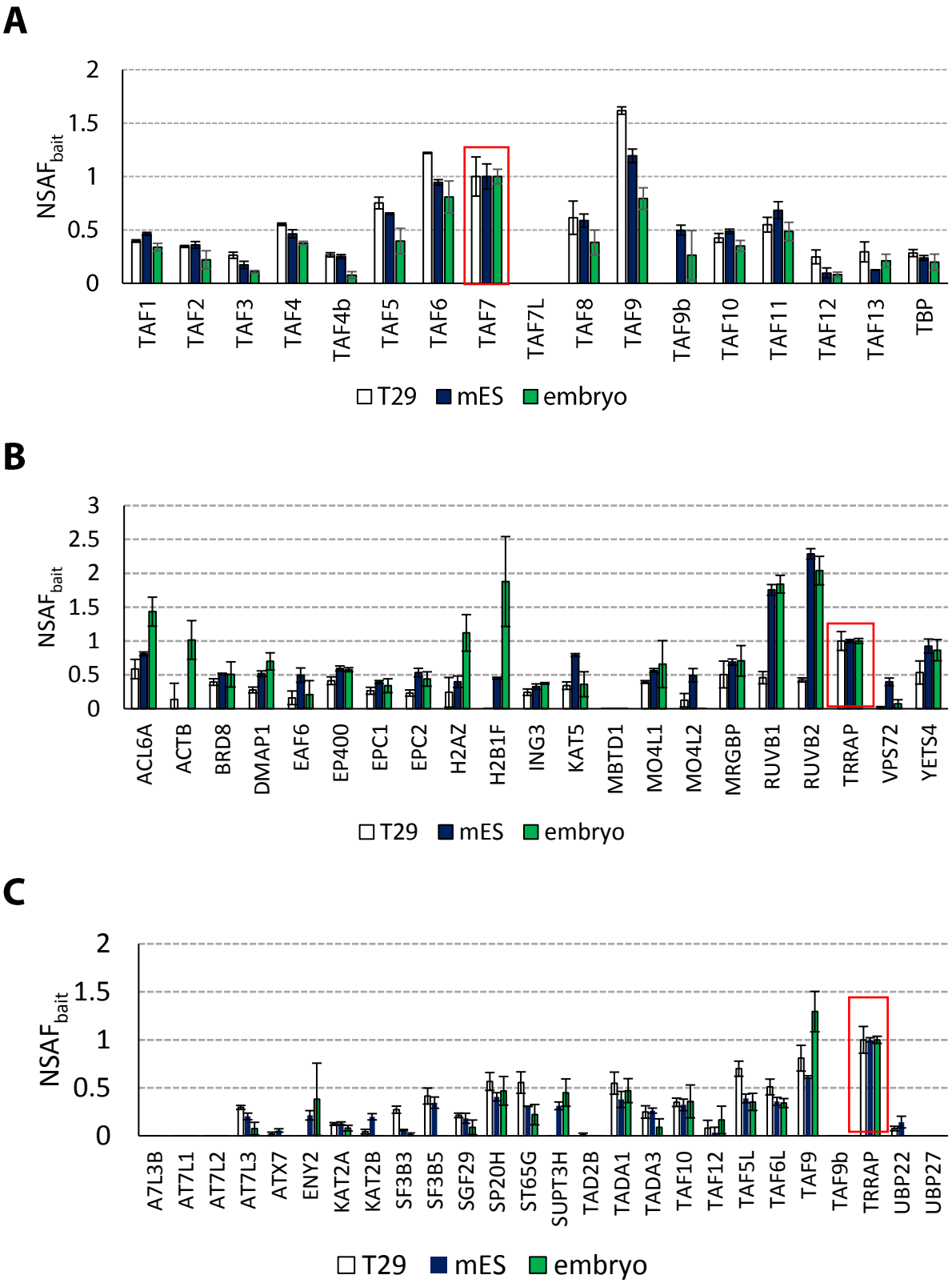



**Figure 15: TAF10-containing complexes characterization in different cellular contexts.** (A) (B) (C) $NSAF_{bait}$ values for (A) TFIID, (B) TIP60/NuA4, (C) SAGA complex subunits of (A) TAF7 IP, (B) (C) TRRAP IP from control whole cell extracts of T29, mES cells (EtOH treated cells collected at day3) or whole E9.5 embryo. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3

Altogether, this data indicates that the composition of the TAF10-containing complexes does not change according the cellular context, at least for the three systems that were used here. Nevertheless, differences concerning the expression level of the subunits of those complexes in T29, mES and the embryo cannot be excluded.

## 3. Phenotype characterization

### a. Experimental workflow

mES cells derived from mouse blastocysts *R26$^{CreERT2/YFP}$; Taf10$^{flox/flox}$; Tg(Mesogenin-YFP/+)* and *R26$^{CreERT2/YFP}$; Taf10$^{flox/flox}$; Tg(Mesogenin-YFP/+)* can be inducibly deleted for *Taf10* upon 4-hydroxy tamoxifen (4OHT) treatment. mES cells were treated with either ethanol (EtOH), the vehicle as a control, or with 4OHT the day after their plating, considered as day 1, for two (day 3) or four days (day 5) (**figure 16 A**). 4OHT was not detrimental for cell growth and viability, as no difference between EtOH- and 4OHT-treated cells was observed in control E14 mES cells (**annexe X**). First, TAF10 depletion was assessed by western-blot from whole cell extracts at day 3 and day 5 (**figure 16 B and C**). At the protein level, TAF10 was nearly completely depleted from day 3 (**figure 16 B**) and no residual trace could be detected at day 5 (**figure 16 C**).

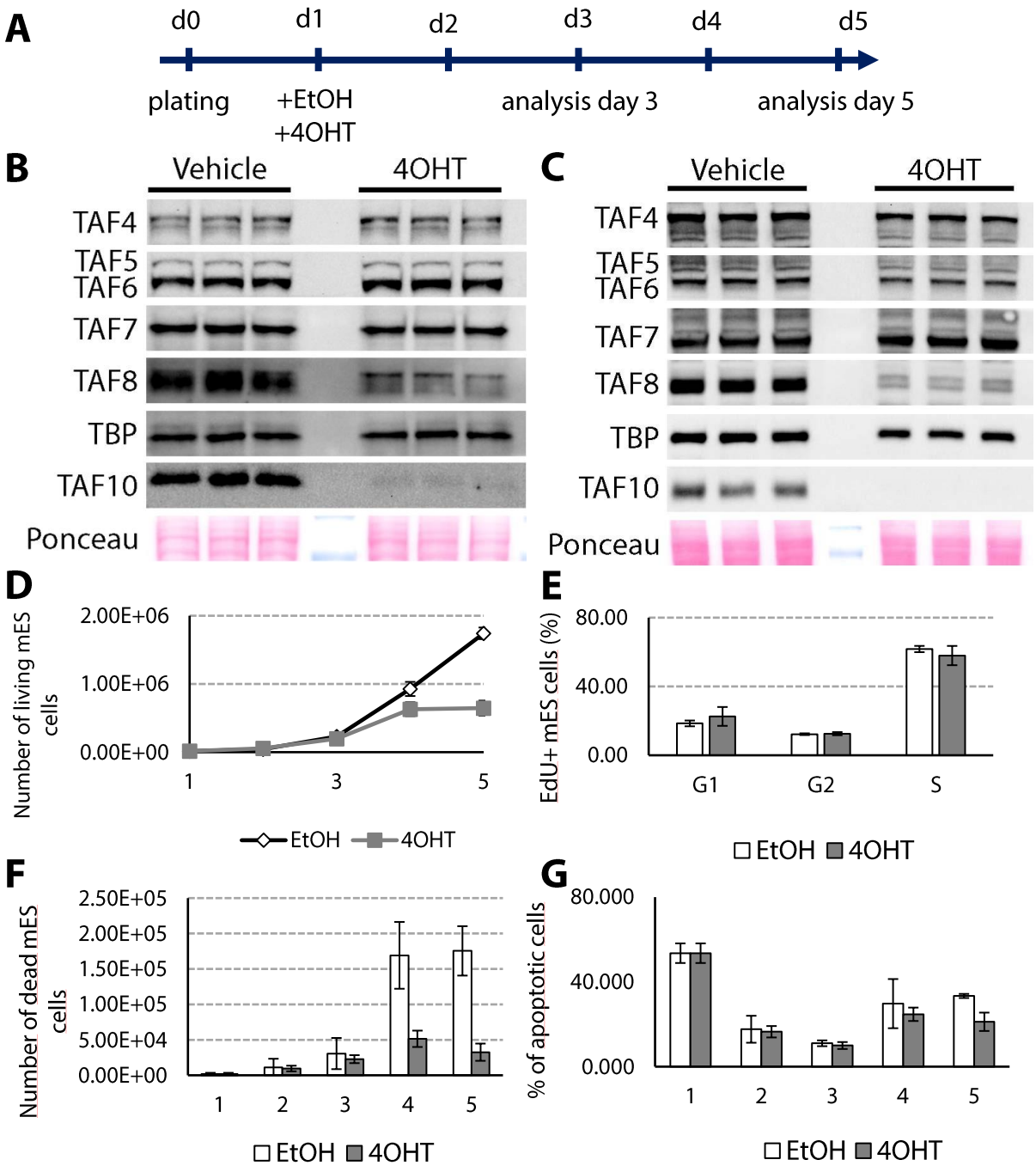



**Figure 16: TAF10 is required for normal mES cell growth.** (A) Experimental workflow. (B) (C) Western-blot analyses for TFIID subunits from mES whole cell extracts treated at day 1 with EtOH (vehicle) or 100 nM 4OHT, collected at (B) day 3 and (C) day 5. (D) Growth curve of mES cells (number of living cells) treated with EtOH (vehicle) or 100 nM 4OHT at day 1. (E) Cell cycle analysis by flow cytometry at day 3 of mES cells cells treated with EtOH (vehicle) or 100 nM 4OHT at day1 and incubated 1 h with 10 mM EdU. (F) Number of dead mES cells treated with EtOH (vehicle) or 100 nM 4OHT at day 1. (G) Percentage of apoptotic mES cells treated with EtOH (vehicle) or 100 nM 4OHT at day 1. Control and mutant conditions are indicated in white and gray, respectively. Error bars indicate s.d. *n*=3

A similar experiment was performed with murine F9 *Taf10*$^{-/-}$ embryonal carcinoma cells (**figure 17**), in which the mouse TAF10 protein is rescued by the doxycycline (Dox) inducible expression of the human *TAF10*, described in (Metzger et al. 1999). These cells were used for comparison with the phenotype of mES cells. F9 cells were cultivated initially in culture in presence of doxycycline to maintain hTAF10 protein expression, and was removed at day 1 (**figure 17 A**). No more hTAF10 protein could be detected from day 4 (**figure 17 B**). Doxycycline was shown not detrimental for cell growth and viability, as no difference between the two conditions was observed in F9 wild-type cells (**annexe XI**).

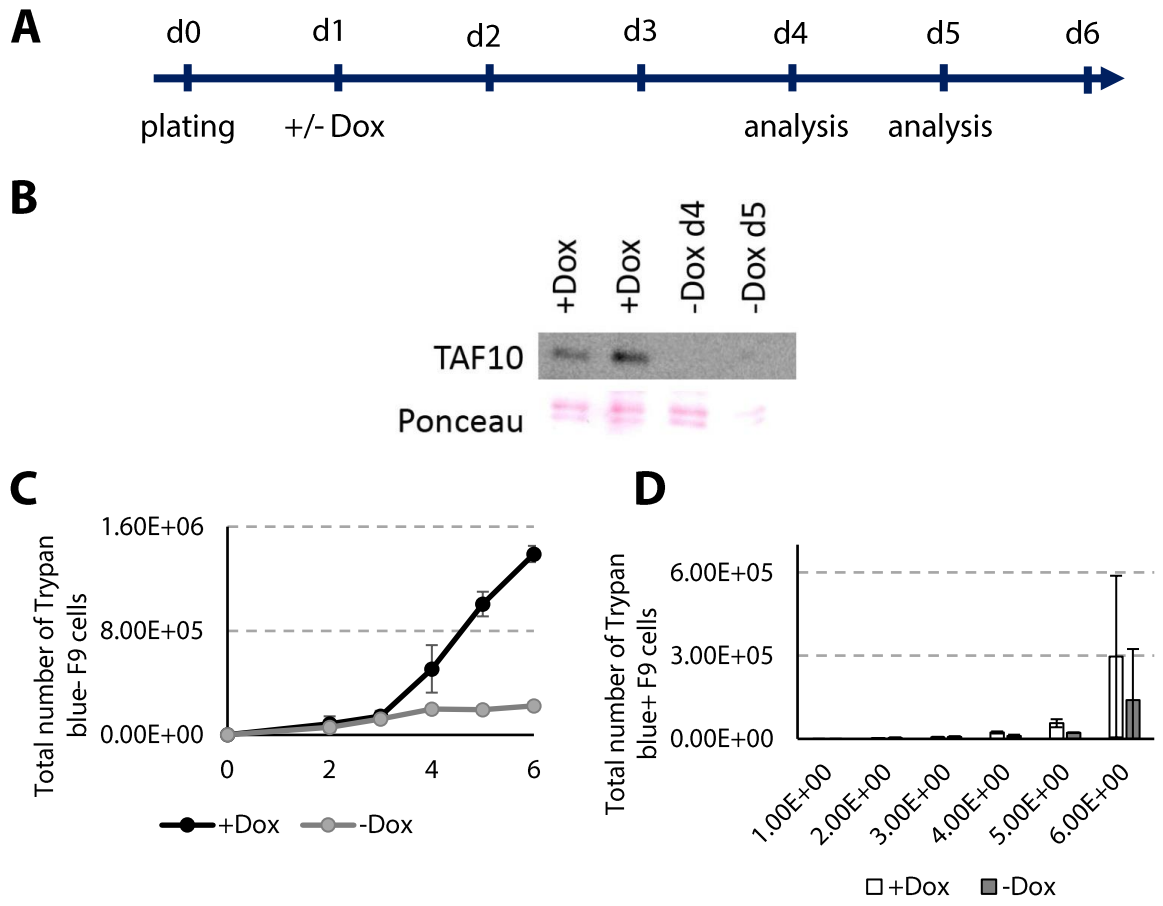



**Figure 17: TAF10 is required for mouse teratocarcinoma F9 cells growth.** (A) Experimental workflow. (B) Western-blot analysis for TAF10 from acidic cellular extracts of F9 cells treated at day 1 with or without doxycyline (Dox) collected at day 4 and day 5. (C) Growth curve of F9 cells treated with or without doxycyline at day 1, and plotted as the number of Trypan blue negative (living) F9 cells. (D) Number of Trypan blue positive (dead) F9 cells treated with or without doxycyline at day 1. Error bars indicate s.d. *n*=3

b. Cellular growth, viability and cell death analyses

In order to check the effect of the depletion of TAF10 on mES cells, an identical number of mES cells was seeded prior treatment with either EtOH or 4OHT. Then, living mES cells were counted every day (**figure 16 D**) and cell viability (**figure 16 F**) was assessed by staining with acridine orange and propidium iodide (which give red fluorescence due to quenching, so all live nucleated cells fluoresce green and all dead nucleated cells fluoresce red). Cellular apoptosis was specifically assessed by the incorporation of a dye (from the APO percentage kit, see Material & methods section) specifically in cells displaying a translocation of phosphatidylserine from the interior to the exterior surface of the membrane. The percentage of apoptotic cells was calculated based on the absorbance measured compared to the absorbance obtained with H202-treated cells (that induces 100 % apoptosis). Until day 3, no difference in the number of living cells was observed between the two conditions (**figure 16 D**). However, from day 4, 4OHT-treated cells already reached a plateau while EtOH-treated cells continued to grow exponentially until day 5 (**figure 16 D**). In parallel, no massive cell death was observed between the two conditions (**figure 16 F**). The increase of apoptosis observed at day 1 (**figure 16 G**) can be explained by the very few number of cells used for the test, that impaired the optimal conditions of the assay (absorbance value too low). Moreover, cell viability assay (**figure 16 F**) showed that from day 4 EtOH-treated cells displayed a higher number of dead cells. This may be due to the fast acidification of the culture medium due to the high confluence reached by EtOH-treated by day 4. The phenotype observed here was confirmed with another mES cell clone which can also be inducibly deleted for *Taf10* upon 4OHT-treatment (**annexe IX**). The growth slow-down was also observed but detected a bit later since this clone was observed to grow faster than the other one in control conditions. No massive cell death was observed for this clone neither, suggesting that *Taf10* deletion impairs cellular proliferation.

In agreement with what was described in (Metzger et al. 1999), F9 cells could not really grow in culture when TAF10 was not rescued anymore by doxycycline expression induction (**figure 17**). No massive cell death could be detected in our conditions (**figure 17 D**). These results indicate a different behavior between different types of cells when *Taf10* is deleted.

c. Cellular proliferation analysis

To check whether the growth arrest phenotype of 4OHT-treated mES cells was due to a proliferation decrease, proliferation was analyzed by measuring the amount of nascent DNA with (EdU). mES cells have been incubated for one hour with EdU to stain newly synthesized DNA, together with propidium iodide to measure the cells in G1 and G2 phases of the cell cycle by flow cytometry (figure 16 E). Only a mild decrease of cells in S phase, together with a slight increase of cells in G1 after 4OHT treatment, was observed compared to EtOH-treated cells. However, no data could be obtained at day 5 because of technical problems. So, it cannot be excluded that cellular proliferation was impaired at day 4 and day 5 that could explain the growth arrest. To support this hypothesis, it has already been shown that F9 carcinomal cells were blocked in G1, 5 days after TAF10 depletion (Metzger et al. 1999).

Together, this data shows that TAF10-depleted mES cells behave differently than *ex-vivo* cells from the *Taf10* null mutant blastocyst inner cell mass or F9 cells (Metzger et al. 1999; Mohan et al. 2003). Since mES cells can survive without TAF10 protein for several days, they represent a very interesting alternative to the embryo for using classical biochemical approaches.

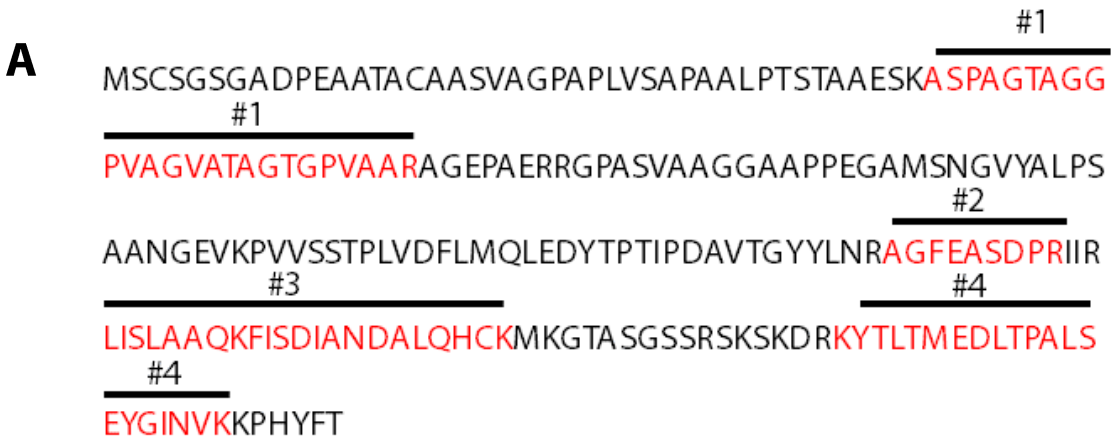
## 4. TAF10 is required for TFIID and SAGA full assembly in pluripotent cells

### a. Residual TAF10 protein detected by mass-spectrometry

In control mES cells, the full length TAF10 protein is detected (**figure 18**), characterized by the detection of four peptides, as shown in control embryos (Bardot et al. 2017). In TFIID complexes in 4OHT-treated mES cells, several TAF10 peptides were detected but mainly the peptide #1 (**figure 18 B**). By contrast, no TAF10 peptides (except one for SUPT3H-immuno-precipitation) were detected in SAGA complexes in 4OHT-treated mES cells (**figure 18 C**). The conditional mutation of *Taf10* deletion deletes exon 2 resulting in an out-of-frame fusion of exon 1 to exon 3. It leads to a premature truncation of TAF10 protein which is expected to produce a truncated N-terminal fragment of TAF10 containing peptide #1, but not the other peptides and not the HFD. The fact that no TAF10 peptides are detected in TRRAP immuno-precipitation suggests that the potential truncated N-terminal peptide remaining in the mutant cannot participate in fully assembled SAGA complexes. In addition, we reported in our paper (Bardot et al. 2017) the TAF10 immuno-precipitation with an antibody raised against the N-terminal part of the protein. It did not pull-down any TFIID subunits in the 4OHT-treated

samples, ruling out that the potential truncated TAF10 protein is integrated into the TFIID complex.

**A**




**B**

| TFIID | | | | | |
|---|---|---|---|---|---|
| | | #1 | #2 | #3 | #4 |
| **IP-TAF7** | EtOH | 6 | 1 | 5 | 14 |
| | 4OHT | 4 | 0 | 1 | 0 |
| **IP-TBP** | EtOH | 6 | 0 | 5 | 15 |
| | 4OHT | 4 | 1 | 0 | 0 |

**C**

| SAGA | | | | | |
|---|---|---|---|---|---|
| | | #1 | #2 | #3 | #4 |
| **IP-TRRAP** | EtOH | 6 | 1 | 4 | 8 |
| | 4OHT | 4 | 0 | 1 | 0 |
| **IP-SUPT3H** | EtOH | 4 | 0 | 4 | 6 |
| | 4OHT | 1 | 0 | 0 | 0 |

**Figure 18: Distribution of TAF10 peptides detected by mass spectrometry in immuno-precipitation experiments.** (A) Localization of the peptides (#1 to #4) on the sequence of the full-length TAF10 protein. The peptides are indicated in red letters. (B) (C) Number of TAF10 detected peptides in EtOH- and 4OHT-treated mES cells at day 5 in each TFIID (B) and SAGA (C) immuno-precipitations.

b. TAF10 is required for TFIID full assembly

Protein levels of TFIID subunits were checked by western-blot in presence and absence of TAF10 (**figure 16 B and C**). The TFIID subunits were stable at the protein level in the absence of TAF10, except TAF8, which is downregulated at the protein level at both day 3 and day 5, in accordance with what was observed in the embryo (Bardot et al. 2017).

In the absence of TAF10, TBP was still able to interact with its non-TFIID partners, which are part of the SL1 and TFIIIB complexes (**figure 19 A**). Contrary to the mutant embryos (Bardot et al. 2017), TBP did not seem to be redistributed to the same extent in RNA Pol I and RNA Pol III complexes, though an increase for TBP in TFIIIB was detected (**figure 19 A**). These results indicate that TAF10 loss does not affect the global integrity of other protein complexes sharing common subunits with TFIID, so TAF10 depletion effect is specific of TFIID and SAGA complexes.

Both TBP and TAF7 immunoprecipitations revealed that all the TFIID subunits were detected in the control samples (EtOH-treated cells), except TAF7L, a TAF paralog, that was not detected (**figure 19 B and C**). When TAF10 was depleted, lower levels of TFIID subunits were detected, and similar results were obtained at day 3 (**figure 19 B and C**) and day 5 (**figure 20 B and C**). These results indicate that as early as day 3, most of the TFIID complexes are not fully assembled anymore.

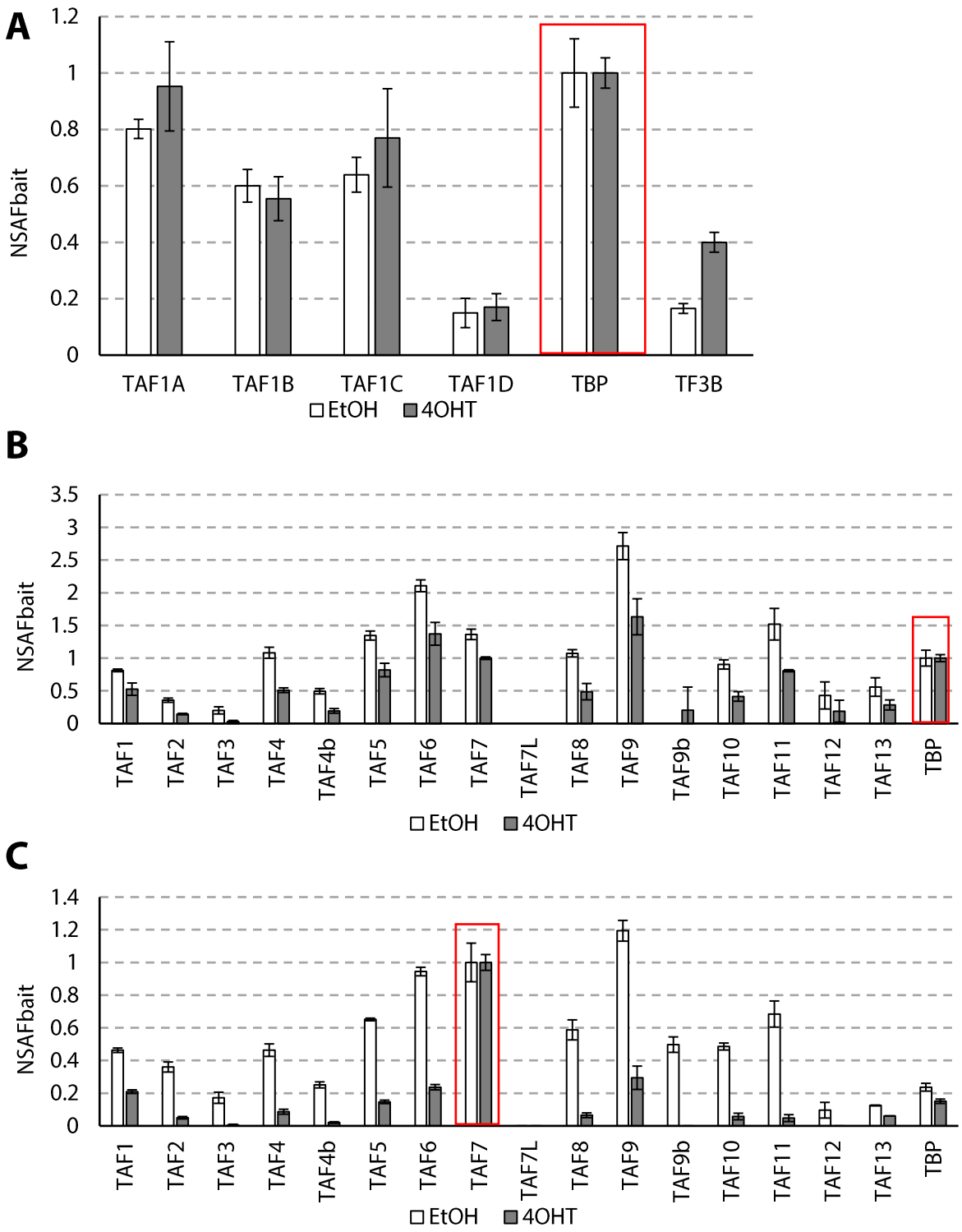



**Figure 19: TFIID assembly defect in mES cells at day 3 after *Taf10* deletion.** (A) $NSAF_{bait}$ values for SL1 complex subunits (TAF1A, TAF1B, TAF1C, TAF1D and TBP) and TF3B-TBP complex at day3. (B) (C) $NSAF_{bait}$ values for TFIID subunits of TBP IP (B) and TAF7 IP (C) from whole cell extracts collected at day3. Control and mutant IPs are indicated in white and gray, respectively. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3

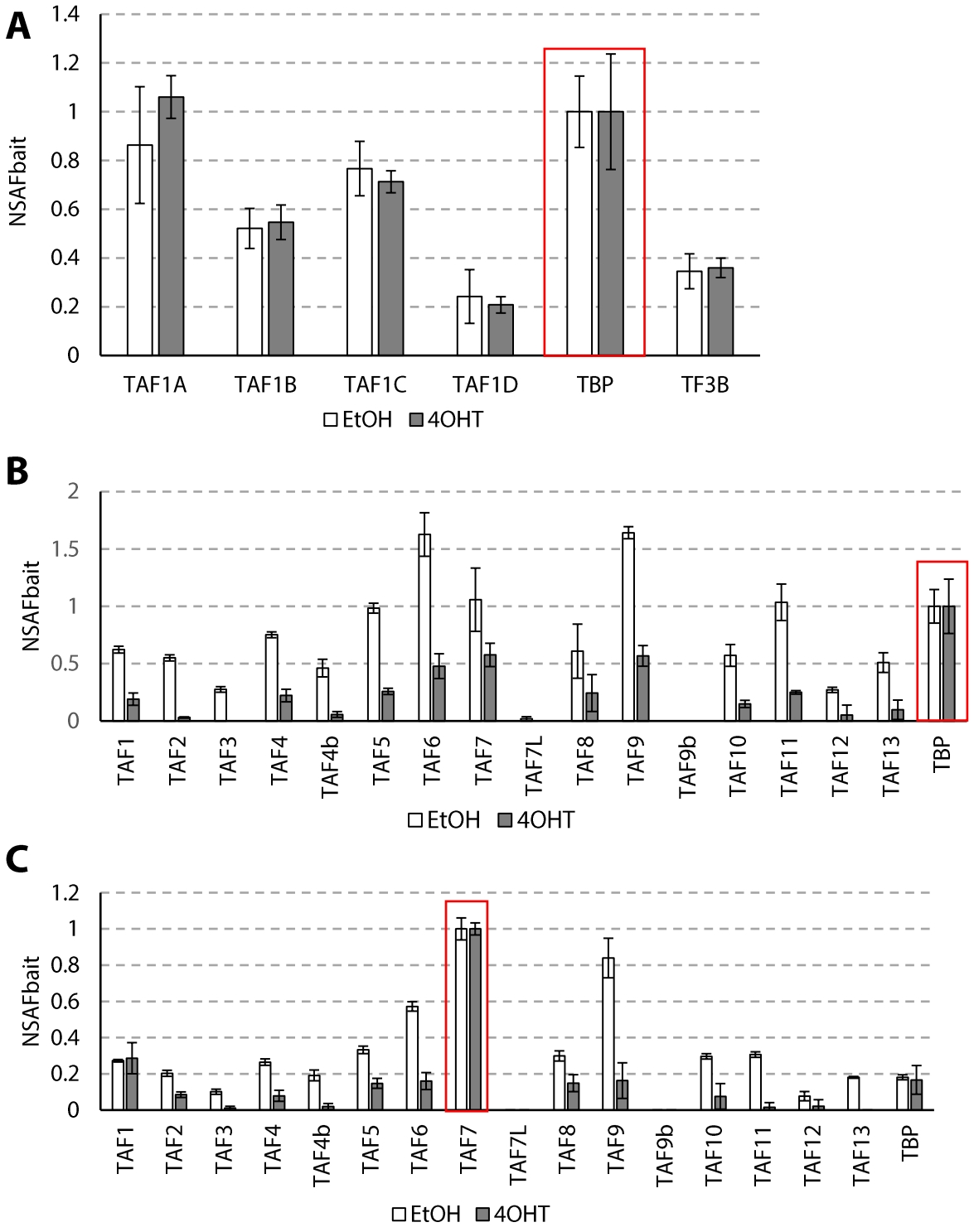



**Figure 20: TFIID assembly defect in mES cells at day 5 after *Taf10* deletion.** (A) $NSAF_{bait}$ values for SL1 complex subunits (TAF1A, TAF1B, TAF1C, TAF1D and TBP) and TF3B-TBP complex at day3. (B) (C) $NSAF_{bait}$ values for TFIID subunits of TBP IP (B) and TAF7 IP (C) from whole cell extracts collected at day3. Control and mutant IPs are indicated in white and gray, respectively. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3

In order to check more carefully the architecture of the TFIID complex, gel filtration analysis from nuclear extracts at day 3 was performed (**figure 21**). This approach was not possible to use previously with the embryo because of the low quantity of protein available, thus mES cells represent an interesting model to analyze deeper the composition and architecture of the TFIID complex. Gel filtration showed that in the TFIID fractions (red rectangle #1), as described in (Langer et al. 2016), only TBP and residual TAF5 remained, while TAF8 together with TAF10 were not detected anymore (**figure 21**).

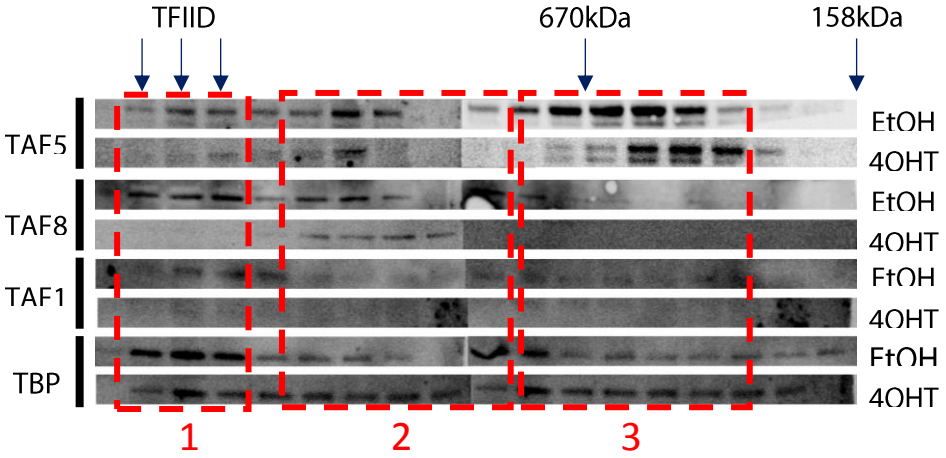



**Figure 21: TFIID assembly defect in mES cells at day 3 after *Taf10* deletion.** Western-blot of gel filtration fraction from nuclear extracts of mES cells at day 3 treated with EtOH or 4OHT. Arrows indicate the corresponding molecular weight, and the expected molecular weight for TFIID. Numbered red rectangles with dashed border indicate potential sub-complexes.

These results confirm that TAF10 is required for the assembly of TFIID, but that holo-TFIID and partially assembled TFIID complexes remain in the nucleus after TAF10 depletion. Interestingly, at a lower molecular weight (approximately 1 MDa) (red rectangle #2), I detected a complex composed of at least TAF5, TAF8 and TBP but with almost undetectable TAF10 in the control (**figure 21**). In the 4OHT-treated cells, this complex was still detected, with an apparent redistribution of TBP, whose signal appeared stronger in those fractions. Similarly, TAF5 and TBP co-eluted at around 670kDa (red rectangle #3) with almost no TAF10 in the control cells, which could indicate the existence of another sub-complex. Together this data indicates that distinct TFIID complexes, or sub-complexes exist in the nucleus of mES cells together with the canonical complex and sub-complexes that could reflect the presence of sub-modules. It was already reported that at least two TFIID complex population might exist from

analyses in HeLa cells (Jacq et al. 1994), which seems to be also the case here. Due to a lack of interpretable data with other antibodies, it is not yet possible to determine more precisely the composition of the remaining TFIID complexes and sub-complexes, and a mass spectrometry analysis has not yet been performed from these fractions.

c. SAGA enzymatic activities are maintained in the absence of TAF10

In order to check whether SAGA enzymatic activities were affected in the absence of TAF10, ubiquitinylation of histone H2B and acetylation of the lysine of the histone H3 levels have been analyzed by western-blot from mES cell acidic extracts at day 3 and day 5 (**figure 22**). No difference could be observed for those two histone marks. Interestingly, transcription inhibition by flavopiridol treatment led to a strong decrease of H2Bub1, as reported in the literature (Davie et al. 1994), but not for 4OHT-treated cells, suggesting a less dramatic effect on transcription with the absence of TAF10. This data suggests that SAGA is still functional.

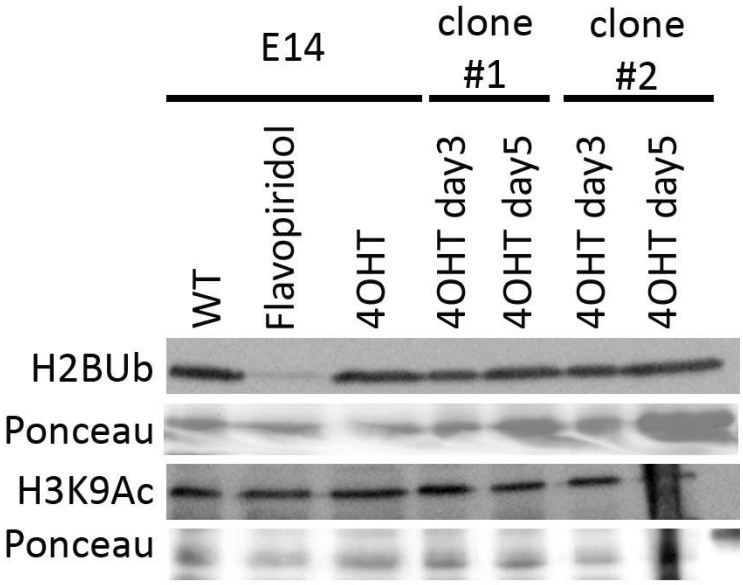



**Figure 22: SAGA enzymatic activities are not affected after *Taf10* deletion.** (A) Western-blot analyses for H2B ubiquitinylation and H3K9 acetylation from acidic extract of mES cells, E14 wild-type, and *R26Cre*$^{ERT2/+}$;*Taf10*$^{fl/fl}$ (clone #1 and #2) mES cells, treated at day 1 with nothing or EtOH or 100 nM 4OHT, collected at day 3 and day 5. E14 mES cells have been treated with flavopiridol for 2 hours before acidic extraction.

d. TAF10 is required for SAGA full assembly

The assembly of SAGA after TAF10 depletion was assessed with TRRAP (figure 23 B) and SUPT3H immuno-precipitations (figure 23 C). Due to technical problems with samples collected at day 3, no data was obtained for this time point concerning. At day 5, TRRAP interaction with TIP60/NuA4 subunits was not affected in mutant mES cells, as reported in (Bardot et al. 2017) (**figure 23 A**). This indicates that TAF10 depletion does not affect the assembly of complexes sharing subunits with SAGA and thus that the effect observed is specific of the depletion of TAF10. Both TRRAP and SUPT3H immunoprecipitations showed that reduced amounts of SAGA subunits were detected in mutant samples, confirming that TAF10 is also required for SAGA assembly. However, from the TRRAP immunoprecipitation (**figure 23 B**), it seems that SAGA disassembly was less severe in mES cells than what was observed in the embryo (Bardot et al. 2017). The HAT enzymes GCN5/KAT2A and PCAF/KAT2B that were not detected (or at low level for GCN5/KAT2A) in the embryo with the TRRAP immuno-precipitation are now found in higher amounts for the same immuno-precipitation, suggesting that PCAF/KAT2B is found in different ratios between cell-types (**figure 23 C**). The protein level of the subunits belonging to the core of the complex were not so dramatically decreased. TAF9 is the only core subunit with a two-fold decrease. Both HAT and DUB modules did not seem dramatically affected. Interestingly, data from the SUPT3H immunoprecipitation (**figure 23 C**) shows that even the amounts of SAGA subunits are even more reduced in the mutant samples than with the TRRAP immunoprecipitation (**figure 23 B**).

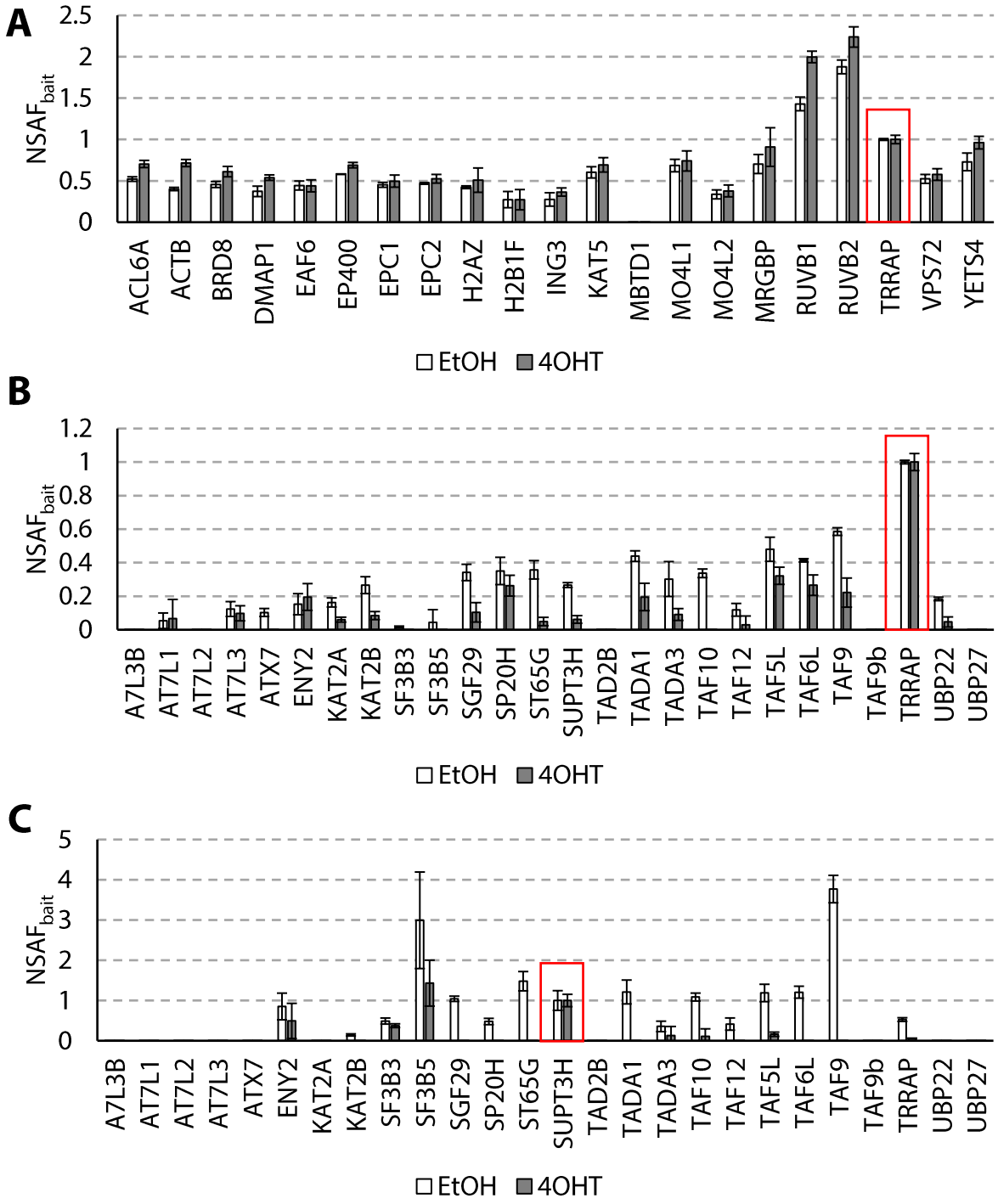



**Figure 23: SAGA assembly defect in mES cells at day 5 after *Taf10* deletion.** (A) $NSAF_{bait}$ values for TIP60/NuA4 complex subunits of TRRAP IP from control or mutant whole cell extracts at day 3. (B) $NSAF_{bait}$ values for SAGA subunits of TRRAP IP from control or mutant whole cell extracts. (C) $NSAF_{bait}$ values for SAGA subunits of SUPT3H IP from control or mutant extracts. Control and mutant IPs are indicated in white and gray, respectively. Error bars indicate s.d. *n*=3

# XI. Analysis of the transcriptional function of TFIID and SAGA

## 1. Gene expression analysis of steady-state mRNA levels in mES cells

To determine the impact of *Taf10* deletion on gene expression in mES cells, steady-state mRNA levels of various genes have been analyzed by RT-qPCR in two independent experiments with EtOH- and 4OHT-treated cells at day 3 only (**figure 24 A and B)** and at day 3 and 5 (**figure 24 C and D**). The total mRNA collected here comes from the nascent transcription experiment (which will be described and detailed thereafter) where mES cells have been treated for 10 minutes with 4-thiouridine (4sU) and spiked-in with *Drosophila* Schneider cells (S2). The cycle-threshold (Ct) values obtained by RT-qPCR were converted in relative concentration with the use of the standard curve calculated for each pairs of primers. The sample relative concentrations were first normalized with the concentrations obtained with the *Drosophila* genes *alphaTUB84B* and *actin42A* to correct any technical bias. Then the relative concentrations were corrected with the RNA Pol I gene *Rn45s*. *Rn45S* did not display significant expression variation, while *Rnu6* expression was variable from one experiment to another with some amplitude (**figure 24 B and D**). This suggests that RNA Pol I transcription is probably not affected, but RNA Pol III transcription remains to be assessed with other genes since much variation has been observed with *Rnu6* (**figure 24 B and D**). The relative fold change was calculated compared to the control (EtOH-treated cells at day3) and set to 1.

In the PSM of *T-Cre;Taf10*$^{flox/flox}$ mutants, several genes were shown to be significantly misregulated (Bardot et al. 2017). In line with this data, *Cdkn1a* and *Cdkn1c* were up-regulated in both experiments, at day 3 and 5 (**figure 24 A and C**) whereas *Gas5*, *Taf1d* were down-regulated at day 3 (not tested at day 5) (**figure 24 C**) in mutant mES cells. The gene *Ccng1*, up-regulated in the PSM of *T-Cre;Taf10*$^{flox/flox}$ mutants, was up-regulated at day 3 but strongly down-regulated at day 5 (**figure 24 C)**. These results indicate that mES cells recapitulate the phenotype observed in the embryo at the gene expression level, at least for those genes. It also indicates that the effect of *Taf10* deletion on gene expression is dynamic, and might give different trends depending on the time point analyzed.

In the first experiment (**figure 24 A**), at day 3, some genes were down-regulated at the steady-state level after *Taf10* deletion including *Rplp0*, *Cdk4* and *Taf8* (**figure 24 A**). The changes observed are moderate and are less than two-fold. It is interesting to note that contrary

to *Tbp* (slightly up-regulated), *Taf8* encoding the interacting partner of TAF10 was decreased, in line with what was observed at the protein level (**figure 16 B and C**). Other genes such as *Gapdh*, *Nanog*, *Cdkn1b*, *Cdkn2a*, *Cdkn2c* and *Ccne2* were up-regulated (**figure 24 A**). At day 3 and day 5 in the second experiment, a similar set of various genes was tested by RT-qPCR (**figure 24 C**). In the mutant mES cells, *Nanog*, *Cdkn1b* and *Cdkn1c* were up-regulated at day 3 and day 5, consistent with the previous experiment (**figure 24 C**). Only *Rplp0* and *Oct4* were following opposite trends between the two experiments (**figure 24 A and C**). This data indicates that global gene expression at the steady-state mRNA level is not impaired, some genes are even up-regulated.

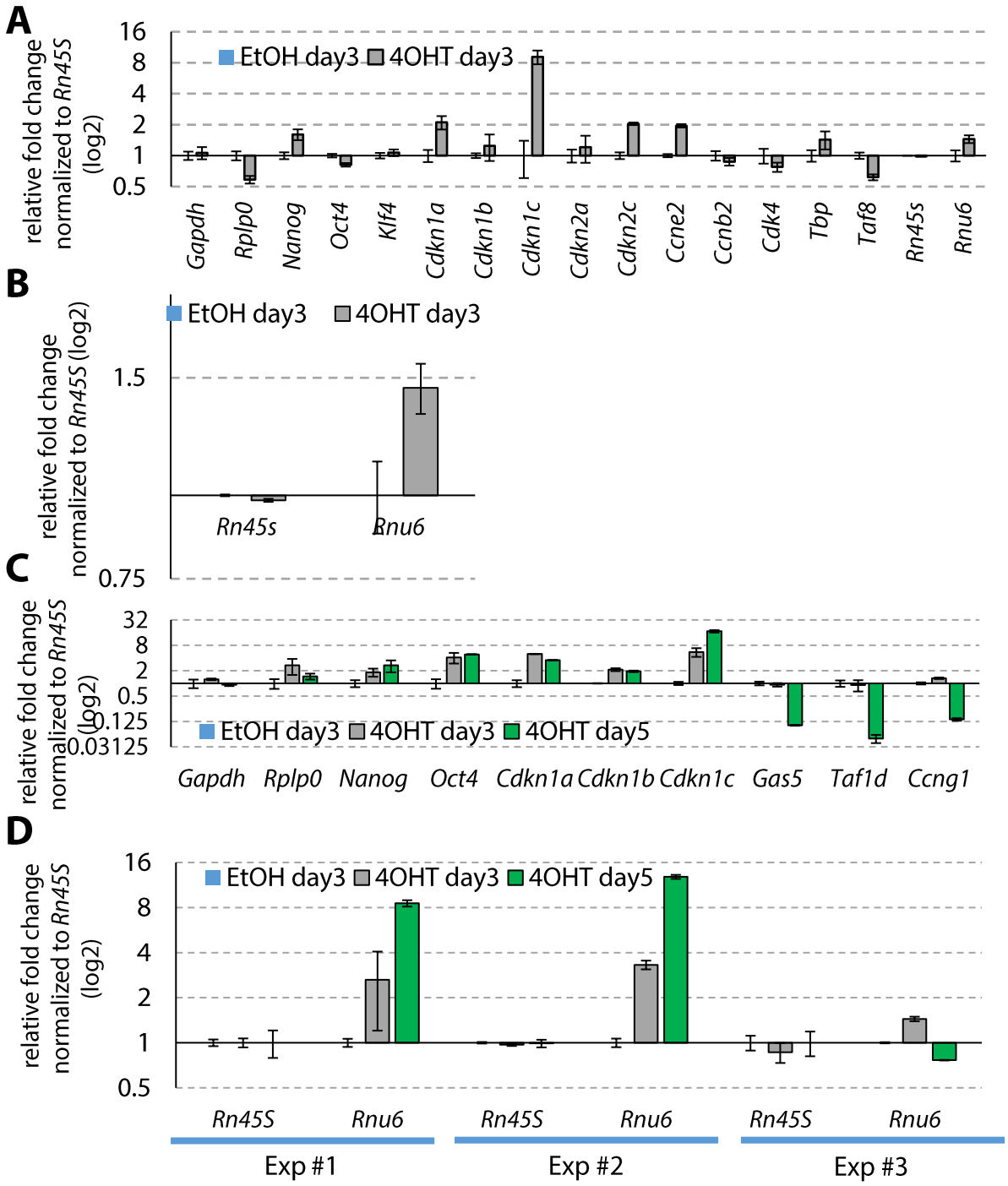



**Figure 24: Steady-state mRNA levels analyses by RT-qPCR**. . EtOH- or 4OHT-treated mES cells at day 3 and/or day 5 were labeled for 10 minutes with 4sU, and spiked-in with S2 cells (ratio mES: S2 cells 3:1). (A) (B) (C) (D) Total mRNA levels were measured by RT-qPCR for samples at (A) (B) day 3 and (C) (D) day 5. Expression values were normalized to spiked-in *Drosophila* S2 cells signal and normalized for *Rn45S* and set to 1 in the control (EtOH-treated cells at day3). *Rn45S* and *Rnu6* were analyzed in three different plates (Exp#1, #2 and #3). Error bars indicate s.d. *n*=3.

## 2. Newly-transcribed mRNA analysis

Previous investigations, including our approach, of the role of TFIID and SAGA *in vivo* were based on steady-state mRNA analyses. Recently, in Làszlò Tora's and Steve Hahn's laboratories, it has been shown that TFIID and SAGA are required for transcription of nearly all genes in *S. cerevisiae* (Baptista et al. 2017; Bonnet et al. 2014; Warfield et al. 2017). These studies shed new light on the role of those complexes in transcription regulation through the analyses of newly-transcribed mRNA. However, little is known about TFIID and SAGA contribution to transcription genome-wide in mammals. In order to better understand how transcription is regulated by those complexes, and to better understand the phenotype we observed in *T-Cre;Taf10*$^{flox/flox}$ mutants, I applied 4sU metabolic labeling of newly-synthesized mRNAs in mES cells (Rädle et al. 2013). This method relies on the labeling of newly-synthesized mRNAs for a short-time of the nucleoside analog 4sU, which is incorporated into the mRNA under synthesis.

### a. Technical validation

The aim of this experiment was to capture the newly-synthesized mRNAs, so I chose a short incubation time for 4sU, here 10 minutes, which was in the range of what was proposed in the protocol published by (Rädle et al. 2013). In order to normalize the data and to avoid any technical bias, *Drosophila* Schneider cells (S2) incubated for 10 minutes with 4sU were added to mES cells just before total RNA extraction with a ratio of 3:1. 4sU-labeled mRNAs were purified and analyzed by qRT-PCR (**figure 14**). Expression levels were normalized and calculated as mentioned above. A set of various protein-coding genes was tested by RT-qPCR (**figure 25**). The results obtained after 10 minutes incubation with 4sU allowed me to detect the expression of all the tested genes as well as the *Drosophila* genes (**figure 25 B and D**). The RNA Pol III gene *Rnu6* displayed a relative variability, and is moderately up-regulated in the mutant samples at both steady-state (**figure 24 B and D**) and newly-synthesized mRNA levels (**figure 25 B and D**). I was also able to detect changes in mRNA synthesis (**figure 25**) in the mutant mES cell samples. The changes of the newly-synthesized mRNA levels follow different, and most of the time opposite, trends than the steady-state levels (**figure 24**). So, this data showed that the technique is working in my conditions and produce enough labeled-mRNAs to be detected by RT-qPCR.

b. Newly-transcribed mRNA is globally affected in the absence of TAF10

At the level of newly-transcribed mRNA levels in the absence of TAF10 (**figure 25**), however, the results indicate a global down-regulation for all these genes, regardless if they were up- or down- regulated at the steady-state mRNA level (**figure 24**). For all of them, at least a 50 % reduction in mRNA synthesis was observed at day 3 and day 5. One exception was noticed for *Cdkn2c* which was up-regulated (**figure 25**). Moreover, for most of them, the decrease was exacerbated at day 5 compared to day 3, suggesting a strong impairment of mRNA synthesis over time. Interestingly, the decrease in mRNA synthesis is associated with an up-regulation of steady-state mRNA levels for many genes, suggesting that there is a cellular compensation at the level of mRNA degradation. Nevertheless, two different behaviors could be observed, since *Gas5*, *Taf1d*, C*cng1* are reduced for both newly-synthesized mRNA and steady-state mRNA levels. Altogether, this data demonstrates that TAF10 is required for RNA Pol II transcription, at least for all the genes tested here, and indicate a potential regulation of mRNA stability depending on the mRNA synthesis rate.

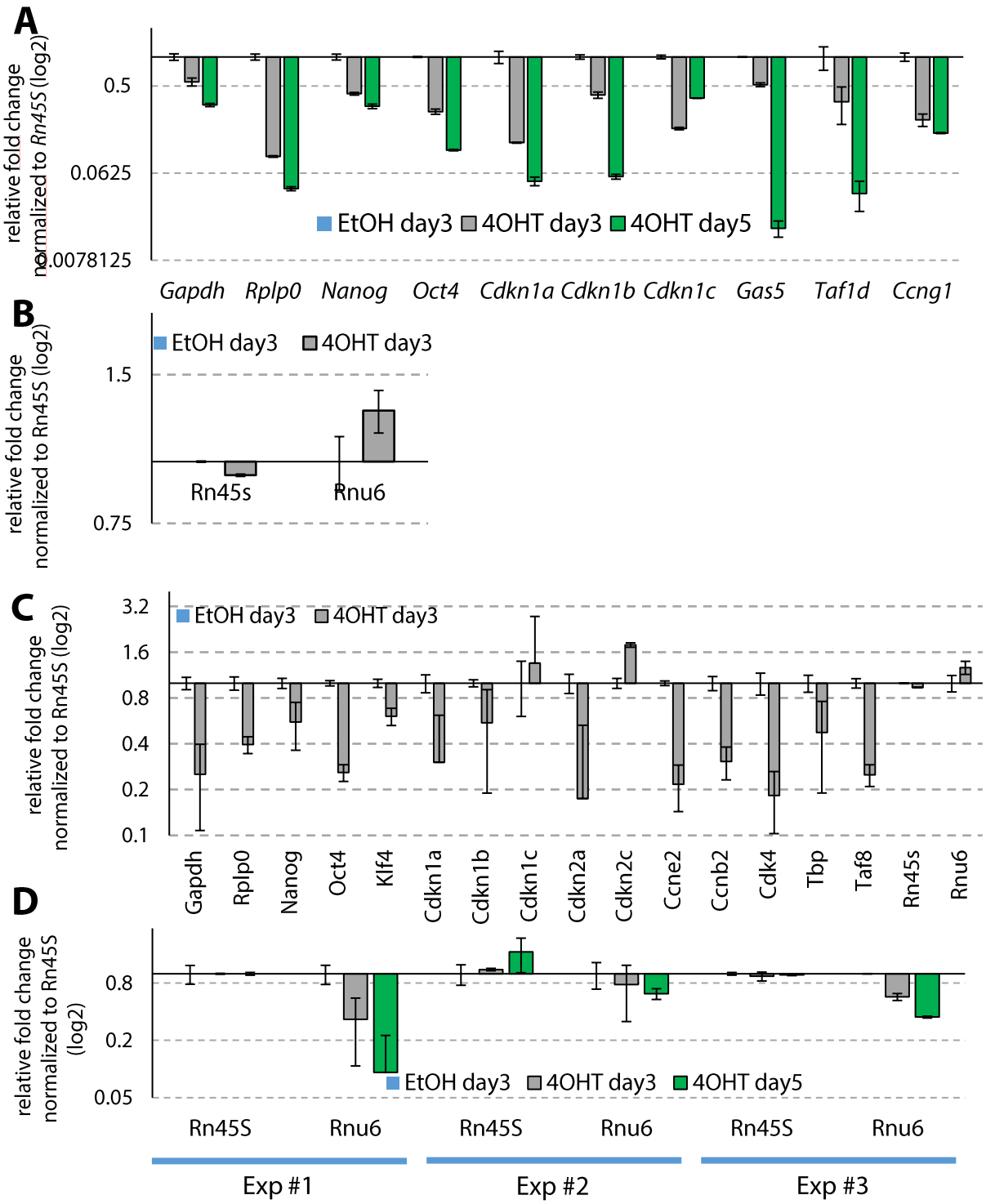



**Figure 25: Newly-synthesized mRNA levels analyses by RT-qPCR**. EtOH or 4OHT-treated mES cells at day3 and day5 were labeled for 10 minutes with 4sU and spiked-in with S2 cells (ratio mES: S2 cells 3:1). (A) (B) Total mRNA levels or (C) (D) newly-synthesized mRNA levels were measured by RT-qPCR. Expression values were normalized to spiked-in *Drosophila* S2 cells signal and normalized for *Rn45S* and set to 1 in the control (EtOH-treated cells at day3). *Rn45S* and *Rnu6* were analyzed in three different runs (Exp #1, #2 and #3). Error bars indicate s.d. *n*=3

c. Genome-wide analysis of nascent transcription

The aim of this approach is to determine the importance of TAF10 for transcription to better understand how TFIID and SAGA control transcription in mammals. So, after having validated the technical approach in mES cells, I used samples collected at day 3 for genome-wide sequencing. At day 3, changes were already detectable and transcription is probable less affected by potential secondary effects resulting from *Taf10* deletion at this time point. To do that, I first tried to use the TT-seq approach described by Patrick Cramer's laboratory (Schwalb et al. 2016). This technique is also based on the metabolic labeling of newly-transcribed mRNAs with 4sU but includes a preliminary step of RNA fragmentation. This step ensures to isolate only the nascent mRNA region and not the 5' preexisting region of the transcript that was already transcribed prior 4sU incorporation (**figure 26**).

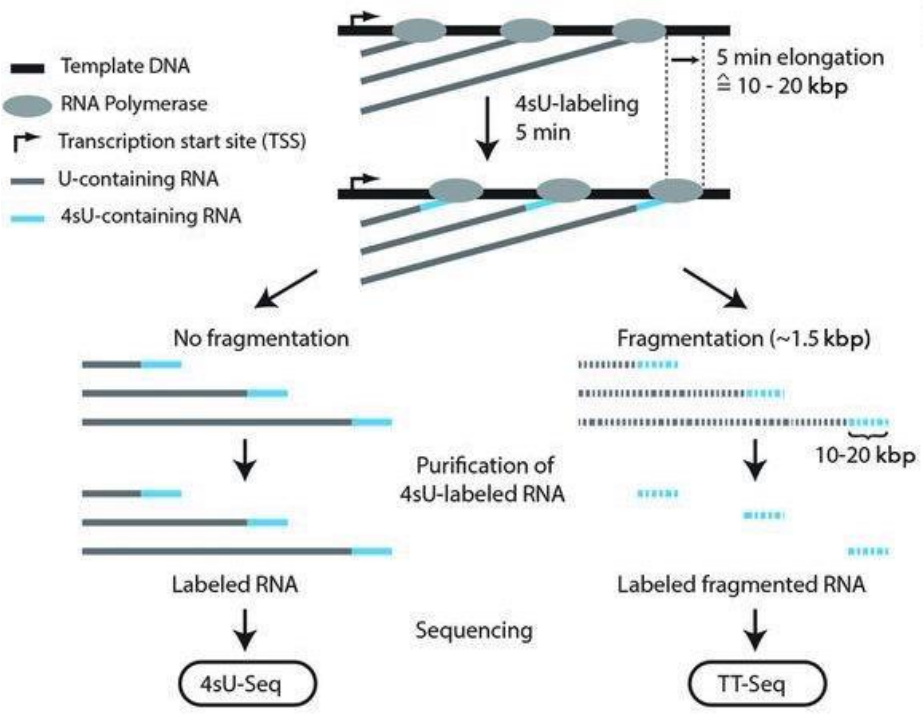



**Figure 26: Comparison of 4sU-seq and TT-seq methods**. 4sU-labeled mRNAs in the 4sU-seq technique give transcripts dominated with a long pre-existing 5' unlabeled region contrary to the TT-seq which allows the isolation of the labeled region of the transcript only (Schwalb et al. 2016).

RNA fragmentation was first optimized with mES cells and with the Covaris E220 for the time needed to obtain a profile similar to what was obtained in the Patrick Cramer's laboratory (personal communication). Total RNA was fragmented for different duration times: 60, 80, 100, 120, 180 or 240 seconds (sec) (**figure 27 and 28**). Similar profiles were obtained, but with a fragmentation time of 80 sec, the profile was the closest to what I was supposed to obtain. When those optimized conditions were applied to 4sU-labeled mRNA samples from day 3 (EtOH- and 4OHT-treated cells), the fragmentation profile was not consistent between triplicate (SNVT26-28 for EtOH and SNVT29-31 for 4OHT samples). After purification, those samples gave a very low yield (between 53 pg to 55 ng), not sufficient for 4sU-labeled mRNA purification. After several unsuccessful trials, it was decided to use the 4sU-seq technique (Rädle et al. 2013) which would still provide information about the status of mRNA synthesis in the absence of TAF10. Unfortunately, I encountered a biotin precipitation issue which delayed the analysis for several months until I was finally able to fix it. So, 4sU-labeled samples in triplicate for EtOH and 4OHT treatment and in duplicate for total mRNA have been sequenced. Results are currently under analysis and cannot yet be presented in this thesis manuscript.

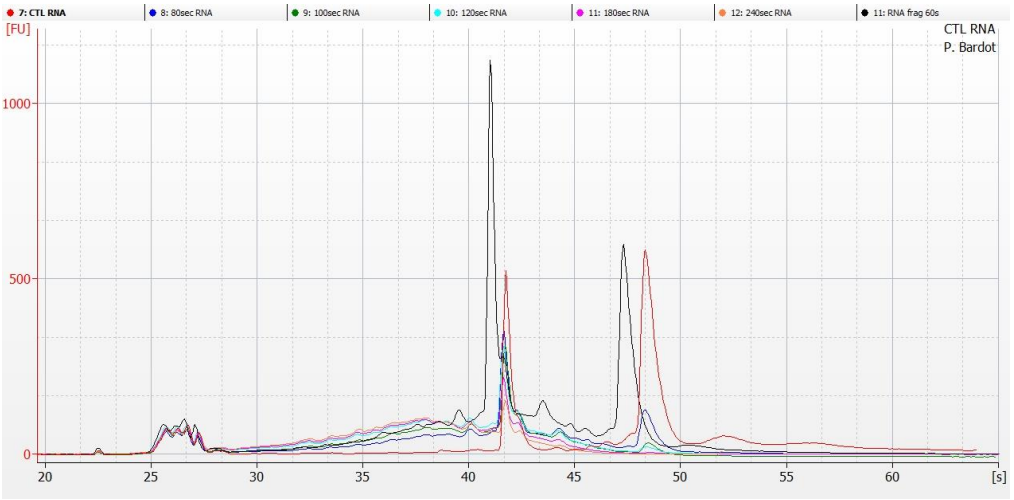



**Figure 27: RNA fragmentation optimization profiles**. 250 ug of total RNA from wild-type mES cells were fragmented using the Covaris E220 and analyzed on the Bioanalyzer. The profile shows the signal intensity (FU) according the elution time (s).

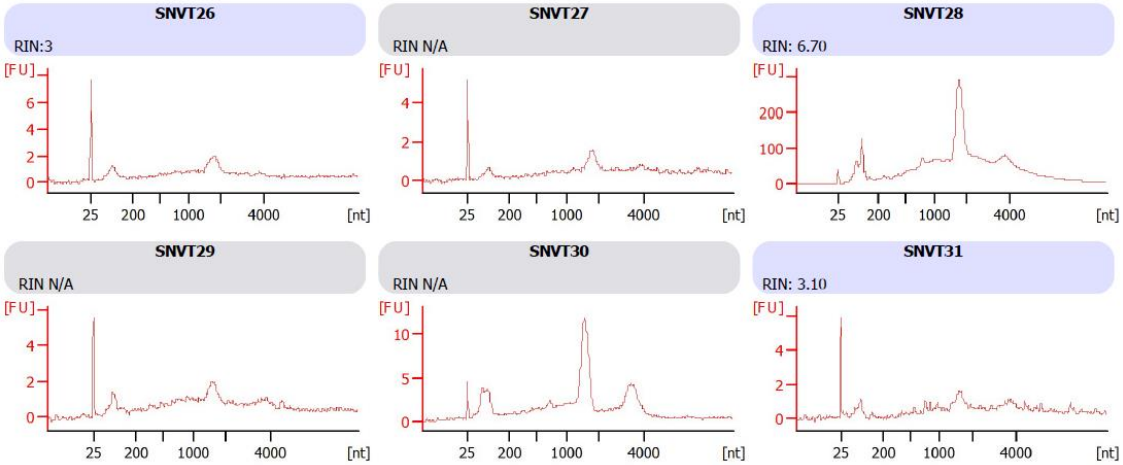



**Figure 28: 4sU-labeled mRNA fragmentation profiles.** 4sU-labeled mRNA from EtOH- (SNVT26-28) and 4OHT- (SNVT 29-31) treated mES cells collected at day 3 were fragmented using the Covaris E220 and analyzed on the Bioanalyzer. The profile shows the signal intensity (FU) according the RNA size (nt).

# General discussion & Perspectives

# XII. Composition of TFIID and SAGA during development

## 1. No alternative TFIID complexes are detected

The characterization of the TFIID complex in different cellular contexts did not reveal an alternative TFIID composition (**figure 15, 19, 20 and 23**). I detected TBP together with TAF1, TAF2, TAF3, TAF4, TAF4b, TAF5, TAF6, TAF7, TAF8, TAF9, TAF9b TAF10, TAF11, TAF12 and TAF13, and is in line with what was described in mES cells (Pijnappel et al. 2013) as well as in mouse and human erythroid cells (Papadopoulos et al. 2015). TAF7L which has been described in male germ cells, in adipocytes and in white fat tissue (Cheng et al. 2007b; Pointud et al. 2003; Zhou et al. 2013a, 2013b) was not detected here. Moreover, no TBP-like proteins were neither detected in these experiments. Together, this data does not reveal alternative TFIID complexes. The ratios between TAF4/B and TAF9/B were shown to change between the three models analyzed here and were also reported in mouse and human erythroid cells (Papadopoulos et al. 2015). However, I compared here: lymphocytes, which are hematopoietic precursors and mES cells, which are pluripotent cells. Since most of the changes in TFIID composition were mainly observed during differentiation processes, it would be interesting to apply this approach to differentiated cells. For instance, it would be very interesting to compare the composition of TFIID in PSM and LPM tissues, which display a differential sensitivity to the conditional deletion of *Taf10* in the embryo. Since it is very challenging to obtain good proteomics data from so low quantity of material, one alternative is to differentiate mES cells into PSM-like and LPM-like cells following the protocol published by (Chal et al. 2015). The preliminary experiments underline the necessity to optimize the protocol in order to obtain the most homogenous cell population as possible.

## 2. Characterization of TFIID sub-complexes in mES cells

The gel filtration experiment from mES nuclear extract (**figure 22**) showed that in wild-type conditions TFIID exist under several forms, and the gel filtration profile is similar to the gel filtration profile published by Irwin Davidson's laboratory based nuclear extracts of mES cells (Langer et al. 2016). The canonical holo-TFIID is found at the highest molecular weight as described previously (Langer et al. 2016). But, TAF5, TAF8, TBP and almost undetectable TAF10 co-eluted at a lower molecular weight (>1 MDa). Furthermore, another complex is

detected around 670 kDa that does not contain TAF10. From what is known about the architecture and the model of assembly of the TFIID complex (Bieniossek et al. 2013; Brou et al. 1993; Jacq et al. 1994; Trowitzsch et al. 2015b) and based on the molecular weight of those sub-complexes, the three TFIID complexes detected here might correspond to the: (1) holo-TFIID complex (TBP and the 13 TAFs), (2) the 8TAF complex (expected to contain of TAF2, 4, 5, 6, 8, 9, 10, 12) and (3) the 5TAF complex (expected to contain TAF4, 5, 6, 9, 12). The data presented here is limited by the number of TFIID subunits tested by western-blot. Therefore, to determine the precise composition of these three forms of TFIID, an analysis by mass-spectrometry is needed. Compared to the model proposed in the literature for the assembly of TFIID (Bieniossek et al. 2013; Trowitzsch et al. 2015b), I found that TBP co-eluted here with the three sub-complexes.

### 3. SAGA composition

To our knowledge, it is the first time that the composition of SAGA is analyzed precisely in the embryo and our data is coherent with the proteomics data published previously in hematopoietic cells (Papadopoulos et al. 2015). All the SAGA subunits were detected. In line with the reported expression pattern of GCN5/KAT2A and PCAF/KAT2B (Xu et al. 2000; Yamauchi et al. 2000), only GCN5/KAT2A was detected at E9.5 in the mouse embryo. It is surprising to detect PCAF/KAT2B in mES cells since they are derived from the blastocyst at a stage preceding the onset expression reported in the mouse embryo (starting at E12.5) (Yamauchi et al. 2000). However, expression of PCAF/KAT2B in mES cells was already reported (Hirsch et al. 2015).

## XIII. The architectural role of TAF10 in TFIID and SAGA assembly

### 1. TAF10 is required for TAF8 stability

Interestingly, in the absence of TAF10 (**figure 16 A**) only the TAF8 protein level is decreased. It suggests that TAF10 regulates the stability of its partner. However, the inverse relationship is not true, since the TAF10 protein level remained unchanged in the absence of TAF8 (El-Saafin et al. 2018). It would be also interesting to analyze the protein levels of both TAF3 and SUPT7L, the two other protein interacting partners reported for TAF10 (Soutoglou

et al. 2005). At the steady-state mRNA level, *Taf8* is also decreased but not *Tbp* (**figure 24 and 25**). Therefore, it seems that the TFIID subunits can regulate each other stability and thus potentially regulating the amount of assembled TFIID. According this hypothesis, if TAF10 is not present in sufficient quantity, the need for TAF8 to incorporate TAF10 decreases and thus TAF8 is not expressed and translated anymore.

### 2. TAF10 is required for TFIID and SAGA full assembly

In both embryo and mES cells, TAF10 is required for the full assembly of TFIID and SAGA complexes. We detected partially assembled TFIID complexes by immuno-precipitations from whole cell extracts in the mutant samples. The fact that both TBP and TAF7 interactions with the other TAFs are reduced when TAF10 is depleted, confirms the central role of TAF10 in nucleating the full assembly of TFIID (Bieniossek et al. 2013). Similarly, TBP and TAF7 immunoprecipitations revealed TFIID assembly defects in fibroblasts harboring the *TAF8:c.781-1G>A* mutation and without detectable TAF8 protein (El-Saafin et al. 2018). Our results are also similar to what was described in the liver (Tatarakis et al. 2008) where TFIID assembly was strongly impaired. However, in F9 cells, it was reported that only TBP was released from the complex (Mohan et al. 2003), which is not the case here in the embryo, neither in mES cells. Our data shows that TBP remains associated to most of the TFIID subunits, though with reduced amounts in the mutant samples. The differences for TFIID integrity observed between F9 cells (Mohan et al. 2003) and post-natal liver (Tatarakis et al. 2008) might come either from the time point used for the analyses and/or from the cellular context itself. While the TFIID complex assembly was assessed only three days after TAF10 removal in F9 cells, it was between 8 and 15 days after complete *Taf10* deletion in the liver (Mohan et al. 2003; Tatarakis et al. 2008). So, one could imagine that it was too early in F9 cells to detect such dramatic differences for TFIID assembly. Nevertheless, I detected that TFIID assembly was compromised as early as 48h (day 3) after 4OHT treatment in mES cells. Therefore, the cellular context might have significant importance on the TFIID architecture and it would be interesting to re-investigate these findings by mass spectrometry.

The gel filtration analysis at day 3 in mES cells from nuclear extracts indicate that the "holo-TFIID" (the three first fractions corresponding to the highest molecular weight) is still present but without several subunits (**figure 21**). In these fractions, when TAF10 is depleted,

only TBP was detected, with reduced level of TAF5 and no TAF8 among the TFIID subunits tested here. The precise composition of the "holo-TFIID" when TAF10 is depleted remains to be tested. This data indicates that the holo-TFIID is capable to sustain a certain flexibility in the absence of several subunits, and there might be a redistribution of certain subunits since TBP co-elutes at the same TFIID molecular weight. Interestingly, the TFIID sub-complexes attributed to the 8TAF and core-TFIID do not seem strongly affected by the TAF10 depletion. Only TAF5 distribution in the 670 kDa sub-complex is shifted downwards. This suggests that mainly the holo-TFIID relies on TAF10 for an efficient assembly of all the TFIID subunits in the complex.

It was reported that the TFTC complex, containing a mixture of SAGA and TFIID subassemblies, displays similar levels of TRRAP, GCN5, and SAP130 (spliceosome-associated protein 130) in the presence or the absence of TAF10 in F9 cells (Mohan et al. 2003). Here, our data clearly shows that both TRRAP and SUPT3 interactions with the SAGA subunits are strongly reduced in the embryo (Bardot et al. 2017). In mES cells, the situation is a bit different since the levels of the SAGA subunits were less decreased in the TRRAP immuno-precipitation. Those differences might be accounted for the cellular context.

Altogether, we show here that TAF10 is generally required for the full assembly of TFIID and SAGA regardless the model (embryo, mES cells, thymocytes).

### 3. Residual TAF10 protein is detected

Residual TAF10 protein could still be detected in TFIID and very rarely in SAGA (**figure 18**), suggesting that TAF10 is more stable when integrated into TFIID. These results raise an important limitation concerning our model based on the inducible *Taf10* deletion. The time between the deletion and the global depletion of the protein takes at least two days, and seems differentially stable when incorporated into TFIID or SAGA. Therefore, it would be very interesting to use an approach based on the conditional rapid degradation of the protein, such as the auxin-inducible degron system which can degrade certain proteins in 30 minutes (Nishimura et al. 2009). I started to establish a mES cell line expressing the plant-specific F box protein OsTIR1 which degrades the aid-degron fused to the N-terminal part of TAF10 in an inducible manner upon auxin treatment. It could be thus possible to examine more directly the impact of TAF10 depletion on TFIID and SAGA assembly. Moreover, this approach would

be very suitable for analyzing the role of TAF10 in PSM-like or LPM-like differentiated cells in order to check whether the phenotype in this two tissues observed in the embryo can be recapitulated and then analyze the effect on TFIID and SAGA assembly in those two tissues.

# XIV. Role of TFIID and SAGA *in vivo*

## 1. TAF10 is essential and required for cellular viability and cellular growth

We showed that *Taf10* deletion in the whole embryo or in the embryonic mesoderm is not viable, mutant embryos could not reach birth (Bardot et al. 2017). The phenotype we observed is in complete accordance with what was described in the literature and confirms that TAF10 is essential for embryonic development (Indra et al. 2005; Mohan et al. 2003; Tatarakis et al. 2008). In particular, the absence of TAF10 at E9.5, impairs the growth of the embryo, and *Taf10* mutant mES cells also stopped growing (**figure 16 D**). Growth of the embryo is completely blocked, and in *T-Cre;Taf10*$^{flox/flox}$ mutants, the allantois and the placenta degenerate which are also strongly deleterious for the survival of the embryo. *Cyclin-dependent kinase inhibitor 1 A* (*p21*) (Xiong et al. 1993) and *Cdkn1C* (*p57*) (Matsuoka et al. 1995) steady-state mRNA levels are up-regulated in the PSM as well as in mES cells in the absence of TAF10. Those genes encode inhibitors of the cell cycle and can explain this phenotype. Their up-regulation could be a consequence of the removal of the repressive action of TFIID proposed by (Tatarakis et al. 2008). This observation is reminiscent with the impairment of *Cyclin E* expression and of the phosphorylation of retinoblastoma protein in F9 cells which are blocked in G1-phase of the cell cycle (Metzger et al. 1999). Here, proliferation assay with EdU-staining coupled to propidium iodide for DNA content measurement in mES cells at day 3 (**figure 16 E**), only showed a slight decrease of cells in S-phase but no data could be obtained at day 5. It is not surprising not to detect major difference at day 3 since the main effect on cell growth is observed at day 4 and day 5. However, compared to the massive apoptosis observed from the inner cell mass in the *Taf10* null mutant (Mohan et al. 2003), I did not detect massive cell death in mES cells and they could be maintained in culture for at least six days. Nevertheless, it was hardly possible to maintain them after trypsinization, they became very small and lost their ability to form massive colonies (data not shown). Such differences between the inner cell mass and mES cells can be explained by the fact that despite a common origin, the blastocyst, they are not identical (Boroviak et al. 2014). mES cells have been first cultivated in 2i+LIF medium

to be maintained and amplified in a pluripotent state prior the inducible deletion of *Taf10*. The results obtained here have been also supported by data obtained from *Taf8* deletion induction in mES cells (El-Saafin et al. 2018). In this study, TAF8 depleted mES cells can be maintained for eight days in culture before they undergo massive cell death, despite a similar phenotype described for *Taf8* null mutation (Voss et al. 2000). This data clearly indicates that there is a differential sensitivity between the inner cell mass and mES cells.

## 2. TAF10 is differentially required between PSM and LPM tissues

At E9.5 in *T-Cre;Taf10*$^{flox/flox}$ mutants, we observed pycnotic nuclei in LPM, which is also targeted by the conditional deletion of *Taf10*, but not in the PSM. One could argue that *Taf10* deletion follow different dynamics between the two tissues, which might be slower in the LPM. However, immuno-fluorescence for TAF10 on PSM and LPM sections at E8.5, showed that TAF10 is equally undetectable from both tissues, ruling out this hypothesis (Bardot et al. 2017).

This observation reveals a differential sensitivity between the two tissues, both expressing *Taf10*. It is not an isolated case, since the survival of the trophoblasts but not the ICM of the blastocyst was described in *Taf10* null mutant embryos (Mohan et al. 2003). To date, the reason of this differential sensitivity has not yet been elucidated. The selective effect observed with the deletion of *Taf10* in the LPM can be explained by several hypotheses.

It is possible that transcription is affected differentially between the PSM and the LPM. For instance, the segmentation clock and the segmentation markers expression is not affected contrary to the LPM markers (*e.g. Hand2*). This suggests already a much more severe effect on gene expression in the LPM. The lack of TAF10 could be functionally compensated in the PSM but not in the LPM. Therefore, investigating the composition of TAF10-containing complexes could be interesting. As mentioned previously, I started to establish an *in cellulo* model for recapitulating the PSM and LPM tissues from mES cell differentiation using the protocol published from Olivier Pourquié's laboratory (Chal et al. 2015). Preliminary results showed that *Taf10* deletion in PSM-like or LPM-like cells recapitulates the phenotype observed in the embryo and in mES cells with a slow-down of the cellular growth for both tissues and no massive cell death (Bardot P. master thesis). To analyze the composition of TAF10-containing complexes, mES cells must be differentiated in PSM and LPM, and sorted with the Mesogenin-

YFP fluorescent reporter (expressed in the PSM progenitors), in order to obtain relative homogenous cellular populations that would allow to recapitulate these two tissues *in cellulo*.

Furthermore, it is known that some TAF proteins can mediate the activation of certain subset of genes by tissue-specific transcription factors. For instance, TAF10 interacts with the hematopoietic transcription factor GATA1 (Papadopoulos et al. 2015), the TAF4–TAF12 heterodimer interacts with HFN4A during post-natal hepatocyte differentiation (Alpern et al. 2014) and with MYB in acute myeloid leukemia (Xu et al. 2018). Genetic inactivation of those *Tafs* phenocopies the effect obtained with their interacting transcription factor. So, it is possible that the loss of TAF10 in the LPM induces the disruption of an interaction with a key transcription factor. To verify this hypothesis, a TAF10 immunoprecipitation in control PSM-like and LPM-like cells could reveal a list of protein candidates, which could be then conditionally depleted with the AID-degron approach to check whether the phenotype is recapitulated.

The phenotype observed in the LPM could also result from secondary effects since *Fgf8* is expressed in the apical ectoderm ridge, and not in the LPM. Here, we found that *Fgf8* is lost in the apical ectoderm ridge of *T-Cre;Taf10*$^{flox/flox}$ mutants at E9.5 and no limb bud was formed, in line with the fact that *Fgf8* is required for normal limb development (Lewandoski et al. 2000; Moon et al. 2000; Sun et al. 2002). It has been proposed that the PSM is important for the limb bud specification in the chicken (Noro et al. 2011). So, the loss of TAF10 in the PSM could affect the level of paracrine signals that are produced from the PSM to specify the limb field. In this study, it was also shown that FGF8 had a protective effect against apoptosis in the limb bud, which could explain the differential sensitivity between PSM and LPM (Noro et al. 2011). In addition, removal of the ectoderm was shown to induce apoptosis in the LPM, but not the PSM, due to an up-regulation of *Bmp4* beyond the normal levels (Schmidt et al. 1998). Importantly, BMP4 negatively regulates *Fgf8* expression, as reported in the chicken prosencephalon and the mouse cortical primordium (Ohkubo et al. 2002; Shimogori 2004). So, it is possible that the loss of TAF10 leads to an up-regulation of BMP4 resulting in LPM apoptosis. Our evidences, however indicate that there is also cell death in the LPM at non-limbs level indicating that FGF8 may not be the cause.

3. Limb bud formation but not vertebrate segmentation is affected at E9.5 in *T-Cre; Taf10* mutants

At E9.5 in *T-Cre;Taf10*$^{flox/flox}$ mutants, we observed the absence of the anterior limb bud but segmentation was not impaired. Expression of *Fgf8* in the presumptive apical ectodermal ridge is not detected at all, while genes of the segmentation clock (*Lfng*, *Hes7*, *Axin2* and *Snai1*) are normally expressed as well as segmentation markers (*Mesp2*, *Uncx4.1*). Our data indicates that the segmentation clock is working properly, the number of somites was not affected in the mutants compared to the wild-type or heterozygous littermates until the embryo becomes completely blocked. However, the size of the somites was much smaller than in controls and was also observed in *Hes7-Cre;Taf10* mutants (data not shown). The fact that somitogenesis can occur normally in the absence of TAF10 might look surprising given that previous studies have shown that SUPT20 plays a role in somitogenesis (Warrier et al. 2017). They both belong to the SAGA complex whose assembly is affected in our study. Moreover, the HAT activity of GCN5 was shown to be required for the paraxial mesoderm survival (Bu et al. 2007) and *Gcn5* and *Supt20* hypomorphs exhibit defects in axial skeletal development, with the decrease of the expression of *Lfng*, *Ripply2*, *Mesp1* and *Dll3* (Warrier et al. 2017). Although we detected at E9.5 a partial SAGA complex, the acetylation levels of H3K9 in mES cells (**figure 22**) were not affected, indicating that the HAT activity remains functional. Based on this observation, it is possible to speculate that the HAT (and the DUB) activity of SAGA is not affected in the *Taf10* mutant embryos. So, a disrupted but functional SAGA complex can explain why somitogenesis can occur with no apparent apoptosis of the paraxial mesoderm.

Concerning the limb bud formation defect, no *Fgf8* was detected at E9.5 in the apical ectoderm ridge. Several hypotheses can explain this phenotype. The expression of *Fgf8* results from a signaling cascade where *Tbx5* induces the expression of *Fgf10*, which establishes a positive feedback loop with *Fgf8*, [reviewed in (Petit et al. 2017)]. So, the absence of signal for *Fgf8* could be the result of defective expression of *Fgf10* or *Tbx5* upstream of the cascade. Preliminary results obtained in the laboratory (Hisler V., master thesis, 2017) indicate that *Tbx5* expression is reduced at E9.5, while still normal at E9.0, and *Fgf10* expression is not detected in the forelimb. Therefore, initial specification seems to be achieved but not maintained. Thus, the absence of *Fgf8* could reflect a defective transcription that would preclude the expression of the downstream targets and as a result impairing the limb bud formation. We showed that *Hand2*, a LPM marker, was down-regulated in the *T-Cre;Taf10*$^{flox/flox}$ mutants contrary to PSM markers, suggesting a more severe effect at the transcription level.

## XV. Role of TFIID and SAGA in mammalian transcription

Given that both TFIID and SAGA assembly was impaired in the absence of TAF10 in the embryo and in mES cells, it was expected that transcription in those systems would be compromised, as it was already evidenced in mouse trophoblasts (Mohan et al. 2003), foetal keratinocytes (Indra et al. 2005) and foetal hepatocytes (Tatarakis et al. 2008).

### 1. TAF10 is required for transcription of many genes with some notable exceptions

Newly-synthesized mRNA analyses in mES cells from a subset of genes showed that RNA Pol II transcription is strongly decreased, by at least 50% (**figure 25**). The on-going analysis of the 4sU-sequencing data obtained from mES cells at day 3 will provide a genome-wide view of the transcriptional status in these cells. Moreover, a RNA Pol II-ChIP sequencing experiment (currently on-going) will be also interesting in order to check whether RNA Pol II recruitment is affected and how it correlates with the changes in gene expression. The results obtained here for the newly-synthesized mRNA analyzed by RT-qPCR are in agreement with what was demonstrated in yeast, where TFIID and SAGA are required for RNA Pol II transcription of nearly all genes (Baptista et al. 2017; Bonnet et al. 2014; Warfield et al. 2017). Warfield *et al.* also showed that TFIID integrity was shown to be compromised upon deletion of several *TAF* genes (Warfield et al. 2017).

However, we found that somitogenesis is still going on at E9.5 while TAF10 was not detected as early as E8.5, with only partial TFIID and SAGA complexes, suggesting that transcription might still work at this stage until the embryo is completely blocked. We also observed the maintenance of the cyclic expression of the *Lfng* reporter *Luvelu*, which is a good readout for transcription initiation. This reporter is based on the expression of the fluorescent protein Venus under the control of *Lfng* promoter (Aulehla et al. 2008). This reporter construct contains also the *Lfng* 3' UTR which destabilizes the transcript, and PEST sequences render the Venus protein very unstable. So, for each cycle of *Luvelu* expression, the mRNA must be transcribed everytime as well as protein translation. Therefore, the maintenance of the *Luvelu* expression pattern indicates that transcription initiation still occurs, at least for a short time. It would have been interesting to know how long these oscillations could be maintained *in vivo*, however, it was not possible to obtain data from a live-imaging experiment of the *Luvelu*

expression in the embryo. An alternative could be to use our mES cell line to obtain PSM-like cells and to monitor gene oscillations as reported recently (Matsumiya et al. 2018). This method is capable to generate oscillations lasting for more than 15 h.

Moreover, the expression of some genes in F9 cells devoid of TAF10 during primitive endoderm differentiation, though not parietal endoderm, was reported (Metzger et al. 1999). In addition, adult skin regenerated normally event in the absence of TAF10 in keratinocytes (Indra et al. 2005). These results indicate that TAF10 might not be required for *de novo* transcription for at least a subset of genes (*eg*. those required for parietal endoderm differentiation and skin regeneration).

Recently, it was reported in our laboratory that TFIID, even with a defective assembly, was able to maintain transcription in fibroblasts, derived from a patient with a *TAF8* mutation (TAF8:c.781-1G>A), but not in *Taf8*$^{lox/lox}$*;Rosa26*$^{CreERT2}$ mES cells treated with 4OHT (El-Saafin et al. 2018). Together, this data indicates that transcription may have different requirements for TFIID composition depending on the genes, the cellular context and the developmental stage.

## 2. Hypothetical model for gene activation by TFIID and transcription maintenance

Tatarakis *et al*. proposed that the canonical TFIID is required for the initial activation of genes, which occurs mainly during development, but then becomes dispensable for transcription maintenance. Given that TAFs were also found downstream of the promoter and could promote transcription reinitiation (Joo et al. 2017), we can reasonably speculate that after their initial activation, many genes retain several PIC components. This would give an advantage for the cell to reinitiate transcription much faster during regeneration or somitogenesis for the cyclic genes. Interestingly, TAF10 was not found enriched in the TAFs found downstream of the promoter (Joo et al. 2017). This observation might indicate that TAF10 would not be part of transcription reinitiation complexes and would also explain why some genes are not sensitive to the loss of TAF10. However, no TAF10 genome-wide ChIP-seq analyses have been performed so far in order to determine its binding profile. Moreover, in other organisms, TAF10 differential requirement for transcription was also described, such as in *C. elegans* development, where transcription of a subset of genes did not require TAF10

(Walker et al. 2001). So, the holo-TFIID complex might not be a general requirement for all genes, under certain conditions, transcription might be able to function with a different kind of TFIID complex.

### 3. Determining the respective contribution of TFIID and SAGA

In the approach used in my thesis, both TFIID and SAGA have been targeted with the deletion of *Taf10*. As a result, the effect observed at the transcriptional level is the potential consequence of the assembly defect of TFIID and SAGA complexes. Therefore, it is not possible to conclude about the individual role of TFIID and SAGA in the control of transcription. Analyses of the levels of H3K9ac and H2Bub1 in the mutant mES cells did not reveal any alteration, suggesting that the HAT and DUB activities of the complex are still functional. However, it cannot be excluded that H3K9ac and removal of H2Bub levels can be achieved by ATAC and independent DUB modules (Atanassov et al. 2016) respectively. The deletion of specific subunits of TFIID and SAGA will clarify the situation. I already crossed the *T-Cre* and $R26^{CreERT2}$ mouse lines with the $Taf7^{flox/flox}$ mouse line in order to analyze the effect of the loss of TAF7 in the embryo and in blastocyst derived ES cells. Moreover, in the laboratory, inducible protein degradation of SAGA and ATAC specific subunits are also ongoing.

### 4. Steady-state gene expression can be sustained with altered TFIID and SAGA complexes

Nascent transcription analyses indicated that TAF10 is required for transcription, while steady-state mRNA analyses of gene expression in the absence of TAF10 did not reveal major changes in the embryonic PSM (Bardot et al. 2017). Instead, only a few genes were misregulated, with some genes that were even up-regulated (*e.g. Cdkn1a* and *Cdkn1c*). Similar observations were made in post-natal keratinocytes and hepatocytes (Indra et al. 2005; Tatarakis et al. 2008), but also in F9 differentiated cells into primitive endoderm where the expression of many genes was unaffected at the mRNA level, with only several genes misregulated (A. Soldatov, W. S. Mohan II, and L. Tora, unpublished data). So, it might look contradictory to claim that TAF10 can be facultative for gene expression given the mRNA synthesis defect observed here. In fact, steady-state levels of mRNA reflect the equilibrium between mRNA

synthesis and degradation. Thus, the observed effect of TAF10 loss/depletion might be only partial, since mRNA degradation can compensate a mRNA synthesis decrease (Sun et al. 2012). Indeed, in mES cells, despite a dramatic decrease in mRNA synthesis rate in mutant samples (**figure 25**), total mRNA levels of many genes were not only stable but were up-regulated. Interestingly, *Cdkn1a* and *Cdkn1c* were up-regulated in mES cells, recapitulating what was observed earlier in the PSM. However, *Gas5* and *Taf1d* total mRNA levels were down-regulated in mES cells, consistent with what was observed in the PSM. This observation is interesting because it shows that not all mRNAs have the same stability dynamics upon *Taf10* deletion. It suggests that mRNA synthesis defect is sensed by the cell, and in response, mRNA stability and decay rate are adjusted to maintain the cellular homeostasis leading to an accumulation of transcripts.

In order to circumvent the mRNA decay compensation and to check the transcription status in the PSM, I planned to analyze newly-synthesized mRNAs directly in the embryo by using the mouse line described in (Gay et al. 2013). The Tg(Uprt;Uprt) mouse line contains a transgene with a ubiquitous chicken *b-actin/CMV* promoter driving a *loxP-GFP-3xstop-loxP* cassette followed by a hemagglutinin epitope-tagged *Uprt* gene. The *Uprt* encodes the uracile phosphoribosyltransferase protein which catalyzes the conversion of uracil to uridine 5'-monophosphate and which can use 4TU for its incorporation into the nascent mRNAs. Crossing this line with our *T-Cre;Taf10*$^{flox/flox}$ mouse line will induce the expression of *Uprt* in the embryo, by removing the STOP cassette, from E7.5 in the mesoderm derivatives. The goal will be to collect the embryonic tail from which it is expected to extract 1 µg of RNA. So, about 100 embryos will be needed for one purification replicate. Nevertheless, several optimization steps are still required, including the incubation time of 4TU (between 4 h and 12 h) and the minimal amount of total RNA required for a yield of several nanograms of labeled mRNA.

### 5. Potential mechanisms of compensation of mRNA decay in response to a decrease in mRNA synthesis

Several evidences suggesting a coupling between mRNA synthesis and decay have been obtained in Patrick Cramer's laboratory in yeast *S. cerevisiae* by using comparative dynamic transcriptome analysis (Sun et al. 2012, 2013). Point mutation in RNA Pol II led to a decrease of the synthesis rate which was compensated by a decrease of the decay rate, and a mutation in

*Ccr4*, encoding the mRNA degradation enzyme Ccr4, induced a decrease of the decay rate which was compensated by a decrease of the synthesis rate (Sun et al. 2012). Mutations of many components of the mRNA decay machinery, all led to a buffering in synthesis rate, except for *Xrn1*, which was proposed to play a major role in this mechanism (Sun et al. 2013) and supported by another study (Haimovich et al. 2013). In yeast, mutations of TFIID and/or SAGA subunits led to such buffering in mRNA decay, though it was only partial upon Taf4 depletion (Baptista et al. 2017; Warfield et al. 2017). In our conditions, up-regulation of steady-state mRNA levels might reflect an accumulation of stabilized transcripts (from day 1 and 2 until day 3 and 5) and also a different degree of mRNA decay compensation. The compensation selectivity is not yet well understood.

In the literature, it has been proposed that promoter elements regulate mRNA decay in yeast and mammal cells (Braun et al. 2014; Bregman et al. 2011; Dori-Bachash et al. 2012; Trcek et al. 2011). So, one speculative hypothesis would be that some genes based on their essentiality for the cellular viability would display specific promoter elements to protect them from mRNA decay when synthesis rate drops. Codon optimality, which refers to the codons that are associated with a faster translation elongation rate, also correlates with transcript stability [reviewed in (Hanson et al. 2018)]. Codons considered as optimal correlate with increased mRNA half-life whereas other codons (sub-optimal) correlate with lower mRNA half-life (Presnyak et al. 2015). Indeed, optimal codons result from the differential concentration of certain tRNAs, that would favor certain codons. In a recent study, the decapping activator and translational regulator Dhh1p was shown to act as a sensor of ribosome speed across the transcriptome, and that promotes mRNA degradation when associated with slow translating ribosomes (Radhakrishnan et al. 2016). Thus, codon optimality could be an explanation for differential mRNA stability upon mRNA synthesis decrease. We can imagine that a decrease in the synthesis rate would give the possibility to sub-optimal codons to be translated more efficiently, due to an increase availability of sub-optimal associated tRNAs, resulting in an increased transcript half-life.

# CONCLUSIONS

The goals of my thesis were to: (1) determine the role of TAF10-containing complexes during somitogenesis at E9.5; (2) develop alternative approaches in order to use classical biochemical approaches; (3) characterize the composition of TFIID and SAGA *in* vivo and in different cellular contexts by mass spectrometry; (4) characterize the composition of TFIID and SAGA in the embryo and in different cellular contexts by mass spectrometry, (5) analyze the contribution of TFIID and SAGA in RNA Pol II transcription in mammals.

The role of TFIID and SAGA has been investigated in the mouse embryo at E9.5 with the conditional *Taf10* deletion in the mesoderm and with the inducible *Taf10* deletion in the whole embryo. We showed that TAF10 is required for embryonic growth, but is differentially required for vertebrate segmentation and limb bud formation. TAF10 is also differentially required between the PSM and LPM tissues and the reason underlining this effect remains to be investigated. These results argue against a generic role of these transcriptional complexes in metazoans.

I provided the detailed composition of the transcriptional TAF10-containing complexes, TFIID and for the first time for SAGA in the embryo. The results obtained here show that the canonical holo-TFIID is found in the embryo as well as in pluripotent cells and immune cells. TAF10 was demonstrated to play a key role in the assembly of these complexes.

The role of TAF10-containing complexes in gene expression has been investigated in the embryo where they are dispensable for steady-state mRNA expression globally in the presomitic mesoderm at E9.5. However, by using newly-synthesized mRNA analyses I showed that TFIID and SAGA are actually required for mRNA synthesis of many genes in mES cells. The results obtained here highlight a potential strong compensation of mRNA decay that results in the stabilization and the accumulation of transcripts.

Altogether, these results clarify the situation for TFIID and SAGA *in vivo* for metazoans, providing new evidences for their role during mammalian development and transcription. Here, I show that the canonical TFIID is found *in vivo*, but that the holo-TFIID is not always required for gene expression and might depend on other parameters that remain to be determined.

# BIBLIOGRAPHY

Adelman, Karen, and John T Lis. 2012. “Promoter-Proximal Pausing of RNA Polymerase II: Emerging Roles in Metazoans.” *Nature Reviews. Genetics* 13 (10). Nature Publishing Group: 720–31. https://doi.org/10.1038/nrg3293.

Aiyar, Sarah E, Jian Long Sun, Ashley L Blair, Christopher A Moskaluk, Yun Zhe Lu, Qi Nong Ye, Yuki Yamaguchi, et al. 2004. “Attenuation of Estrogen Receptor α-Mediated Transcription through Estrogen-Stimulated Recruitment of a Negative Elongation Factor.” *Genes and Development* 18 (17). Cold Spring Harbor Laboratory Press: 2134–46. https://doi.org/10.1101/gad.1214104.

Alekseev, Sergey, Zita Nagy, Jérémy Sandoz, Amélie Weiss, Jean Marc Egly, Nicolas Le May, and Frederic Coin. 2017. “Transcription without XPB Establishes a Unified Helicase-Independent Mechanism of Promoter Opening in Eukaryotic Gene Expression.” *Molecular Cell* 65 (3). Elsevier: 504–514.e4. https://doi.org/10.1016/j.molcel.2017.01.012.

Alpern, Daniil, Diana Langer, Benoit Ballester, Stephanie Le Gras, Christophe Romier, Gabrielle Mengus, and Irwin Davidson. 2014. “TAF4, a Subunit of Transcription Factor II D, Directs Promoter Occupancy of Nuclear Receptor HNF4A during Post-Natal Hepatocyte Differentiation.” *ELife* 3 (January): 1–23. https://doi.org/10.7554/eLife.03613.

Amir-Zilberstein, L., E. Ainbinder, L. Toube, Y. Yamaguchi, H. Handa, and R. Dikstein. 2007. “Differential Regulation of NF- B by Elongation Factors Is Determined by Core Promoter Type.” *Molecular and Cellular Biology* 27 (14): 5246–59. https://doi.org/10.1128/MCB.00586-07.

Andel, Frank, Andreas G Ladurner, Carla Inouye, Robert Tjian, and Eva Nogales. 1999. “Three-Dimensional Structure of the Human TFIID-IIA-IIB Complex.” *Science* 286 (5447): 2153–56. https://doi.org/10.1126/science.286.5447.2153.

Arents, G., R. W. Burlingame, B. C. Wang, W. E. Love, and E. N. Moudrianakis. 1991. “The Nucleosomal Core Histone Octamer at 3.1 A Resolution: A Tripartite Protein Assembly and a Left-Handed Superhelix.” *Proceedings of the National Academy of Sciences* 88 (22): 10148–52. https://doi.org/10.1073/pnas.88.22.10148.

Asahara, H, B Santoso, E Guzman, K Du, P A Cole, I Davidson, and M Montminy. 2001. “Chromatin-Dependent Cooperativity between Constitutive and Inducible Activation Domains in CREB.” *Molecular and Cellular Biology* 21 (23): 7892–7900. https://doi.org/10.1128/MCB.21.23.7892-7900.2001.

Atanassov, Boyko S., Ryan D. Mohan, Xianjiang Lan, Xianghong Kuang, Yue Lu, Kevin Lin, Elizabeth McIvor, et al. 2016. “ATXN7L3 and ENY2 Coordinate Activity of Multiple H2B Deubiquitinases Important for Cellular Proliferation and Tumor Growth.” *Molecular Cell* 62 (4): 558–71. https://doi.org/10.1016/j.molcel.2016.03.030.

Auble, D. T., and S. Hahn. 1993. “An ATP-Dependent Inhibitor of TBP Binding to DNA.” *Genes and Development* 7 (5). Cold Spring Harbor Laboratory Press: 844–56. https://doi.org/10.1101/gad.7.5.844.

Aulehla, Alexander, and Olivier Pourquié. 2010. “Signaling Gradients during Paraxial

Mesoderm Development." *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/cshperspect.a000869.

Aulehla, Alexander, Winfried Wiegraebe, Valerie Baubet, Matthias B. Wahl, Chuxia Deng, Makoto Taketo, Mark Lewandoski, and Olivier Pourquié. 2008. "A β-Catenin Gradient Links the Clock and Wavefront Systems in Mouse Embryo Segmentation." *Nature Cell Biology* 10 (2). Nature Publishing Group: 186–93. https://doi.org/10.1038/ncb1679.

Balasubramanian, Ramakrishnan, Marilyn G Pray-Grant, William Selleck, Patrick A Grant, and Song Tan. 2002. "Role of the Ada2 and Ada3 Transcriptional Coactivators in Histone Acetylation." *Journal of Biological Chemistry* 277 (10). American Society for Biochemistry and Molecular Biology: 7989–95. https://doi.org/10.1074/jbc.M110849200.

Bannister, Andrew J, and Tony Kouzarides. 2011. "Regulation of Chromatin by Histone Modifications." *Cell Research*. Nature Publishing Group. https://doi.org/10.1038/cr.2011.22.

Baptista, Tiago, Sebastian Grünberg, Nadège Minoungou, Maria J.E. Koster, H.T. T.Marc Timmers, Steve Hahn, Didier Devys, and László Tora. 2017. "SAGA Is a General Cofactor for RNA Polymerase II Transcription." *Molecular Cell* 68 (1). Cell Press: 130–143.e5. https://doi.org/10.1016/j.molcel.2017.08.016.

Barberis, A, C W Müller, S C Harrison, and M Ptashne. 1993. "Delineation of Two Functional Regions of Transcription Factor TFIIB." *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences. https://doi.org/10.1073/pnas.90.12.5628.

Bardot, Paul, Stéphane D Vincent, Marjorie Fournier, Alexis Hubaud, Mathilde Joint, László Tora, and Olivier Pourquié. 2017. "The TAF10-Containing TFIID and SAGA Transcriptional Complexes Are Dispensable for Early Somitogenesis in the Mouse Embryo." *Development (Cambridge, England)* 144 (20). Oxford University Press for The Company of Biologists Limited: 3808–18. https://doi.org/10.1242/dev.146902.

Barlev, Nickolai A, Alexander V Emelyanov, Paola Castagnino, Philip Zegerman, Andrew J Bannister, Manuel A Sepulveda, Flavie Robert, et al. 2003. "A Novel Human Ada2 Homologue Functions with Gcn5 or Brg1 to Coactivate Transcription." *Molecular and Cellular Biology* 23 (19). American Society for Microbiology: 6944–57. https://doi.org/10.1128/MCB.23.19.6944-6957.2003.

Bártfai, Richárd, Carolin Balduf, Traci Hilton, Yvonne Rathmann, Yavor Hadzhiev, László Tora, László Orbán, and Ferenc Müller. 2004. "TBP2, a Vertebrate-Specific Member of the TBP Family, Is Required in Embryonic Development of Zebrafish." *Current Biology* 14 (7): 593–98. https://doi.org/10.1016/j.cub.2004.03.034.

Bartholomew, B, M E Dahmus, and C F Meares. 1986. "RNA Contacts Subunits IIo and IIc in HeLa RNA Polymerase II Transcription Complexes." *Journal of Biological Chemistry* 261 (30): 14226–31. http://www.ncbi.nlm.nih.gov/pubmed/2429953.

Basehoar, Andrew D., Sara J. Zanton, and B. Franklin Pugh. 2004. "Identification and Distinct Regulation of Yeast TATA Box-Containing Genes." *Cell* 116 (5). Elsevier: 699–709. https://doi.org/10.1016/S0092-8674(04)00205-3.

Bénazéraf, Bertrand, Paul Francois, Ruth E. Baker, Nicolas Denans, Charles D. Little, and Olivier Pourquié. 2010. “A Random Cell Motility Gradient Downstream of FGF Controls Elongation of an Amniote Embryo.” *Nature* 466 (7303). Nature Publishing Group: 248–52. https://doi.org/10.1038/nature09151.

Bessho, Y., R Sakata, S Komatsu, K Shiota, S Yamada, and R Kageyama. 2001. “Dynamic Expression and Essential Functions of Hes7 in Somite Segmentation.” *Genes and Development* 15 (20): 2642–47. https://doi.org/10.1101/gad.930601.

Bessho, Yasumasa, Hiromi Hirata, Yoshito Masamizu, and Ryoichiro Kageyama. 2003. “Periodic Repression by the BHLH Factor Hes7 Is an Essential Mechanism for the Somite Segmentation Clock.” *Genes and Development* 17 (12). Cold Spring Harbor Laboratory Press: 1451–56. https://doi.org/10.1101/gad.1092303.

Bhaumik, S R, and M R Green. 2001. “SAGA Is an Essential in Vivo Target of the Yeast Acidic Activator Gal4p.” *Genes and Development* 15 (15). Cold Spring Harbor Laboratory Press: 1935–45. https://doi.org/10.1101/gad.911401.

Bieniossek, Christoph, Gabor Papai, Christiane Schaffitzel, Frederic Garzoni, Maxime Chaillet, Elisabeth Scheer, Petros Papadopoulos, Laszlo Tora, Patrick Schultz, and Imre Berger. 2013. “The Architecture of Human General Transcription Factor TFIID Core Complex.” *Nature* 493 (7434). Nature Publishing Group: 699–702. https://doi.org/10.1038/nature11791.

Birck, Catherine, Olivier Poch, Christophe Romier, Marc Ruff, Gabrielle Mengus, Anne Claire Lavigne, Irwin Davidson, and Dino Moras. 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 (2). Elsevier: 239–49. https://doi.org/10.1016/S0092-8674(00)81423-3.

Blackwood, E M, and J T Kadonaga. 1998. “Going the Distance: A Current View of Enhancer Action.” *Science (New York, N.Y.)* 281 (5373): 60–63. http://www.ncbi.nlm.nih.gov/pubmed/9679020.

Boffelli, Dario, Sachiko Takayama, and David I.K. Martin. 2014. “Now You See It: Genome Methylation Makes a Comeback in Drosophila.” *BioEssays* 36 (12). Wiley-Blackwell: 1138–44. https://doi.org/10.1002/bies.201400097.

Bonev, Boyan, Peter Stanley, and Nancy Papalopulu. 2012. “MicroRNA-9 Modulates Hes1 Ultradian Oscillations by Forming a Double-Negative Feedback Loop.” *Cell Reports* 2 (1). Elsevier: 10–18. https://doi.org/10.1016/j.celrep.2012.05.017.

Bonnet, Jacques, Chen-Yi Wang, Tiago Baptista, Stéphane D Vincent, Wei-Chun Hsiao, Matthieu Stierle, Cheng-Fu Kao, László Tora, and Didier Devys. 2014. “The SAGA Coactivator Complex Acts on the Whole Transcribed Genome and Is Required for RNA Polymerase II Transcription.” *Genes & Development* 28 (18): 1999–2012. https://doi.org/10.1101/gad.250225.114.

Boroviak, Thorsten, Remco Loos, Paul Bertone, Austin Smith, and Jennifer Nichols. 2014. “The Ability of Inner-Cell-Mass Cells to Self-Renew as Embryonic Stem Cells Is Acquired Following Epiblast Specification.” *Nature Cell Biology* 16 (6): 513–25.

https://doi.org/10.1038/ncb2965.

Boube, Muriel, Laurent Joulia, David L Cribbs, and Henri Marc Bourbon. 2002. “Evidence for a Mediator of RNA Polymerase II Transcriptional Regulation Conserved from Yeast to Man.” *Cell*. https://doi.org/10.1016/S0092-8674(02)00830-9.

Bourbon, Henri Marc, Andres Aguilera, Aseem Z. Ansari, Francisco J. Asturias, Arnold J. Berk, Stefan Bjorklund, T. Keith Blackwell, et al. 2004. “A Unified Nomenclature for Protein Subunits of Mediator Complexes Linking Transcriptional Regulators to RNA Polymerase II.” *Molecular Cell*. https://doi.org/10.1016/j.molcel.2004.05.011.

Boutet, Stéphane C, Stefano Biressi, Kevin Iori, Vanita Natu, and Thomas a Rando. 2010. “Taf1 Regulates Pax3 Protein by Monoubiquitination in Skeletal Muscle Progenitors.” *Molecular Cell* 40 (5): 749–61. https://doi.org/10.1016/j.molcel.2010.09.029.

Boyer, Thomas G., Michelle E.D. Martin, Emma Lees, Robert P. Ricciardi, and Arnold J. Berk. 1999. “Mammalian Srb/Mediator Complex Is Targeted by Adenovirus E1A Protein.” *Nature* 399 (6733). Nature Publishing Group: 276–79. https://doi.org/10.1038/20466.

Brand, M., J G Moggs, M Oulad-Abdelghani, F Lejeune, F J Dilworth, J Stevenin, G Almouzni, and L Tora. 2001. “UV-Damaged DNA-Binding Protein in the TFTC Complex Links DNA Damage Recognition to Nucleosome Acetylation.” *The EMBO Journal* 20 (12): 3187–96. https://doi.org/10.1093/emboj/20.12.3187.

Brand, Marjorie, Claire Leurent, Véronique Mallouh, Làszlò Tora, and Patrick Schultz. 1999a. “Three-Dimensional Structures of the TAF(II)-Containing Complexes TFIID and TFTC.” *Science* 286 (5447): 2151–53. https://doi.org/10.1126/science.286.5447.2151.

Brand, Marjorie, Ken Yamamoto, Adrien Staub, and Laszlo Tora. 1999b. “Identification of TATA-Binding Protein-Free TAF(II)-Containing Complex Subunits Suggests a Role in Nucleosome Acetylation and Signal Transduction.” *Journal of Biological Chemistry* 274 (26). American Society for Biochemistry and Molecular Biology: 18285–89. https://doi.org/10.1074/jbc.274.26.18285.

Braun, Katherine A, and Elton T Young. 2014. “Coupling MRNA Synthesis and Decay.” *Molecular and Cellular Biology* 34 (22). American Society for Microbiology (ASM): 4078–87. https://doi.org/10.1128/MCB.00535-14.

Bregman, Almog, Moran Avraham-Kelbert, Oren Barkai, Lea Duek, Adi Guterman, and Mordechai Choder. 2011. “Promoter Elements Regulate Cytoplasmic MRNA Decay.” *Cell* 147 (7). Cell Press: 1473–83. https://doi.org/10.1016/j.cell.2011.12.005.

Brou, C, S Chaudhary, I Davidson, Y Lutz, J Wu, J M Egly, L Tora, and P Chambon. 1993. “Distinct TFIID Complexes Mediate the Effect of Different Transcriptional Activators.” *The EMBO Journal* 12 (2): 489–99. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=413232&tool=pmcentrez&rendertype=abstract.

Brown, C. E., L. Howe, K. Sousa, S. C. Alley, M. J. Carrozza, S. Tan, and J. L. Workman. 2001. “Recruitment of HAT Complexes by Direct Activator Interactions with the ATM-Related Tra1 Subunit.” *Science* 292 (5525). American Association for the Advancement

of Science: 2333–37. https://doi.org/10.1126/science.1060214.

Brownell, J. E., and C. D. Allis. 1995. “An Activity Gel Assay Detects a Single, Catalytically Active Histone Acetyltransferase Subunit in Tetrahymena Macronuclei.” *Proceedings of the National Academy of Sciences* 92 (14). National Academy of Sciences: 6364–68. https://doi.org/10.1073/pnas.92.14.6364.

Brownell, J. E., J. Zhou, T. Ranalli, R. Kobayashi, D. G. Edmondson, S. Y. Roth, and C. D. Allis. 1996. “Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5p Linking Histone Acetylation to Gene Activation.” *Cell* 84 (6): 843–51. https://doi.org/10.1016/S0092-8674(00)81063-6.

Bu, Ping, Yvonne A Evrard, Guillermina Lozano, and Sharon Y R Dent. 2007. “Loss of Gcn5 Acetyltransferase Activity Leads to Neural Tube Closure Defects and Exencephaly in Mouse Embryos.” *Molecular and Cellular Biology* 27 (9). American Society for Microbiology (ASM): 3405–16. https://doi.org/10.1128/MCB.00066-07.

Buratowski, S, and H Zhou. 1993. “Functional Domains of Transcription Factor TFIIB.” *Proceedings of the National Academy of Sciences* 90 (12). National Academy of Sciences: 5633–37. https://doi.org/10.1073/pnas.90.12.5633.

Buratowski, Stephen, Steven Hahn, Leonard Guarente, and Phillip A. Sharp. 1989. “Five Intermediate Complexes in Transcription Initiation by RNA Polymerase II.” *Cell* 56 (4): 549–61. https://doi.org/10.1016/0092-8674(89)90578-3.

Buratowski, Stephen, Steven Hahn, Phillip A. Sharp, and Leonard Guarente. 1988. “Function of a Yeast TATA Element-Binding Protein in a Mammalian Transcription System.” *Nature* 334 (6177): 37–42. https://doi.org/10.1038/334037a0.

Burton, Z F, M Killeen, M Sopta, L G Ortolan, and J Greenblatt. 1988. “RAP30/74: A General Initiation Factor That Binds to RNA Polymerase II.” *Molecular and Cellular Biology* 8 (4): 1602–13. https://doi.org/10.1128/MCB.8.4.1602.

Cadena, D L, and M E Dahmus. 1987. “Messenger RNA Synthesis in Mammalian Cells Is Catalyzed by the Phosphorylated Form of RNA Polymerase II.” *Journal of Biological Chemistry* 262 (26): 12468–74. http://www.ncbi.nlm.nih.gov/pubmed/3624268.

Cambray, Noemí, and Valerie Wilson. 2002. “Axial Progenitors with Extensive Potency Are Localised to the Mouse Chordoneural Hinge.” *Development (Cambridge, England)* 129 (20): 4855–66. https://doi.org/10.1016/s0925-4773(98)00015-x.

Carles, C, I Treich, F Bouet, M Riva, and A Sentenac. 1991. “Two Additional Common Subunits, ABC10 Alpha and ABC10 Beta, Are Shared by Yeast RNA Polymerases.” *The Journal of Biological Chemistry* 266 (35): 24092–96. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=1748681%5Cnhttp://www.jbc.org/content/266/35/24092.long.

Cavallini, B, J Huet, J L Plassat, A Sentenac, J M Egly, and P Chambon. 1988. “A Yeast Activity Can Substitute for the HeLa Cell TATA Box Factor.” *Nature* 334 (6177). Nature Publishing Group: 77–80. https://doi.org/10.1038/334077a0.

Chal, Jerome, Masayuki Oginuma, Ziad Al Tanoury, Benedicte Gobert, Olga Sumara, Aurore

Hick, Fanny Bousson, et al. 2015. “Differentiation of Pluripotent Stem Cells to Muscle Fiber to Model Duchenne Muscular Dystrophy.” *Nature Biotechnology* 33 (9). Nature Publishing Group: 962–69. https://doi.org/10.1038/nbt.3297.

Chen, Kai, Jeff Johnston, Wanqing Shao, Samuel Meier, Cynthia Staber, and Julia Zeitlinger. 2013. “A Global Change in RNA Polymerase II Pausing during the Drosophila Midblastula Transition.” *ELife* 2013 (2). eLife Sciences Publications Limited: e00861. https://doi.org/10.7554/eLife.00861.

Chen, Z, and J L Manley. 2000. “Robust MRNA Transcription in Chicken DT40 Cells Depleted of TAF(II)31 Suggests Both Functional Degeneracy and Evolutionary Divergence.” *Molecular and Cellular Biology* 20 (14). American Society for Microbiology: 5064–76. https://doi.org/10.1128/MCB.20.14.5064-5076.2000.

Cheng, Bo, and David H. Price. 2007a. “Properties of RNA Polymerase II Elongation Complexes Before and After the P-TEFb-Mediated Transition into Productive Elongation.” *Journal of Biological Chemistry* 282 (30): 21901–12. https://doi.org/10.1074/jbc.M702936200.

Cheng, Y., M. G. Buffone, M. Kouadio, M. Goodheart, D. C. Page, G. L. Gerton, I. Davidson, and P. J. Wang. 2007b. “Abnormal Sperm in Mice Lacking the Taf7l Gene.” *Molecular and Cellular Biology* 27 (7). American Society for Microbiology: 2582–89. https://doi.org/10.1128/MCB.01722-06.

Chesnut, J. D., J. H. Stephens, and M. E. Dahmus. 1992. “The Interaction of RNA Polymerase II with the Adenovirus-2 Major Late Promoter Is Precluded by Phosphorylation of the C-Terminal Domain of Subunit IIa.” *Journal of Biological Chemistry* 267 (15): 10500–506. http://www.ncbi.nlm.nih.gov/pubmed/1316903.

Cheung, Alan C M, Sarah Sainsbury, and Patrick Cramer. 2011. “Structural Basis of Initial RNA Polymerase II Transcription.” *The EMBO Journal* 30 (23). EMBO Press: 4755–63. https://doi.org/10.1038/emboj.2011.396.

Chiang, Cheng Ming, and Robert G. Roeder. 1995. “Cloning of an Intrinsic Human TFIID Subunit That Interacts with Multiple Transcriptional Activators.” *Science* 267 (5197): 531–36. https://doi.org/10.1126/science.7824954.

Chicca, J J, D T Auble, and B F Pugh. 1998. “Cloning and Biochemical Characterization of TAF-172, a Human Homolog of Yeast Mot1.” *Molecular and Cellular Biology* 18 (3). American Society for Microbiology (ASM): 1701–10. https://doi.org/10.1128/MCB.18.3.1701.

Christ, Bodo, and Charles P. Ordahl. 1995. “Early Stages of Chick Somite Development.” *Anatomy and Embryology* 191 (5): 381–96. https://doi.org/10.1007/BF00304424.

Cojocaru, Marilena, Célia Jeronimo, Diane Forget, Annie Bouchard, Dominique Bergeron, Pierre Côte, Guy G. Poirier, Jack Greenblatt, and Benoit Coulombe. 2008. “Genomic Location of the Human RNA Polymerase II General Machinery: Evidence for a Role of TFIIF and Rpb7 at Both Early and Late Stages of Transcription.” *Biochemical Journal* 409 (1): 139–47. https://doi.org/10.1042/BJ20070751.

Coleman, Robert A., Andrew K.P. Taggart, Sandeep Burma, John J. Chicca, and B. Franklin Pugh. 1999. “TFIIA Regulates TBP and TFIID Dimers.” *Molecular Cell* 4 (3): 451–57. https://doi.org/10.1016/S1097-2765(00)80453-0.

Compe, Emmanuel, and Jean-Marc Egly. 2016. “Nucleotide Excision Repair and Transcriptional Regulation: TFIIH and Beyond.” *Annual Review of Biochemistry* 85 (1): 265–90. https://doi.org/10.1146/annurev-biochem-060815-014857.

Compe, Emmanuel, and Jean Marc Egly. 2012. “TFIIH: When Transcription Met DNA Repair.” *Nature Reviews Molecular Cell Biology* 13 (6). Nature Publishing Group: 343–54. https://doi.org/10.1038/nrm3350.

Conaway, Ronald C, and Joan Weliky Conaway. 1989. “An RNA Polymerase II Transcription Factor Has an Associated DNA-Dependent ATPase (DATPase) Activity Strongly Stimulated by the TATA Region of Promoters (Messenger RNA Synthesis/Core Promoter/Run-off Transcription).” *Biochemistry* 86 (19). National Academy of Sciences: 7356–60. https://doi.org/10.1073/pnas.86.19.7356.

Cooke, J., and E. C. Zeeman. 1976. “A Clock and Wavefront Model for Control of the Number of Repeated Structures during Animal Morphogenesis.” *Journal of Theoretical Biology* 58 (2): 455–76. https://doi.org/10.1016/S0022-5193(76)80131-2.

Core, Leighton J., Joshua J. Waterfall, and John T. Lis. 2008. “Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters.” *Science* 322 (5909): 1845–48. https://doi.org/10.1126/science.1162228.

Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. “Structural Basis of Transcription: RNA Polymerase II at 2.8 Ångstrom Resolution.” *Science* 292 (5523): 1863–76. https://doi.org/10.1126/science.1059493.

Crowley, Thomas E., Timothy Hoey, Jen-Kuei Liu, Yuh Nung Jan, Lily Y. Jan, and Robert Tjian. 1993. “A New Factor Related to TATA-Binding Protein Has Highly Restricted Expression Patterns in Drosophila.” *Nature* 361 (6412). Nature Publishing Group: 557–61. https://doi.org/10.1038/361557a0.

D’Alessio, Joseph A, Raymond Ng, Holger Willenbring, and Robert Tjian. 2011. “Core Promoter Recognition Complex Changes Accompany Liver Development.” *Proceedings of the National Academy of Sciences* 108 (10). National Academy of Sciences: 3906–11. https://doi.org/10.1073/pnas.1100640108.

Dale, J. K., M. Maroto, M. L. Dequeant, P. Malapert, M. McGrew, and O. Pourquie. 2003. “Periodic Notch Inhibition by Lunatic Fringe Underlies the Chick Segmentation Clock.” *Nature* 421 (6920). Nature Publishing Group: 275–78. https://doi.org/10.1038/nature01244.

Dantonel, Jean Christophe, Sophie Quintin, Lòrànt Lakatos, Michel Labouesse, and Làszlò Tora. 2000. “TBP-like Factor Is Required for Embryonic RNA Polymerase II Transcription in C. Elegans.” *Molecular Cell* 6 (3). Elsevier: 715–22. https://doi.org/10.1016/S1097-2765(00)00069-1.

Darzacq, Xavier, Yaron Shav-Tal, Valeria de Turris, Yehuda Brody, Shailesh M Shenoy,

Robert D Phair, and Robert H Singer. 2007. “In Vivo Dynamics of RNA Polymerase II Transcription.” *Nature Structural & Molecular Biology* 14 (9). NIH Public Access: 796–806. https://doi.org/10.1038/nsmb1280.

Davie, James R., and Leigh C. Murphy. 1994. “Inhibition of Transcription Selectively Reduces the Level of Ubiquitinated Histone H2B in Chromatin.” *Biochemical and Biophysical Research Communications* 203 (1): 344–50. https://doi.org/10.1006/bbrc.1994.2188.

Deato, Maria Divina E, and Robert Tjian. 2007. “Switching of the Core Transcription Machinery during Myogenesis.” *Genes & Development* 21 (17): 2137–49. https://doi.org/10.1101/gad.1583407.

DeJong, J., and R. G. Roeder. 1993. “A Single CDNA, HTFIIA/α, Encodes Both the P35 and P19 Subunits of Human TFIIA.” *Genes and Development* 7 (11): 2220–34. https://doi.org/10.1101/gad.7.11.2220.

Demény, Màté a, Evi Soutoglou, Zita Nagy, Elisabeth Scheer, Agnes Jànoshàzi, Magalie Richardot, Manuela Argentini, Pascal Kessler, and Laszlo Tora. 2007. “Identification of a Small TAF Complex and Its Role in the Assembly of TAF-Containing Complexes.” *PloS One* 2 (3): e316. https://doi.org/10.1371/journal.pone.0000316.

Dequéant, Mary L., and Olivier Pourquié. 2008. “Segmental Patterning of the Vertebrate Embryonic Axis.” *Nature Reviews Genetics* 9 (5): 370–82. https://doi.org/10.1038/nrg2320.

Dequéant, Mary Lee, Earl Glynn, Karin Gaudenz, Matthias Wahl, Jie Chen, Arcady Mushegian, and Olivier Pourquié. 2006. “A Complex Oscillating Network of Signaling Genes Underlies the Mouse Segmentation Clock.” *Science* 314 (5805). American Association for the Advancement of Science: 1595–98. https://doi.org/10.1126/science.1133141.

Dikstein, Rivka, Siegfried Ruppert, and Robert Tjian. 1996. “TAFII250 Is a Bipartite Protein Kinase That Phosphorylates the Base Transcription Factor RAP74.” *Cell* 84 (5): 781–90.

Dori-Bachash, Mally, Ophir Shalem, Yair S. Manor, Yitzhak Pilpel, and Itay Tirosh. 2012. “Widespread Promoter-Mediated Coordination of Transcription and MRNA Degradation.” *Genome Biology* 13 (12): R114. https://doi.org/10.1186/gb-2012-13-12-r114.

Dubrulle, Julien, Michael J McGrew, and Olivier Pourquié. 2001. “FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation.” *Cell* 106 (2): 219–32. https://doi.org/10.1016/S0092-8674(01)00437-8.

Dubrulle, Julien, and Olivier Pourquié. 2004. “Fgf8 MRNA Decay Establishes a Gradient That Couples Axial Elongation to Patterning in the Vertebrate Embryo.” *Nature* 427 (6973). Nature Publishing Group: 419–22. https://doi.org/10.1038/nature02216.

Duda, C T. 1976. “Plant RNA Polymerases.” *Annual Review of Plant Physiology* 27 (1): 119–32. https://doi.org/10.1146/annurev.pp.27.060176.001003.

Dudley, Aimée M, Claire Rougeulle, and Fred Winston. 1999. “The Spt Components of SAGA Facilitate TBP Binding to a Promoter at a Post-Activator-Binding Step in Vivo.” *Genes*

*and Development* 13 (22). Cold Spring Harbor Laboratory Press: 2940–45. https://doi.org/10.1101/gad.13.22.2940.

Dumortier, Alexis, Robin Jeannet, Peggy Kirstetter, Eva Kleinmann, MacLean Sellars, Nuno R dos Santos, Christelle Thibault, et al. 2006. "Notch Activation Is an Early and Critical Event during T-Cell Leukemogenesis in Ikaros-Deficient Mice." *Molecular and Cellular Biology* 26 (1). American Society for Microbiology: 209–20. https://doi.org/10.1128/MCB.26.1.209-220.2006.

Dunphy, E L, T Johnson, S S Auerbach, and E H Wang. 2000. "Requirement for TAF(II)250 Acetyltransferase Activity in Cell Cycle Progression." *Molecular and Cellular Biology* 20 (4): 1134–39. https://doi.org/10.1128/MCB.20.4.1134-1139.2000.

Dvir, A, R C Conaway, and J W Conaway. 1997. "A Role for TFIIH in Controlling the Activity of Early RNA Polymerase II Elongation Complexes." *Proceedings of the National Academy of Sciences of the United States of America* 94 (17): 9006–10. http://www.ncbi.nlm.nih.gov/pubmed/9256425.

Dynlacht, Brian David, Timothy Hoey, and Robert Tjian. 1991. "Isolation of Coactivators Associated with the TATA-Binding Protein That Mediate Transcriptional Activation." *Cell* 66 (3). Elsevier: 563–76. https://doi.org/10.1016/0092-8674(81)90019-2.

Eberharter, A, D E Sterner, D Schieltz, A Hassan, J R Yates, S L Berger, and J L Workman. 1999. "The ADA Complex Is a Distinct Histone Acetyltransferase Complex in Saccharomyces Cerevisiae." *Molecular and Cellular Biology* 19 (10). American Society for Microbiology: 6621–31. https://doi.org/10.1128/MCB.19.10.6621.

Edwards, A. M., C. M. Kane, R. A. Young, and R. D. Kornberg. 1991. "Two Dissociable Subunits of Yeast RNA Polymerase II Stimulate the Initiation of Transcription at a Promoter in Vitro." *Journal of Biological Chemistry* 266 (1): 71–75. http://www.ncbi.nlm.nih.gov/pubmed/1985924.

Eisenmann, David M., Karen M. Arndt, Stephanie L. Ricupero, John W. Rooney, and Fred Winston. 1992. "SPT3 Interacts with TFIID to Allow Normal Transcription in Saccharomyces Cerevisiae." *Genes and Development* 6 (7). Cold Spring Harbor Laboratory Press: 1319–31. https://doi.org/10.1101/gad.6.7.1319.

El-Saafin, Farrah, Cynthia Curry, Tao Ye, Jean-Marie Garnier, Isabelle Kolb-Cheynel, Matthieu Stierle, Natalie L Downer, et al. 2018. "Homozygous TAF8 Mutation in a Patient with Intellectual Disability Results in Undetectable TAF8 Protein, but Preserved RNA Polymerase II Transcription." *Human Molecular Genetics*, April. https://doi.org/10.1093/hmg/ddy126.

Fairley, Jennifer A, Rachel Evans, Nicola A Hawkes, and Stefan G E Roberts. 2002. "Core Promoter-Dependent TFIIB Conformation and a Role for TFIIB Conformation in Transcription Start Site Selection." *Molecular and Cellular Biology* 22 (19). American Society for Microbiology (ASM): 6697–6705. https://doi.org/10.1128/MCB.22.19.6697-6705.2002.

Filion, Guillaume J., Joke G. van Bemmel, Ulrich Braunschweig, Wendy Talhout, Jop Kind, Lucas D. Ward, Wim Brugman, et al. 2010. "Systematic Protein Location Mapping

Reveals Five Principal Chromatin Types in Drosophila Cells." *Cell* 143 (2): 212–24. https://doi.org/10.1016/j.cell.2010.09.009.

Flores, O., E. Maldonado, and D. Reinberg. 1989. "Factors Involved in Specific Transcription by Mammalian RNA Polymerase II. Factors IIE and IIF Independently Interact with RNA Polymerase II." *Journal of Biological Chemistry* 264 (15): 8913–21. http://www.ncbi.nlm.nih.gov/pubmed/2566609.

Flores, O, H Lu, M Killeen, J Greenblatt, Z F Burton, and D Reinberg. 1991. "The Small Subunit of Transcription Factor IIF Recruits RNA Polymerase II into the Preinitiation Complex." *Proceedings of the National Academy of Sciences of the United States of America* 88 (22). National Academy of Sciences: 9999–10003. http://www.ncbi.nlm.nih.gov/pubmed/1946469.

Flores, O, H Lu, and D Reinberg. 1992. "Factors Involved in Specific Transcription by Mammalian RNA Polymerase II. Identification and Characterization of Factor IIH." *J Biol Chem* 267 (4): 2786–93. http://www.ncbi.nlm.nih.gov/pubmed/1733973.

Flores, Osvaldo, Hua Lu, and Danny Reinberg. 1990. "Factors Involved in Specific Transcription by Mammalian RNA Polymerase II." *Proceedings of the National Academy of Sciences of the United States of America* 87 (23): 9158–62. https://doi.org/10.1073/pnas.87.23.9158.

Fondell, J D, H Ge, and R G Roeder. 1996. "Ligand Induction of a Transcriptionally Active Thyroid Hormone Receptor Coactivator Complex." *Proceedings of the National Academy of Sciences of the United States of America* 93 (16): 8329–33. https://doi.org/10.1073/pnas.93.16.8329.

Forget, Diane, Marie-France Langelier, Cynthia Thérien, Vincent Trinh, and Benoit Coulombe. 2004. "Photo-Cross-Linking of a Purified Preinitiation Complex Reveals Central Roles for the RNA Polymerase II Mobile Clamp and TFIIE in Initiation Mechanisms." *Molecular and Cellular Biology* 24 (3): 1122–31. https://doi.org/10.1128/MCB.24.3.1122.

Freiman, R. N., S. R. Albright, S. Zheng, W. C. Sha, R. E. Hammer, and R. Tjian. 2001. "Requirement of Tissue-Selective TBP-Associated Factor TAFII105 in Ovarian Development." *Science* 293 (5537): 2084–87. https://doi.org/10.1126/science.1061935.

Gaertner, Bjoern, Jeff Johnston, Kai Chen, Nina Wallaschek, Ariel Paulson, Alexander S. Garruss, Karin Gaudenz, Bony De Kumar, Robb Krumlauf, and Julia Zeitlinger. 2012. "Poised RNA Polymerase II Changes over Developmental Time and Prepares Genes for Future Expression." *Cell Reports* 2 (6): 1670–83. https://doi.org/10.1016/j.celrep.2012.11.024.

Gamper, A. M., J. Kim, and R. G. Roeder. 2009. "The STAGA Subunit ADA2b Is an Important Regulator of Human GCN5 Catalysis." *Molecular and Cellular Biology* 29 (1): 266–80. https://doi.org/10.1128/MCB.00315-08.

Gangloff, Y G, J C Pointud, S Thuault, L Carré, C Romier, S Muratoglu, M Brand, L Tora, J L Couderc, and I Davidson. 2001. "The TFIID Components Human TAF(II)140 and Drosophila BIP2 (TAF(II)155) Are Novel Metazoan Homologues of Yeast TAF(II)47 Containing a Histone Fold and a PHD Finger." *Molecular and Cellular Biology* 21 (15).

American Society for Microbiology (ASM): 5109–21. https://doi.org/10.1128/MCB.21.15.5109-5121.2001.

Gangloff, Y G, S Werten, C Romier, L Carré, O Poch, D Moras, and I Davidson. 2000. "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." *Molecular and Cellular Biology* 20 (1). American Society for Microbiology: 340–51. https://doi.org/10.1128/MCB.20.1.340-351.2000.

Gannon, F, K O'Hare, F Perrin, J P LePennec, C Benoist, M Cochet, R Breathnach, et al. 1979. "Organisation and Sequences at the 5' End of a Cloned Complete Ovalbumin Gene." *Nature* 278 (5703): 428–34. http://www.ncbi.nlm.nih.gov/pubmed/450048.

Gantier, Michael P., Claire E. McCoy, Irina Rusinova, Damien Saulep, Die Wang, Dakang Xu, Aaron T. Irving, et al. 2011. "Analysis of MicroRNA Turnover in Mammalian Cells Following Dicer1 Ablation." *Nucleic Acids Research* 39 (13): 5692–5703. https://doi.org/10.1093/nar/gkr148.

Garbett, K. A., M. K. Tripathi, B. Cencki, J. H. Layer, and P. A. Weil. 2007. "Yeast TFIID Serves as a Coactivator for Rap1p by Direct Protein-Protein Interaction." *Molecular and Cellular Biology* 27 (1): 297–311. https://doi.org/10.1128/MCB.01558-06.

Gariglio, P, M Bellard, and P Chambon. 1981. "Clustering of RNA Polymerase B Molecules in the 5′ Moiety of the Adult β-Globin Gene of Hen Erythrocytes." *Nucleic Acids Research* 9 (11): 2589–98. https://doi.org/10.1093/nar/9.11.2589.

Gay, Leslie, Michael R. Miller, P. Britten Ventura, Vidusha Devasthali, Zer Vue, Heather L. Thompson, Sally Temple, et al. 2013. "Mouse TU Tagging: A Chemical/Genetic Intersectional Method for Purifying Cell Type-Specific Nascent RNA." *Genes and Development* 27 (1): 98–115. https://doi.org/10.1101/gad.205278.112.

Gazdag, Emese, Aleksandar Rajkovic, Maria Elena Torres-Padilla, and Làszlò Tora. 2007. "Analysis of TATA-Binding Protein 2 (TBP2) and TBP Expression Suggests Different Roles for the Two Proteins in Regulation of Gene Expression during Oogenesis and Early Mouse Development." *Reproduction* 134 (1). Society for Reproduction and Fertility: 51–62. https://doi.org/10.1530/REP-06-0337.

Gazdag, Emese, Angèle Santenard, Céline Ziegler-Birling, Gioia Altobelli, Olivier Poch, Làszlò Tora, and Maria Elena Torres-Padilla. 2009. "TBP2 Is Essential for Germ Cell Development by Regulating Transcription and Chromatin Condensation in the Oocyte." *Genes and Development* 23 (18). Cold Spring Harbor Laboratory Press: 2210–23. https://doi.org/10.1101/gad.535209.

Ge, H, E Martinez, C M Chiang, and R G Roeder. 1996. "Activator-Dependent Transcription by Mammalian RNA Polymerase II: In Vitro Reconstitution with General Transcription Factors and Cofactors." *Methods Enzymol* 274: 57–71. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8902796.

Gegonne, Anne, Xuguang Tai, Jinghui Zhang, Gang Wu, Jianjian Zhu, Aki Yoshimoto, Jeffrey Hanson, et al. 2012. "The General Transcription Factor TAF7 Is Essential for Embryonic

Development but Not Essential for the Survival or Differentiation of Mature T Cells." *Molecular and Cellular Biology* 32 (10). American Society for Microbiology: 1984–97. https://doi.org/10.1128/MCB.06305-11.

Gerard, Matthieu, Laurent Fischer, Vincent Moncollin, Jean Marc Chipoulet, Pierre Chambon, and Jean Marc Egly. 1991. "Purification and Interaction Properties of the Human RNA Polymerase B(II) General Transcription Factor BTF2." *Journal of Biological Chemistry* 266 (31): 20940–45. http://www.ncbi.nlm.nih.gov/pubmed/1939143.

Ghazy, M. A., S. A. Brodie, M. L. Ammerman, L. M. Ziegler, and A. S. Ponticelli. 2004. "Amino Acid Substitutions in Yeast TFIIF Confer Upstream Shifts in Transcription Initiation and Altered Interaction with RNA Polymerase II." *Molecular and Cellular Biology* 24 (24): 10975–85. https://doi.org/10.1128/MCB.24.24.10975-10985.2004.

Gill, G, E Pascal, Z H Tseng, and R Tjian. 1994. "A Glutamine-Rich Hydrophobic Patch in Transcription Factor Sp1 Contacts the DTAFII110 Component of the Drosophila TFIID Complex and Mediates Transcriptional Activation." *Proceedings of the National Academy of Sciences of the United States of America* 91 (1): 192–96. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=42912&tool=pmcentrez&rendertype=abstract.

Giraldez, Antonio J., Ryan M. Cinalli, Margaret E. Glasner, Anton J. Enright, J. Michael Thomson, Scott Baskerville, Scott M. Hammond, David P. Bartel, and Alexander F. Schier. 2005. "MicroRNAs Regulate Brain Morphogenesis in Zebrafish." *Science* 308 (5723). American Association for the Advancement of Science: 833–38. https://doi.org/10.1126/science.1109020.

Goldman-Levi, R, C Miller, J Bogoch, and N B Zak. 1996. "Expanding the Mot1 Subfamily: 89B Helicase Encodes a New Drosophila Melanogaster SNF2-Related Protein Which Binds to Multiple Sites on Polytene Chromosomes." *Nucleic Acids Res* 24 (16). Oxford University Press: 3121–28. https://doi.org/6c0103 [pii].

Goodrich, James A., Timothy Hoey, Catherine J. Thut, Arie Admon, and Robert Tjian. 1993. "Drosophila TAFII40 Interacts with Both a VP16 Activation Domain and the Basal Transcription Factor TFIIB." *Cell* 75 (3). Cell Press: 519–30. https://doi.org/10.1016/0092-8674(93)90386-5.

Grant, Patrick A., Laura Duggan, Jacques Côté, Shannon M. Roberts, James E. Brownell, Reyes Candau, Reiko Ohba, et al. 1997. "Yeast Gcn5 Functions in Two Multisubunit Complexes to Acetylate Nucleosomal Histones: Characterization of an Ada Complex and the Saga (Spt/Ada) Complex." *Genes and Development* 11 (13). Cold Spring Harbor Laboratory Press: 1640–50. https://doi.org/10.1101/gad.11.13.1640.

Grant, Patrick A, David Schieltz, Marilyn G Pray-Grant, David J Steger, Joseph C Reese, John R Yates, and Jerry L Workman. 1998. "A Subset of TAF(II)s Are Integral Components of the SAGA Complex Required for Nucleosome Acetylation and Transcriptional Stimulation." *Cell* 94 (1). Elsevier: 45–53. https://doi.org/10.1016/S0092-8674(00)81220-9.

Green, Michael R. 2000. "TBP-Associated Factors (TAF(II)s): Multiple, Selective Transcriptional Mediators in Common Complexes." *Trends in Biochemical Sciences* 25

(2): 59–63. https://doi.org/10.1016/S0968-0004(99)01527-3.

Greer, Eric Lieberman, Mario Andres Blanco, Lei Gu, Erdem Sendinc, Jianzhao Liu, David Aristizábal-Corrales, Chih Hung Hsu, L. Aravind, Chuan He, and Yang Shi. 2015. "DNA Methylation on N6-Adenine in C. Elegans." *Cell* 161 (4): 868–78. https://doi.org/10.1016/j.cell.2015.04.005.

Grob, Patricia, Michael J. Cruse, Carla Inouye, Marian Peris, Pawel A. Penczek, Robert Tjian, and Eva Nogales. 2006. "Cryo-Electron Microscopy Studies of Human TFIID: Conformational Breathing in the Integration of Gene Regulatory Cues." *Structure* 14 (3): 511–20. https://doi.org/10.1016/j.str.2005.11.020.

Grünberg, Sebastian, Steven Henikoff, Steven Hahn, and Gabriel E Zentner. 2016. "Mediator Binding to UASs Is Broadly Uncoupled from Transcription and Cooperative with TFIID Recruitment to Promoters." *The EMBO Journal* 35 (22). EMBO Press: 2435–46. https://doi.org/10.15252/embj.201695020.

Grünberg, Sebastian, and Gabriel E. Zentner. 2017. "Genome-Wide Characterization of Mediator Recruitment, Function, and Regulation." *Transcription* 8 (3). Taylor & Francis: 169–74. https://doi.org/10.1080/21541264.2017.1291082.

Guelman, S., K. Kozuka, Y. Mao, V. Pham, M. J. Solloway, J. Wang, J. Wu, J. R. Lill, and J. Zha. 2009. "The Double-Histone-Acetyltransferase Complex ATAC Is Essential for Mammalian Development." *Molecular and Cellular Biology* 29 (5): 1176–88. https://doi.org/10.1128/MCB.01599-08.

Guelman, S., T. Suganuma, L. Florens, V. Weake, S. K. Swanson, M. P. Washburn, S. M. Abmayr, and J. L. Workman. 2006a. "The Essential Gene Wda Encodes a WD40 Repeat Subunit of Drosophila SAGA Required for Histone H3 Acetylation." *Molecular and Cellular Biology* 26 (19). American Society for Microbiology: 7178–89. https://doi.org/10.1128/MCB.00130-06.

Guelman, Sebastián, Tamaki Suganuma, Laurence Florens, Selene K Swanson, Cheri L Kiesecker, Thomas Kusch, Scott Anderson, et al. 2006b. "Host Cell Factor and an Uncharacterized SANT Domain Protein Are Stable Components of ATAC, a Novel DAda2A/DGcn5-Containing Histone Acetyltransferase Complex in Drosophila." *Molecular and Cellular Biology* 26 (3). American Society for Microbiology: 871–82. https://doi.org/10.1128/MCB.26.3.871-882.2006.

Guenther, Matthew G., Stuart S. Levine, Laurie A. Boyer, Rudolf Jaenisch, and Richard A. Young. 2007. "A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells." *Cell* 130 (1): 77–88. https://doi.org/10.1016/j.cell.2007.05.042.

Guermah, Mohamed, Kai Ge, Cheng Ming Chiang, and Robert G. Roeder. 2003. "The TBN Protein, Which Is Essential for Early Embryonic Mouse Development, Is an Inducible TAFII Implicated in Adipogenesis." *Molecular Cell* 12 (4): 991–1001. https://doi.org/10.1016/S1097-2765(03)00396-4.

Ha, Ilho, William S. Lane, and Danny Reinberg. 1991. "Cloning of a Human Gene Encoding the General Transcription Initiation Factor IIB." *Nature* 352 (6337): 689–95. https://doi.org/10.1038/352689a0.

Hahn, Steven, Stephen Buratowski, Phillip A. Sharp, and Leonard Guarente. 1989. “Isolation of the Gene Encoding the Yeast TATA Binding Protein TFIID: A Gene Identical to the SPT15 Suppressor of Ty Element Insertions.” *Cell* 58 (6). Cell Press: 1173–81. https://doi.org/10.1016/0092-8674(89)90515-1.

Haimovich, Gal, Daniel A. Medina, Sebastien Z. Causse, Manuel Garber, Gonzalo Millán-Zambrano, Oren Barkai, Sebastián Chávez, José E. Pérez-Ortín, Xavier Darzacq, and Mordechai Choder. 2013. “XGene Expression Is Circular: Factors for MRNA Degradation Also Foster MRNA Synthesis.” *Cell* 153 (5). Cell Press: 1000–1011. https://doi.org/10.1016/j.cell.2013.05.012.

Hall, Daniel B., and Kevin Struhl. 2002. “The VP16 Activation Domain Interacts with Multiple Transcriptional Components as Determined by Protein-Protein Cross-Linking in Vivo.” *Journal of Biological Chemistry* 277 (48). American Society for Biochemistry and Molecular Biology: 46043–50. https://doi.org/10.1074/jbc.M208911200.

Hampsey, M. 1998. “Molecular Genetics of the RNA Polymerase II General Transcriptional Machinery.” *Microbiology and Molecular Biology Reviews : MMBR* 62 (2). American Society for Microbiology: 465–503. https://doi.org/10.1158/0008-5472.CAN-07-2721.

Han, Yan, Jie Luo, Jeffrey Ranish, and Steven Hahn. 2014. “Architecture of the Saccharomyces Cerevisiae SAGA Transcription Coactivator Complex.” *The EMBO Journal* 33 (21): 2534–46. https://doi.org/10.15252/embj.201488638.

Hanisch, A., M. V. Holder, S. Choorapoikayil, M. Gajewski, E. M. Ozbudak, and J. Lewis. 2013. “The Elongation Rate of RNA Polymerase II in Zebrafish and Its Significance in the Somite Segmentation Clock.” *Development* 140 (2): 444–53. https://doi.org/10.1242/dev.077230.

Hanson, Gavin, and Jeff Coller. 2018. “Translation and Protein Quality Control: Codon Optimality, Bias and Usage in Translation and MRNA Decay.” *Nature Reviews Molecular Cell Biology*. Nature Publishing Group. https://doi.org/10.1038/nrm.2017.91.

He, Yuan, Chunli Yan, Jie Fang, Carla Inouye, Robert Tjian, Ivaylo Ivanov, and Eva Nogales. 2016. “Near-Atomic Resolution Visualization of Human Transcription Promoter Opening.” *Nature* 533 (7603). Nature Publishing Group: 359–65. https://doi.org/10.1038/nature17970.

Helmlinger, Dominique, Sara Hardy, Souphatta Sasorith, Fabrice Klein, Flavie Robert, Chantal Weber, Laurent Miguet, et al. 2004. “Ataxin-7 Is a Subunit of GCN5 Histone Acetyltransferase-Containing Complexes.” *Human Molecular Genetics* 13 (12): 1257–65. https://doi.org/10.1093/hmg/ddh139.

Helmlinger, Dominique, Samuel Marguerat, Judit Villén, Danielle L. Swaney, Steven P. Gygi, Jürg Bähler, and Fred Winston. 2011. “Tra1 Has Specific Regulatory Roles, Rather than Global Functions, within the SAGA Co-Activator Complex.” *EMBO Journal* 30 (14). European Molecular Biology Organization: 2843–52. https://doi.org/10.1038/emboj.2011.181.

Helmlinger, Dominique, and László Tora. 2017. “Sharing the SAGA.” *Trends in Biochemical Sciences*. https://doi.org/10.1016/j.tibs.2017.09.001.

Hendrix, D. A., J.-W. Hong, J. Zeitlinger, D. S. Rokhsar, and M. S. Levine. 2008. “Promoter Elements Associated with RNA Pol II Stalling in the Drosophila Embryo.” *Proceedings of the National Academy of Sciences* 105 (22). National Academy of Sciences: 7762–67. https://doi.org/10.1073/pnas.0802406105.

Henry, Karl W., Anastasia Wyce, Wan Sheng Lo, Laura J. Duggan, N. C Tolga Emre, Cheng Fu Kao, Lorraine Pillus, Ali Shilatifard, Mary Ann Osley, and Shelley L. Berger. 2003. “Transcriptional Activation via Sequential Histone H2B Ubiquitylation and Deubiquitylation, Mediated by SAGA-Associated Ubp8.” *Genes and Development* 17 (21). Cold Spring Harbor Laboratory Press: 2648–63. https://doi.org/10.1101/gad.1144003.

Hernández-Hernández, A., and A. Ferrús. 2001. “Prodos Is a Conserved Transcriptional Regulator That Interacts with DTAF(II)16 in Drosophila Melanogaster.” *Molecular and Cellular Biology* 21 (2): 614–23. https://doi.org/10.1128/MCB.21.2.614-623.2001.

Hernandez, N. 1993. “TBP, a Universal Eukaryotic Transcription Factor?. [Review].” *Genes Dev.* 7 (7B): 1291–1308. https://doi.org/10.1101/gad.7.7b.1291.

Herrera, Francisco J., Teppei Yamaguchi, Henk Roelink, and Robert Tjian. 2014. “Core Promoter Factor TAF9B Regulates Neuronal Gene Expression.” *ELife* 3 (3). eLife Sciences Publications Ltd: e02559. https://doi.org/10.7554/eLife.02559.

Hiller, M. 2004. “Testis-Specific TAF Homologs Collaborate to Control a Tissue-Specific Transcription Program.” *Development* 131 (21): 5297–5308. https://doi.org/10.1242/dev.01314.

Hilton, Traci L, Yun Li, Elizabeth L Dunphy, and Edith H Wang. 2005. “TAF1 Histone Acetyltransferase Activity in Sp1 Activation of the Cyclin D1 Promoter.” *Molecular and Cellular Biology* 25 (10). American Society for Microbiology (ASM): 4321–32. https://doi.org/10.1128/MCB.25.10.4321-4332.2005.

Hirata, Hiromi, Shigeki Yoshiura, Toshiyuki Ohtsuka, Yasumasa Bessho, Takahiro Harada, Kenichi Yoshikawa, and Ryoichiro Kageyama. 2002. “Oscillatory Expression of the BHLH Factor Hes1 Regulated by a Negative Feedback Loop.” *Science* 298 (5594). American Association for the Advancement of Science: 840–43. https://doi.org/10.1126/science.1074560.

Hirsch, Calley L, Zeynep Coban Akdemir, Li Wang, Gowtham Jayakumaran, Dan Trcka, Alexander Weiss, J Javier Hernandez, et al. 2015. “Myc and SAGA Rewire an Alternative Splicing Network during Early Somatic Cell Reprogramming.” *Genes & Development* 29 (8). Cold Spring Harbor Laboratory Press: 803–16. https://doi.org/10.1101/gad.255109.114.

Hisatake, K, T Ohta, R Takada, M Guermah, M Horikoshi, Y Nakatani, and R G Roeder. 1995. “Evolutionary Conservation of Human TATA-Binding-Polypeptide-Associated Factors TAFII31 and TAFII80 and Interactions of TAFII80 with Other TAFs and with General Transcription Factors.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (18): 8195–99. https://doi.org/10.1073/pnas.92.18.8195.

Hoey, T, B D Dynlacht, M G Peterson, B F Pugh, and R Tjian. 1990. “Isolation and

Characterization of the Drosophila Gene Encoding the TATA Box Binding Protein, TFIID." *Cell* 61 (7): 1179–86. https://doi.org/10.1016/0092-8674(90)90682-5.

Hoey, Timothy, Robert O.J. Weinzierl, Grace Gill, Jin Long Chen, Brian David Dynlacht, and Robert Tjian. 1993. "Molecular Cloning and Functional Analysis of Drosophila TAF110 Reveal Properties Expected of Coactivators." *Cell* 72 (2): 247–60. https://doi.org/10.1016/0092-8674(93)90664-C.

Hoffman, A, E Sinn, T Yamamoto, J Wang, A Roy, M Horikoshi, and R G Roeder. 1990. "Highly Conserved Core Domain and Unique N Terminus with Presumptive Regulatory Motifs in a Human TATA Factor (TFIID)." *Nature* 346 (6282). Nature Publishing Group: 387–90. https://doi.org/10.1038/346387a0.

Hoffmann, Alexander, and Robert G. Roeder. 1996. "Cloning and Characterization of Human TAF20/15. Multiple Interactions Suggest a Central Role in TFIID Complex Formation." *Journal of Biological Chemistry* 271 (30): 18194–202. https://doi.org/10.1074/jbc.271.30.18194.

Høiby, Torill, Huiqing Zhou, Dimitra J. Mitsiou, and Hendrik G. Stunnenberg. 2007. "A Facelift for the General Transcription Factor TFIIA." *Biochimica et Biophysica Acta - Gene Structure and Expression* 1769 (7–8): 429–36. https://doi.org/10.1016/j.bbaexp.2007.04.008.

Holstege, F. C., P C. van der Vliet, and H. T. Timmers. 1996. "Opening of an RNA Polymerase II Promoter Occurs in Two Distinct Steps and Requires the Basal Transcription Factors IIE and IIH." *The EMBO Journal* 15 (7): 1666–77. http://www.ncbi.nlm.nih.gov/pubmed/8612591.

Holstege, Frank C.P., Ezra G. Jennings, John J. Wyrick, Tong Ihn Lee, Christoph J. Hengartner, Michael R. Green, Todd R. Golub, Eric S. Lander, and Richard A. Young. 1998. "Dissecting the Regulatory Circuitry of a Eukaryotic Genome." *Cell* 95 (5): 717–28. https://doi.org/10.1016/S0092-8674(00)81641-4.

Horikoshi, M, C K Wang, H Fujii, J a Cromlish, P a Weil, and R G Roeder. 1989. "Purification of a Yeast TATA Box-Binding Protein That Exhibits Human Transcription Factor IID Activity." *Proceedings of the National Academy of Sciences of the United States of America* 86 (13). National Academy of Sciences: 4843–47. https://doi.org/10.1073/pnas.86.13.4843.

Horiuchi, J, N Silverman, B Piña, G a Marcus, and L Guarente. 1997. "ADA1, a Novel Component of the ADA/GCN5 Complex, Has Broader Effects than GCN5, ADA2, or ADA3." *Molecular and Cellular Biology* 17 (6): 3220–28. https://doi.org/10.1128/MCB.17.6.3220.

Hou, S Y, S Y Wu, T Zhou, M C Thomas, and C M Chiang. 2000. "Alleviation of Human Papillomavirus E2-Mediated Transcriptional Repression via Formation of a TATA Binding Protein (or TFIID)-TFIIB-RNA Polymerase II-TFIIF Preinitiation Complex." *Molecular and Cellular Biology* 20 (1). American Society for Microbiology (ASM): 113–25. http://www.ncbi.nlm.nih.gov/pubmed/10594014.

Hoyle, Nathaniel P, and David Ish-Horowicz. 2013. "Transcript Processing and Export Kinetics

Are Rate-Limiting Steps in Expressing Vertebrate Segmentation Clock Genes." *Proceedings of the National Academy of Sciences* 110 (46). National Academy of Sciences: E4316–24. https://doi.org/10.1073/pnas.1308811110.

Hubaud, Alexis, and Olivier Pourquié. 2014. "Signalling Dynamics in Vertebrate Segmentation." *Nature Reviews Molecular Cell Biology* 15 (11). Nature Publishing Group: 709–21. https://doi.org/10.1038/nrm3891.

Huisinga, Kathryn L, and B Franklin Pugh. 2004. "A Genome-Wide Housekeeping Role for TFIID and a Highly Regulated Stress-Related Role for SAGA in Saccharomyces Cerevisiae." *Molecular Cell* 13 (4): 573–85. https://doi.org/10.1016/S1097-2765(04)00087-5.

Imhof, Axel, Xiang-Jiao Yang, Vasily V Ogryzko, Yoshihiro Nakatani, Alan P Wolffe, and Hui Ge. 1997. "Acetylation of General Transcription Factors by Histone Acetyltransferases." *Current Biology* 7 (9): 689–92. https://doi.org/10.1016/S0960-9822(06)00296-X.

Indra, Arup Kumar, William S Mohan, Mattia Frontini, Elisabeth Scheer, Nadia Messaddeq, Daniel Metzger, and Làszlò Tora. 2005. "TAF10 Is Required for the Establishment of Skin Barrier Function in Foetal, but Not in Adult Mouse Epidermis." *Developmental Biology* 285 (1): 28–37. https://doi.org/10.1016/j.ydbio.2005.05.043.

Ingvarsdottir, Kristin, Nevan J. Krogan, N. C. Tolga Emre, Anastasia Wyce, Natalie J. Thompson, Andrew Emili, Timothy R. Hughes, Jack F. Greenblatt, and Shelley L. Berger. 2005. "H2B Ubiquitin Protease Ubp8 and Sgf11 Constitute a Discrete Functional Module within the Saccharomyces Cerevisiae SAGA Complex." *Molecular and Cellular Biology* 25 (3): 1162–72. https://doi.org/10.1128/MCB.25.3.1162-1172.2005.

Inostroza, A., Fred H. Mermelstein, Ilho Ha, William S. Lane, and Danny Reinberg. 1992. "Dr1, a TATA-Binding Protein-Associated Phosphoprotein and Inhibitor of Class II Gene Transcription." *Cell* 70 (3): 477–89. https://doi.org/10.1016/0092-8674(92)90172-9.

Ito, Mitsuhiro, Hirotaka J. Okano, Robert B. Darnell, and Robert G. Roeder. 2002. "The TRAP100 Component of the TRAP/Mediator Complex Is Essential in Broad Transcriptional Events and Development." *EMBO Journal* 21 (13). European Molecular Biology Organization: 3464–75. https://doi.org/10.1093/emboj/cdf348.

Ito, Mitsuhiro, Chao Xing Yuan, Hirotaka J. Okano, Robert B. Darnell, and Robert G. Roeder. 2000. "Involvement of the TRAP220 Component of the TRAP/SMCC Coactivator Complex in Embryonic Development and Thyroid Hormone Action." *Molecular Cell* 5 (4): 683–93. https://doi.org/10.1016/S1097-2765(00)80247-6.

Iwafuchi-Doi, Makiko, and Kenneth S Zaret. 2016. "Cell Fate Control by Pioneer Transcription Factors." *Development (Cambridge, England)* 143 (11). Oxford University Press for The Company of Biologists Limited: 1833–37. https://doi.org/10.1242/dev.133900.

Jacobson, Raymond H., Andreas G. Ladurner, David S. King, and Robert Tjian. 2000. "Structure and Function of a Human TAF(II)250 Double Bromodomain Module." *Science* 288 (5470). American Association for the Advancement of Science: 1422–25. https://doi.org/10.1126/science.288.5470.1422.

Jacq, Xavier, Christel Brou, Yves Lutz, Irwin Davidson, Pierre Chambon, and Laszlo Tora. 1994. “Human TAFII30 Is Present in a Distinct TFIID Complex and Is Required for Transcriptional Activation by the Estrogen Receptor.” *Cell* 79 (1): 107–17. https://doi.org/10.1016/0092-8674(94)90404-9.

Jallow, Zainab, Ulrike G Jacobi, Daniel L Weeks, Igor B Dawid, and Gert Jan C Veenstra. 2004. “Specialized and Redundant Roles of TBP and a Vertebrate-Specific TBP Paralog in Embryonic Gene Regulation in Xenopus.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (37). National Academy of Sciences: 13525–30. https://doi.org/10.1073/pnas.0405536101.

Jenuwein, T., and C D Allis. 2001. “Translating the Histone Code.” *Science*. https://doi.org/10.1126/science.1063127.

Jiang, Cizhong, and B. Franklin Pugh. 2009. “Nucleosome Positioning and Gene Regulation: Advances through Genomics.” *Nature Reviews Genetics* 10 (3): 161–72. https://doi.org/10.1038/nrg2522.

Jiang, Y. W., P. Veschambre, H. Erdjument-Bromage, P. Tempst, J. W. Conaway, R. C. Conaway, and R. D. Kornberg. 1998. “Mammalian Mediator of Transcriptional Regulation and Its Possible Role as an End-Point of Signal Transduction Pathways.” *Proceedings of the National Academy of Sciences* 95 (15). National Academy of Sciences: 8538–43. https://doi.org/10.1073/pnas.95.15.8538.

Jin, Victor X., Gregory A.C. Singer, Francisco J. Agosto-Pérez, Sandya Liyanarachchi, and Ramana V. Davuluri. 2006. “Genome-Wide Analysis of Core Promoter Elements from Conserved Human and Mouse Orthologous Pairs.” *BMC Bioinformatics* 7 (1): 114. https://doi.org/10.1186/1471-2105-7-114.

Joo, Yoo Jin, Scott B Ficarro, Luis M Soares, Yujin Chun, Jarrod A Marto, and Stephen Buratowski. 2017. “Downstream Promoter Interactions of TFIID TAFs Facilitate Transcription Reinitiation.” *Genes and Development* 31 (21). Cold Spring Harbor Laboratory Press: 2162–74. https://doi.org/10.1101/gad.306324.117.

Juven-Gershon, Tamar, Susan Cheng, and James T. Kadonaga. 2006. “Rational Design of a Super Core Promoter That Enhances Gene Expression.” *Nature Methods* 3 (11): 917–22. https://doi.org/10.1038/nmeth937.

Kadonaga, James T. 2012. “Perspectives on the RNA Polymerase II Core Promoter.” *Wiley Interdisciplinary Reviews: Developmental Biology* 1 (1). NIH Public Access: 40–51. https://doi.org/10.1002/wdev.21.

Kamenova, I., L. Warfield, and S. Hahn. 2014. “Mutations on the DNA Binding Surface of TBP Discriminate between Yeast TATA and TATA-Less Gene Transcription.” *Molecular and Cellular Biology* 34 (15): 2929–43. https://doi.org/10.1128/MCB.01685-13.

Kao, C Cheng, Paul M Lieberman, Martin C Schmidt, Qiang Zhou, Rui Pei, and Arnold J Berk. 1990. “Cloning of a Transcriptionally Active Human TATA Binding Factor.” *Science (New York, N.Y.)* 248 (4963): 1646–50. https://doi.org/10.1126/science.2194289.

Kedinger, C., M. Gniazdowski, J. L. Mandel, F. Gissinger, and P. Chambon. 1970. “α-

Amanitin: A Specific Inhibitor of One of Two DNA-Dependent RNA Polymerase Activities from Calf Thymus." *Biochemical and Biophysical Research Communications* 38 (1). Academic Press: 165–71. https://doi.org/10.1016/0006-291X(70)91099-5.

Kelly, W. G., M. E. Dahmus, and G. W. Hart. 1993. "RNA Polymerase II Is a Glycoprotein. Modification of the COOH-Terminal Domain by O-GlcNAc." *Journal of Biological Chemistry* 268 (14): 10416–24. http://www.ncbi.nlm.nih.gov/pubmed/8486697.

Kim, Young Joon, Stefan Björklund, Yang Li, Michael H. Sayre, and Roger D. Kornberg. 1994. "A Multiprotein Mediator of Transcriptional Activation and Its Interaction with the C-Terminal Repeat Domain of RNA Polymerase II." *Cell* 77 (4): 599–608. https://doi.org/10.1016/0092-8674(94)90221-6.

Kimura, Kouichi, Ai Wakamatsu, Yutaka Suzuki, Toshio Ota, Tetsuo Nishikawa, Riu Yamashita, Jun Ichi Yamamoto, et al. 2006. "Diversification of Transcriptional Modulation: Large-Scale Identification and Characterization of Putative Alternative Promoters of Human Genes." *Genome Research* 16 (1): 55–65. https://doi.org/10.1101/gr.4039406.

Klein, Joachim, Mark Nolden, Steven L. Sanders, Jay Kirchner, P. Anthony Weil, and Karsten Melcher. 2003. "Use of a Genetically Introduced Cross-Linker to Identify Interaction Sites of Acidic Activators within Native Transcription Factor IID and SAGA." *Journal of Biological Chemistry* 278 (9). American Society for Biochemistry and Molecular Biology: 6779–86. https://doi.org/10.1074/jbc.M212514200.

Klemm, R D, J A Goodrich, S Zhou, and R Tjian. 1995. "Molecular Cloning and Expression of the 32-KDa Subunit of Human TFIID Reveals Interactions with VP16 and TFIIB That Mediate Transcriptional Activation." *Proc Natl Acad Sci U S A* 92 (13): 5788–92. https://doi.org/10.1073/pnas.92.13.5788.

Knaap, J. A. van der, J. W. Borst, P. C. van der Vliet, R. Gentz, and H. T. M. Timmers. 1997. "Cloning of the CDNA for the TATA-Binding Protein-Associated FactorII170 Subunit of Transcription Factor B-TFIID Reveals Homology to Global Transcription Regulators in Yeast and Drosophila." *Proceedings of the National Academy of Sciences of the United States of America* 94 (22): 11827–32. https://doi.org/10.1073/pnas.94.22.11827.

Kobayashi, N, T G Boyer, and A J Berk. 1995. "A Class of Activation Domains Interacts Directly with TFIIA and Stimulates TFIIA-TFIID-Promoter Complex Assembly." *Molecular and Cellular Biology* 15 (11): 6465–73. https://doi.org/10.1128/MCB.15.11.6465.

Kobayashi, N, P J Horn, S M Sullivan, S J Triezenberg, T G Boyer, and a J Berk. 1998. "DA-Complex Assembly Activity Required for VP16C Transcriptional Activation." *Molecular and Cellular Biology* 18 (7): 4023–31. https://doi.org/10.1128/MCB.18.7.4023.

Köhler, Alwin, Pau Pascual-García, Ana Llopis, Meritxell Zapater, Francesc Posas, Ed Hurt, and Susana Rodríguez-Navarro. 2006. "The MRNA Export Factor Sus1 Is Involved in Spt/Ada/Gcn5 Acetyltransferase-Mediated H2B Deubiquitinylation through Its Interaction with Ubp8 and Sgf11." *Molecular Biology of the Cell* 17 (10): 4228–36. https://doi.org/10.1091/mbc.E06-02-0098.

Köhler, Alwin, Erik Zimmerman, Maren Schneider, Ed Hurt, and Ning Zheng. 2010. “Structural Basis for Assembly and Activation of the Heterotetrameric SAGA Histone H2B Deubiquitinase Module.” *Cell* 141 (4). Elsevier: 606–17. https://doi.org/10.1016/j.cell.2010.04.026.

Kokubo, T, D W Gong, J C Wootton, M Horikoshi, R G Roeder, and Y Nakatani. 1994. “Molecular Cloning of Drosophila TFIID Subunits.” *Nature* 367 (6462): 484–87. https://doi.org/10.1038/367484a0.

Kokubo, T, M J Swanson, J I Nishikawa, a G Hinnebusch, and Y Nakatani. 1998. “The Yeast TAF145 Inhibitory Domain and TFIIA Competitively Bind to TATA-Binding Protein.” *Molecular and Cellular Biology* 18 (2). American Society for Microbiology (ASM): 1003–12. https://doi.org/10.1128/MCB.18.2.1003.

Kosinsky, Robyn L, Florian Wegwitz, Nicole Hellbach, Matthias Dobbelstein, Ahmed Mansouri, Tanja Vogel, Yvonne Begus-Nahrmann, and Steven A Johnsen. 2015. “Usp22 Deficiency Impairs Intestinal Epithelial Lineage Specification in Vivo.” *Oncotarget* 6 (35). Impact Journals, LLC: 37906–18. https://doi.org/10.18632/oncotarget.5412.

Kraemer, S M, R T Ranallo, R C Ogg, and L A Stargell. 2001. “TFIIA Interacts with TFIID via Association with TATA-Binding Protein and TAF40.” *Molecular and Cellular Biology* 21 (5): 1737–46. https://doi.org/10.1128/MCB.21.5.1737-1746.2001.

Krebs, Arnaud R., Krishanpal Karmodiya, Marianne Lindahl-Allen, Kevin Struhl, and Làszlò Tora. 2011. “SAGA and ATAC Histone Acetyl Transferase Complexes Regulate Distinct Sets of Genes and ATAC Defines a Class of P300-Independent Enhancers.” *Molecular Cell* 44 (3). NIH Public Access: 410–23. https://doi.org/10.1016/j.molcel.2011.08.037.

Krol, Aurélie J, Daniela Roellig, M.-L. Dequeant, Olivier Tassy, Earl Glynn, Gaye Hattem, Arcady Mushegian, Andrew C Oates, and O. Pourquie. 2011. “Evolutionary Plasticity of Segmentation Clock Networks.” *Development* 138 (13). Oxford University Press for The Company of Biologists Limited: 2783–92. https://doi.org/10.1242/dev.063834.

Kuo, M. H., J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, S. Y. Roth, and C. D. Allis. 1996. “Transcription-Linked Acetylation by Gcn5p of Histones H3 and H4 at Specific Lysines.” *Nature* 383 (6597): 269–72. https://doi.org/10.1038/383269a0.

Kuras, Laurent, Peter Kosa, Mario Mencia, and Kevin Struhl. 2000. “TAF-Containing and TAF-Independent Forms of Transcriptionally Active TBP in Vivo.” *Science* 288 (5469): 1244–48. https://doi.org/10.1126/science.288.5469.1244.

Kusch, Thomas, Susan M Abmayr, and Jerry L Workman. 2003. “Two Drosophila Ada2 Homologues Function in Different Multiprotein Complexes.” *Molecular and Cellular Biology* 23 (9). American Society for Microbiology: 3305–19. https://doi.org/10.1128/MCB.23.9.3305.

Lagha, Mounia, Jacques P Bothma, Emilia Esposito, Samuel Ng, Laura Stefanik, Chiahao Tsui, Jeffrey Johnston, et al. 2013. “Paused Pol II Coordinates Tissue Morphogenesis in the Drosophila Embryo.” *Cell* 153 (5). NIH Public Access: 976–87. https://doi.org/10.1016/j.cell.2013.04.045.

Lagrange, Thierry, Achillefs N. Kapanidis, Hong Tang, Danny Reinberg, and Richard H. Ebright. 1998. “New Core Promoter Element in RNA Polymerase II-Dependent Transcription: Sequence-Specific DNA Binding by Transcription Factor IIB.” *Genes and Development* 12 (1). Cold Spring Harbor Laboratory Press: 34–44. https://doi.org/10.1101/gad.12.1.34.

Lalonde, M.-E., X. Cheng, and J. Cote. 2014. “Histone Target Selection within Chromatin: An Exemplary Case of Teamwork.” *Genes & Development* 28 (10). Cold Spring Harbor Laboratory Press: 1029–41. https://doi.org/10.1101/gad.236331.113.

Langer, Diana, Igor Martianov, Daniel Alpern, Muriel Rhinn, Céline Keime, Pascal Dollé, Gabrielle Mengus, and Irwin Davidson. 2016. “Essential Role of the TFIID Subunit TAF4 in Murine Embryogenesis and Embryonic Stem Cell Differentiation.” *Nature Communications* 7 (March). Nature Publishing Group. https://doi.org/10.1038/ncomms11063.

Laprade, Lisa, David Rose, and Fred Winston. 2007. “Characterization of New Spt3 and TATA-Binding Protein Mutants of Saccharomyces Cerevisiae: Spt3-TBP Allele-Specific Interactions and Bypass of Spt8.” *Genetics* 177 (4). Genetics: 2007–17. https://doi.org/10.1534/genetics.107.081976.

Larschan, E., and F. Winston. 2001. “The S. Cerevisiae SAGA Complex Functions in Vivo as a Coactivator for Transcriptional Activation by Gal4.” *Genes and Development* 15 (15). Cold Spring Harbor Laboratory Press: 1946–56. https://doi.org/10.1101/gad.911501.

Laybourn, P. J., and M. E. Dahmus. 1990. “Phosphorylation of RNA Polymerase IIA Occurs Subsequent to Interaction with the Promoter and before the Initiation of Transcription.” *Journal of Biological Chemistry* 265 (22): 13165–73. http://www.ncbi.nlm.nih.gov/pubmed/2376591.

Lee, D.-H., N. Gershenzon, M. Gupta, I. P. Ioshikhes, D. Reinberg, and B. A. Lewis. 2005a. “Functional Characterization of Core Promoter Elements: The Downstream Core Element Is Recognized by TAF1.” *Molecular and Cellular Biology* 25 (21). American Society for Microbiology (ASM): 9674–86. https://doi.org/10.1128/MCB.25.21.9674-9686.2005.

Lee, K, L. Florens, S Swanson, M Washburn, and J Workman. 2005b. “The Deubiquitylation Activity of Ubp8 Is Dependent upon Sgf11 and Its Association with the SAGA Complex.” *Molecular and Cellular Biology* 25 (3): 1173–82. https://doi.org/10.1128/MCB.25.3.1173-1182.2005.

Lee, K K, P Prochasson, L Florens, S K Swanson, M P Washburn, and J L Workman. 2004. “Proteomic Analysis of Chromatin-Modifying Complexes in Saccharomyces Cerevisiae Identifies Novel Subunits.” *Biochemical Society Transactions* 32 (Pt 6): 899–903. https://doi.org/10.1042/BST0320899.

Lee, Kenneth K., Mihaela E. Sardiu, Selene K. Swanson, Joshua M. Gilmore, Michael Torok, Patrick A. Grant, Laurence Florens, Jerry L. Workman, and Michael P. Washburn. 2011. “Combinatorial Depletion Analysis to Assemble the Network Architecture of the SAGA and ADA Chromatin Remodeling Complexes.” *Molecular Systems Biology* 7 (1). EMBO Press: 503. https://doi.org/10.1038/msb.2011.40.

Lee, Kenneth K, Selene K Swanson, Laurence Florens, Michael P Washburn, and Jerry L Workman. 2009. "Yeast Sgf73/Ataxin-7 Serves to Anchor the Deubiquitination Module into Both SAGA and Slik(SALSA) HAT Complexes." *Epigenetics & Chromatin* 2 (1): 2. https://doi.org/10.1186/1756-8935-2-2.

Lee, Kenneth K, and Jerry L Workman. 2007. "Histone Acetyltransferase Complexes: One Size Doesn't Fit All." *Nature Reviews. Molecular Cell Biology* 8 (4). Nature Publishing Group: 284–95. https://doi.org/10.1038/nrm2145.

Lee, Tong I., Helen C. Causton, Frank C.P. Holstege, Wu Cheng Shen, Nancy Hannett, Ezra G. Jennings, Fred Winston, Michael R. Green, and Richard A. Young. 2000. "Redundant Roles for the TFIID and SAGA Complexes in Global Transcription." *Nature* 405 (6787). Nature Publishing Group: 701–4. https://doi.org/10.1038/35015104.

Lenstra, Tineke L, Joris J Benschop, Taesoo Kim, Julia M Schulze, Nathalie A C H Brabers, Thanasis Margaritis, Loes A L van de Pasch, et al. 2011. "The Specificity and Topology of Chromatin Interaction Pathways in Yeast." *Molecular Cell* 42 (4). NIH Public Access: 536–49. https://doi.org/10.1016/j.molcel.2011.03.026.

Levine, Michael, Claudia Cattoglio, and Robert Tjian. 2014. "Looping Back to Leap Forward: Transcription Enters a New Era." *Cell* 157 (1). Elsevier Inc.: 13–25. https://doi.org/10.1016/j.cell.2014.02.009.

Lewandoski, Mark, Xin Sun, and Gail R. Martin. 2000. "Fgf8 Signalling from the AER Is Essential for Normal Limb Development." *Nature Genetics* 26 (4): 460–63. https://doi.org/10.1038/82609.

Lewis, Julian. 2003. "Autoinhibition with Transcriptional Delay: A Simple Mechanism for the Zebrafish Somitogenesis Oscillator." *Current Biology* 13 (16): 1398–1408. https://doi.org/10.1016/S0960-9822(03)00534-7.

Li, En, and Yi Zhang. 2014. "DNA Methylation in Mammals." *Cold Spring Harbor Perspectives in Biology* 6 (5). Cold Spring Harbor Laboratory Press: a019133. https://doi.org/10.1101/cshperspect.a019133.

Li, Li, Silvia Sanchez Martinez, Wenxin Hu, Zhe Liu, and Robert Tjian. 2015. "A Specific E3 Ligase/Deubiquitinase Pair Modulates TBP Protein Levels during Muscle Differentiation." *ELife* 4 (September 2015). eLife Sciences Publications Limited: e08536. https://doi.org/10.7554/eLife.08536.

Li, Xuanying, Christopher W. Seidel, Leanne T. Szerszen, Jeffrey J. Lange, Jerry L. Workman, and Susan M. Abmayr. 2017. "Enzymatic Modules of the SAGA Chromatin-Modifying Complex Play Distinct Roles in Drosophila Gene Expression and Development." *Genes and Development* 31 (15). Cold Spring Harbor Laboratory Press: 1588–1600. https://doi.org/10.1101/gad.300988.117.

Lieberman, P. 1994. "Identification of Functional Targets of the Zta Transcriptional Activator by Formation of Stable Preinitiation Complex Intermediates." *Mol Cell Biol* 14 (12): 8365–75. https://doi.org/10.1128/MCB.14.12.8365.

Lieberman, P M, J Ozer, and D B Gursel. 1997. "Requirement for Transcription Factor IIA

(TFIIA)-TFIID Recruitment by an Activator Depends on Promoter Structure and Template Competition." *Mol Cell Biol* 17 (11): 6624–32. https://doi.org/10.1128/MCB.17.11.6624.

Lin, Wenchu, Geraldine Srajer, Yvonne A. Evrard, Huy M. Phan, Yas Furuta, and Sharon Y R Dent. 2007. "Developmental Potential of Gcn5(-/-) Embryonic Stem Cells in Vivo and in Vitro." *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 236 (6): 1547–57. https://doi.org/10.1002/dvdy.21160.

Lin, Zhenghong, Heeyoung Yang, Qingfei Kong, Jinping Li, Sang Myeong Lee, Beixue Gao, Hongxin Dong, et al. 2012. "USP22 Antagonizes P53 Transcriptional Activation by Deubiquitinating Sirt1 to Suppress Cell Apoptosis and Is Required for Mouse Embryonic Development." *Molecular Cell* 46 (4): 484–94. https://doi.org/10.1016/j.molcel.2012.03.024.

Liu, Jianzhao, Yuanxiang Zhu, Guan Zheng Luo, Xinxia Wang, Yanan Yue, Xiaona Wang, Xin Zong, et al. 2016. "Abundant DNA 6mA Methylation during Early Embryogenesis of Zebrafish and Pig." *Nature Communications* 7 (October): 13052. https://doi.org/10.1038/ncomms13052.

Liu, Xin, David A. Bushnell, Daniel Adriano Silva, Xuhui Huang, and Roger D. Kornberg. 2011. "Initiation Complex Structure and Promoter Proofreading." *Science* 333 (6042). American Association for the Advancement of Science: 633–37. https://doi.org/10.1126/science.1206629.

Liu, Xiuli, W. Lee Kraus, and Xiaoying Bai. 2015. "Ready, Pause, Go: Regulation of RNA Polymerase II Pausing and Release by Cellular Signaling Pathways." *Trends in Biochemical Sciences* 40 (9). Elsevier: 516–25. https://doi.org/10.1016/j.tibs.2015.07.003.

Louder, Robert K., Yuan He, José Ramón López-Blanco, Jie Fang, Pablo Chacón, and Eva Nogales. 2016. "Structure of Promoter-Bound TFIID and Model of Human Pre-Initiation Complex Assembly." *Nature* 531 (7596). Nature Publishing Group: 604–9. https://doi.org/10.1038/nature17394.

Luecke, Hans F., and Keith R. Yamamoto. 2005. "The Glucocorticoid Receptor Blocks P-TEFb Recruitment by NFκB to Effect Promoter-Specific Transcriptional Repression." *Genes and Development* 19 (9). Cold Spring Harbor Laboratory Press: 1116–27. https://doi.org/10.1101/gad.1297105.

Luse, Donal S., and Robert G. Roeder. 1980. "Accurate Transcription Initiation on a Purified Mouse β-Globin DNA Fragment in a Cell-Free System." *Cell* 20 (3): 691–99. https://doi.org/10.1016/0092-8674(80)90315-3.

Maldonado, E, I Ha, P Cortes, L Weis, and D Reinberg. 1990. "Factors Involved in Specific Transcription by Mammalian RNA Polymerase II: Role of Transcription Factors IIA, IID, and IIB during Formation of a Transcription-Competent Complex." *Molecular and Cellular Biology* 10 (12): 6335–47. https://doi.org/10.1128/MCB.10.12.6335.

Malecova, Barbora, Alessandra Dall'Agnese, Luca Madaro, Sole Gatto, Paula Coutinho Toto, Sonia Albini, Tammy Ryan, Làszlò Tora, and Pier Lorenzo Puri. 2016. "TBP/TFIID-Dependent Activation of MyoD Target Genes in Skeletal Muscle Cells." *ELife* 5 (February). eLife Sciences Publications Limited: e12534.

https://doi.org/10.7554/eLife.12534.

Malik, S., H. J. Baek, W. Wu, and R. G. Roeder. 2005. “Structural and Functional Characterization of PC2 and RNA Polymerase II-Associated Subpopulations of Metazoan Mediator.” *Molecular and Cellular Biology* 25 (6): 2117–29. https://doi.org/10.1128/MCB.25.6.2117-2129.2005.

Malik, S, D K Lee, and R G Roeder. 1993. “Potential RNA Polymerase II-Induced Interactions of Transcription Factor TFIIB.” *Molecular and Cellular Biology.* 13 (10): 6253–59. https://doi.org/10.1128/MCB.13.10.6253.

Martianov, Igor, Stefano Brancorsini, Anne Gansmuller, Martti Parvinen, Irwin Davidson, and Paolo Sassone-Corsi. 2002a. “Distinct Functions of TBP and TLF/TRF2 during Spermatogenesis: Requirement of TLF for Heterochromatic Chromocenter Formation in Haploid Round Spermatids.” *Development (Cambridge, England)* 129 (4): 945–55. http://www.ncbi.nlm.nih.gov/pubmed/11861477.

Martianov, Igor, Gian Maria Fimia, Andrée Dierich, Martti Parvinen, Paolo Sassone-Corsi, and Irwin Davidson. 2001. “Late Arrest of Spermiogenesis and Germ Cell Apoptosis in Mice Lacking the TBP-like TLF/TRF2 Gene.” *Molecular Cell* 7 (3): 509–15. https://doi.org/10.1016/S1097-2765(01)00198-8.

Martianov, Igor, Stephane Viville, and Irwin Davidson. 2002b. “RNA Polymerase II Transcription in Murine Cells Lacking the TATA Binding Protein.” *Science* 298 (5595): 1036–39. https://doi.org/10.1126/science.1076327.

Martínez-Cerdeño, Verónica, Jessica M. Lemen, Vanessa Chan, Alice Wey, Wenchu Lin, Sharon R. Dent, and Paul S. Knoepfler. 2012. “N-Myc and GCN5 Regulate Significantly Overlapping Transcriptional Programs in Neural Stem Cells.” *PLoS ONE* 7 (6). Public Library of Science: e39456. https://doi.org/10.1371/journal.pone.0039456.

Martinez, E., V. B. Palhan, A. Tjernberg, E. S. Lymar, A. M. Gamper, T. K. Kundu, B. T. Chait, and R. G. Roeder. 2001. “Human STAGA Complex Is a Chromatin-Acetylating Transcription Coactivator That Interacts with Pre-MRNA Splicing and DNA Damage-Binding Factors In Vivo.” *Molecular and Cellular Biology* 21 (20): 6782–95. https://doi.org/10.1128/MCB.21.20.6782-6795.2001.

Martinez, Ernest, Tapas K Kundu, Jack Fu, and Robert G Roeder. 1998. “A Human SPT3-TAF(II)31-GCN5-L Acetylase Complex Distinct from Transcription Factor IID.” *Journal of Biological Chemistry* 273 (37). American Society for Biochemistry and Molecular Biology: 23781–85. https://doi.org/10.1074/jbc.273.37.23781.

Maston, Glenn a, Lihua Julie Zhu, Lynn Chamberlain, Ling Lin, Minggang Fang, and Michael R Green. 2012. “Non-Canonical TAF Complexes Regulate Active Promoters in Human Embryonic Stem Cells.” *ELife* 1 (January): e00068. https://doi.org/10.7554/eLife.00068.

Matsumiya, Marina, Takehito Tomita, Kumiko Yoshioka-Kobayashi, Akihiro Isomura, and Ryoichiro Kageyama. 2018. “ES Cell-Derived Presomitic Mesoderm-like Tissues for Analysis of Synchronized Oscillations in the Segmentation Clock.” *Development (Cambridge, England)* 145 (4). Oxford University Press for The Company of Biologists Limited: dev156836. https://doi.org/10.1242/dev.156836.

Matsuoka, S, M C Edwards, C Bai, S Parker, P Zhang, A Baldini, J W Harper, and S J Elledge. 1995. “P57KIP2, a Structurally Distinct Member of the P21CIP1 Cdk Inhibitor Family, Is a Candidate Tumor Suppressor Gene.” *Genes & Development* 9 (6): 650–62. https://doi.org/10.1101/gad.9.6.650.

Maxon, Mary E, James A Goodrich, and Robert Tjian. 1994. “Transcription Factor IIE Binds Preferentially to RNA Polymerase IIa and Recruits TFIIH: A Model for Promoter Clearance.” *Genes and Development* 8 (5): 515–24. https://doi.org/10.1101/gad.8.5.515.

May, M, G Mengus, A C Lavigne, P Chambon, and I Davidson. 1996. “Human TAF(II28) Promotes Transcriptional Stimulation by Activation Function 2 of the Retinoid X Receptors.” *The EMBO Journal* 15 (12). European Molecular Biology Organization: 3093–3104. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=450252&tool=pmcentrez&rendertype=abstract.

Mayer, Andreas, Heather M Landry, and L Stirling Churchman. 2017. “Pause & Go: From the Discovery of RNA Polymerase Pausing to Its Functional Implications.” *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2017.03.002.

Mengus, G, M May, L Carré, P Chambon, and I Davidson. 1997. “Human TAF(II)135 Potentiates Transcriptional Activation by the AF-2s of the Retinoic Acid, Vitamin D3, and Thyroid Hormone Receptors in Mammalian Cells.” *Genes & Development* 11 (11): 1381–95. http://www.ncbi.nlm.nih.gov/pubmed/9192867.

Mengus, G, M May, X Jacq, A Staub, L Tora, P Chambon, and I Davidson. 1995. “Cloning and Characterization of HTAFII18, HTAFII20 and HTAFII28: Three Subunits of the Human Transcription Factor TFIID.” *The EMBO Journal* 14 (7). European Molecular Biology Organization: 1520–31. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=398239&tool=pmcentrez&rendertype=abstract.

Merino, Alejandro, Knut R. Madden, William S. Lane, James J. Champoux, and Danny Reinberg. 1993. “DNA Topoisomerase I Is Involved in Both Repression and Activation of Transcription.” *Nature* 365 (6443). Nature Publishing Group: 227–32. https://doi.org/10.1038/365227a0.

Metzger, Daniel, Elisabeth Scheer, A. Soldatov, and László Tora. 1999. “Mammalian TAF II 30 Is Required for Cell Cycle Progression and Specific Cellular Differentiation Programmes.” *The EMBO Journal* 18 (17): 4823–34.

Michel, Bertha, Philip Komarnitsky, and Stephen Buratowski. 1998. “Histone-like TAFs Are Essential for Transcription in Vivo.” *Molecular Cell* 2 (5): 663–73. https://doi.org/10.1016/S1097-2765(00)80164-1.

Mizzen, Craig A., Xiang Jiao Yang, Tetsuro Kokubo, James E. Brownell, Andrew J. Bannister, Tom Owen-Hughes, Jerry Workman, et al. 1996. “The TAF(II)250 Subunit of TFIID Has Histone Acetyltransferase Activity.” *Cell* 87 (7): 1261–70. https://doi.org/10.1016/S0092-8674(00)81821-8.

Mohan, William S, Elisabeth Scheer, Olivia Wendling, Daniel Metzger, and Làszlò Tora. 2003.

"TAF10 (TAF(II)30) Is Necessary for TFIID Stability and Early Embryogenesis in Mice." *Molecular and Cellular Biology* 23 (12): 4307–18. https://doi.org/10.1128/MCB.23.12.4307.

Mohibullah, Neeman, and Steven Hahn. 2008. "Site-Specific Cross-Linking of TBP in Vivo and in Vitro Reveals a Direct Functional Interaction with the SAGA Subunit Spt3." *Genes and Development* 22 (21). Cold Spring Harbor Laboratory Press: 2994–3006. https://doi.org/10.1101/gad.1724408.

Moon, Anne M., and Mario R. Capecchi. 2000. "Fgf8 Is Required for Outgrowth and Patterning of the Limbs." *Nature Genetics* 26 (4). Nature Publishing Group: 455–59. https://doi.org/10.1038/82601.

Moore, P a, J Ozer, M Salunek, G Jan, D Zerby, S Campbell, and P M Lieberman. 1999. "A Human TATA Binding Protein-Related Protein with Altered DNA Binding Specificity Inhibits Transcription from Multiple Promoters and Activators." *Molecular and Cellular Biology* 19 (11). American Society for Microbiology: 7610–20. https://doi.org/10.1128/MCB.19.11.7610.

Moqtaderi, Z., Y. Bai, D. Poon, P. A. Well, and K. Struhl. 1996a. "TBP-Associated Factors Are Not Generally Required for Transcriptional Activation in Yeast." *Nature* 383 (6596). Nature Publishing Group: 188–91. https://doi.org/10.1038/383188a0.

Moqtaderi, Z, J D Yale, K Struhl, and S Buratowski. 1996b. "Yeast Homologues of Higher Eukaryotic TFIID Subunits." *Proc Natl Acad Sci U S A* 93 (25). National Academy of Sciences: 14654–58. https://doi.org/10.1073/pnas.93.25.14654.

Moreland, R J, F Tirode, Q Yan, J W Conaway, J M Egly, and R C Conaway. 1999. "A Role for the TFIIH XPB DNA Helicase in Promoter Escape by RNA Polymerase II." *The Journal of Biological Chemistry* 274 (32): 22127–30. http://www.ncbi.nlm.nih.gov/pubmed/10428772.

Muhich, M L, C T Iida, M Horikoshi, R G Roeder, and C S Parker. 1990. "CDNA Clone Encoding Drosophila Transcription Factor TFIID." *Proceedings of the National Academy of Sciences of the United States of America* 87 (23). National Academy of Sciences: 9148–52. https://doi.org/10.1073/pnas.87.23.9148.

Müller, Ferenc, and László Tora. 2009. "TBP2 Is a General Transcription Factor Specialized for Female Germ Cells." *Journal of Biology* 8 (11). BioMed Central: 97. https://doi.org/10.1186/jbiol196.

Müller, Ferenc, Andreas Zaucker, and Làszlò Tora. 2010. "Developmental Regulation of Transcription Initiation: More than Just Changing the Actors." *Current Opinion in Genetics & Development* 20 (5): 533–40. https://doi.org/10.1016/j.gde.2010.06.004.

Muratoglu, Selen, Sofia Georgieva, Gábor Pápai, Elisabeth Scheer, Izzet Enünlü, Orbán Komonyi, Imre Cserpán, et al. 2003. "Two Different Drosophila ADA2 Homologues Are Present in Distinct GCN5 Histone Acetyltransferase-Containing Complexes." *Molecular and Cellular Biology* 23 (1). American Society for Microbiology (ASM): 306–21. https://doi.org/10.1128/MCB.23.1.306-321.2003.

Myers, Lawrence C., Claes M. Gustafsson, David A. Bushnell, Mary Lui, Hediye Erdjument-Bromage, Paul Tempst, and Roger D. Kornberg. 1998. “The Med Proteins of Yeast and Their Function through the RNA Polymerase II Carboxy-Terminal Domain.” *Genes and Development* 12 (1): 45–54. https://doi.org/10.1101/gad.12.1.45.

Näär, Anders M., Pierre A. Beaurang, Sharleen Zhou, Shaji Abraham, William Solomon, and Robert Tjian. 1999. “Composite Co-Activator ARC Mediates Chromatin-Directed Transcriptional Activation.” *Nature* 398 (6730): 828–32. https://doi.org/10.1038/19789.

Nagy, Z., and L. Tora. 2007. “Distinct GCN5/PCAF-Containing Complexes Function as Co-Activators and Are Involved in Transcription Factor and Global Histone Acetylation.” *Oncogene* 26 (37). Nature Publishing Group: 5341–57. https://doi.org/10.1038/sj.onc.1210604.

Nagy, Zita, Anne Riss, Sally Fujiyama, Arnaud Krebs, Meritxell Orpinell, Pascal Jansen, Adrian Cohen, Henk G. Stunnenberg, Shigeaki Kato, and Làszlò Tora. 2010. “The Metazoan ATAC and SAGA Coactivator HAT Complexes Regulate Different Sets of Inducible Target Genes.” *Cellular and Molecular Life Sciences* 67 (4). SP Birkhäuser Verlag Basel: 611–28. https://doi.org/10.1007/s00018-009-0199-8.

Narlikar, Leelavati, and Ivan Ovcharenko. 2009. “Identifying Regulatory Elements in Eukaryotic Genomes.” *Briefings in Functional Genomics & Proteomics* 8 (4). Oxford University Press: 215–30. https://doi.org/10.1093/bfgp/elp014.

Natarajan, Krishnamurthy, Belinda M. Jackson, Eugene Rhee, and Alan G. Hinnebusch. 1998. “YTAFII61 Has a General Role in RNA Polymerase II Transcription and Is Required by Gcn4p to Recruit the SAGA Coactivator Complex.” *Molecular Cell* 2 (5). Elsevier: 683–92. https://doi.org/10.1016/S1097-2765(00)80166-5.

Niederreither, Karen, Peter McCaffery, Ursula C. Dräger, Pierre Chambon, and Pascal Dollé. 1997. “Restricted Expression and Retinoic Acid-Induced Downregulation of the Retinaldehyde Dehydrogenase Type 2 (RALDH-2) Gene during Mouse Development.” *Mechanisms of Development* 62 (1): 67–78. https://doi.org/10.1016/S0925-4773(96)00653-3.

Nishimura, Kohei, Tatsuo Fukagawa, Haruhiko Takisawa, Tatsuo Kakimoto, and Masato Kanemaki. 2009. “An Auxin-Based Degron System for the Rapid Depletion of Proteins in Nonplant Cells.” *Nature Methods* 6 (12). Nature Publishing Group: 917–22. https://doi.org/10.1038/nmeth.1401.

Nitanda, Yasuhide, Takaaki Matsui, Tatsuro Matta, Aya Higami, Kenji Kohno, Yasukazu Nakahata, and Yasumasa Bessho. 2014. “3′-UTR-Dependent Regulation of MRNA Turnover Is Critical for Differential Distribution Patterns of Cyclic Gene MRNAs.” *FEBS Journal* 281 (1): 146–56. https://doi.org/10.1111/febs.12582.

Niwa, Yasutaka, Yoshito Masamizu, Tianxiao Liu, Rika Nakayama, Chu Xia Deng, and Ryoichiro Kageyama. 2007. “The Initiation and Propagation of Hes7 Oscillation Are Cooperatively Regulated by Fgf and Notch Signaling in the Somite Segmentation Clock.” *Developmental Cell* 13 (2). Cell Press: 298–304. https://doi.org/10.1016/j.devcel.2007.07.013.

Noro, Miyuki, Hiroki Yuguchi, Taeko Sato, Takanobu Tsuihiji, Sayuri Yonei-Tamura, Hitoshi Yokoyama, Yoshio Wakamatsu, and Koji Tamura. 2011. "Role of Paraxial Mesoderm in Limb/Flank Regionalization of the Trunk Lateral Plate." *Developmental Dynamics* 240 (7). Wiley-Blackwell: 1639–49. https://doi.org/10.1002/dvdy.22666.

Nuland, Rick van, Andrea W. Schram, Frederik M. A. van Schaik, Pascal W. T. C. Jansen, Michiel Vermeulen, and H. T. Marc Timmers. 2013. "Multivalent Engagement of TFIID to Nucleosomes." Edited by Pierre-Antoine Defossez. *PLoS ONE* 8 (9): e73495. https://doi.org/10.1371/journal.pone.0073495.

O'Brien, T, and R Tjian. 1998. "Functional Analysis of the Human TAFII250 N-Terminal Kinase Domain." *Molecular Cell* 1 (6): 905–11. https://doi.org/10.1016/S1097-2765(00)80089-1.

Oates, Andrew C., and Robert K. Ho. 2002. "Hairy/E(Spl)-Related (Her) Genes Are Central Components of the Segmentation Oscillator and Display Redundancy with the Delta/Notch Signaling Pathway in the Formation of Anterior Segmental Boundaries in the Zebrafish." *Development* 129 (12): 2929–46. http://dev.biologists.org/content/129/12/2929%5Cnhttp://dev.biologists.org/content/129/12/2929.full.pdf%5Cnhttp://dev.biologists.org/content/129/12/2929.short%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/12050140.

Ogryzko, Vasily V, Tomohiro Kotani, Xiaolong Zhang, R Louis Schiltz, Tazuko Howard, Xiang Jiao Yang, Bruce H Howard, Jun Qin, and Yoshihiro Nakatani. 1998. "Histone-like TAFS within the PCAF Histone Acetylase Complex." *Cell* 94 (1): 35–44. https://doi.org/10.1016/S0092-8674(00)81219-2.

Ohbayashi, Tetsuya, Yasutaka Makino, and Taka Aki Tamura. 1999. "Identification of a Mouse TBP-like Protein (TLP) Distantly Related to the Drosophila TBP-Related Factor." *Nucleic Acids Research* 27 (3): 750–55. https://doi.org/10.1093/nar/27.3.750.

Ohkubo, Y., C. Chiang, and J. L.R. Rubenstein. 2002. "Coordinate Regulation and Synergistic Actions of BMP4, SHH and FGF8 in the Rostral Prosencephalon Regulate Morphogenesis of the Telencephalic and Optic Vesicles." *Neuroscience* 111 (1): 1–17. https://doi.org/10.1016/S0306-4522(01)00616-9.

Ohkuma, Y, S Hashimoto, C K Wang, M Horikoshi, and Robert G Roeder. 1995. "Analysis of the Role of TFIIE in Basal Transcription and TFIIH-Mediated Carboxy-Terminal Domain Phosphorylation through Structure-Function Studies of TFIIE-α." *Molecular and Cellular Biology* 15 (9): 4856–66. https://doi.org/10.1128/MCB.15.9.4856.

Ohkuma, Yoshiaki, and Robert G. Roeder. 1994. "Regulation of TFIIH ATPase and Kinase Activities by TFIIE during Active Initiation Complex Formation." *Nature* 368 (6467). Nature Publishing Group: 160–63. https://doi.org/10.1038/368160a0.

Orphanides, G, T Lagrange, and D Reinberg. 1996. "The General Transcription Factors of RNA Polymerase II." *Genes & Development* 10 (21): 2657–83. http://www.ncbi.nlm.nih.gov/pubmed/8946909.

Orpinell, Meritxell, Marjorie Fournier, Anne Riss, Zita Nagy, Arnaud R Krebs, Mattia Frontini, and Làszlò Tora. 2010. "The ATAC Acetyl Transferase Complex Controls Mitotic

Progression by Targeting Non-Histone Substrates." *The EMBO Journal* 29 (14): 2381–94. https://doi.org/10.1038/emboj.2010.125.

Ozer, Josef, Paul A. Moore, Arthur H. Bolden, Arianna Lee, Craig A. Rosen, and Paul M. Lieberman. 1994. "Molecular Cloning of the Small (γ) Subunit of Human TFIIA Reveals Functions Critical for Activated Transcription." *Genes and Development* 8 (19): 2324–35. https://doi.org/10.1101/gad.8.19.2324.

Palmeirim, Isabel, Domingos Henrique, David Ish-Horowicz, and Olivier Pourquié. 1997. "Avian Hairy Gene Expression Identifies a Molecular Clock Linked to Vertebrate Segmentation and Somitogenesis." *Cell* 91 (5): 639–48. https://doi.org/10.1016/S0092-8674(00)80451-1.

Pan, G, and J Greenblatt. 1994. "Initiation of Transcription by RNA Polymerase II Is Limited by Melting of the Promoter DNA in the Region Immediately Upstream of the Initiation Site." *The Journal of Biological Chemistry* 269 (48): 30101–4. http://www.ncbi.nlm.nih.gov/pubmed/7982911.

Pankotai, Tibor, Orbán Komonyi, László Bodai, Zsuzsanna Ujfaludi, Selen Muratoglu, Anita Ciurciu, László Tora, János Szabad, and Imre Boros. 2005. "The Homologous Drosophila Transcriptional Adaptors ADA2a and ADA2b Are Both Required for Normal Development but Have Different Functions." *Molecular and Cellular Biology* 25 (18). American Society for Microbiology: 8215–27. https://doi.org/10.1128/MCB.25.18.8215-8227.2005.

Papadopoulos, Petros, Laura Gutiérrez, Jeroen Demmers, Elisabeth Scheer, Farzin Pourfarzad, Dimitris N. Papageorgiou, Elena Karkoulia, et al. 2015. "TAF10 Interacts with the GATA1 Transcription Factor and Controls Mouse Erythropoiesis." *Molecular and Cellular Biology* 35 (12). American Society for Microbiology: 2103–18. https://doi.org/10.1128/MCB.01370-14.

Perantoni, Alan O, Olga Timofeeva, Florence Naillat, Charmaine Richman, Sangeeta Pajni-Underwood, Catherine Wilson, Seppo Vainio, Lee F Dove, and Mark Lewandoski. 2005. "Inactivation of FGF8 in Early Mesoderm Reveals an Essential Role in Kidney Development." *Development (Cambridge, England)* 132 (17): 3859–71. https://doi.org/10.1242/dev.01945.

Perletti, Lucia, Eliezer Kopf, Lucie Carré, and Irwin Davidson. 2001. "Coordinate Regulation of RARgamma2, TBP, and TAFII135 by Targeted Proteolysis during Retinoic Acid-Induced Differentiation of F9 Embryonal Carcinoma Cells." *BMC Molecular Biology* 2: 4. https://doi.org/10.1186/1471-2199-2-4.

Persengiev, Stephan P, Xiaochun Zhu, Bharat L Dixit, Glenn A Maston, Ellen L W Kittler, and Michael R Green. 2003. "TRF3, a TATA-Box-Binding Protein-Related Factor, Is Vertebrate-Specific and Widely Expressed." *Proceedings of the National Academy of Sciences of the United States of America* 100 (25): 14887–91. https://doi.org/10.1073/pnas.2036440100.

Peterson, M G, J Inostroza, M E Maxon, O Flores, a Admon, D Reinberg, and R Tjian. 1991. "Structure and Functional Properties of Human General Transcription Factor IIE." *Nature* 354 (6352). Nature Publishing Group: 369–73. https://doi.org/10.1038/354369a0.

Peterson, M G, N Tanese, B F Pugh, and R Tjian. 1990a. “Functional Domains and Upstream Activation Properties of Cloned Human TATA Binding Protein.” *Science (New York, N.Y.)* 248 (4963): 1625–30. https://doi.org/10.1126/science.844-b.

———. 1990b. “Functional Domains and Upstream Activation Properties of Cloned Human TATA Binding Protein.” *Science (New York, N.Y.)* 248 (4963): 1625–30. https://doi.org/10.1126/science.844-b.

Petit, Florence, Karen E. Sears, and Nadav Ahituv. 2017. “Limb Development: A Paradigm of Gene Regulation.” *Nature Reviews Genetics* 18 (4): 245–58. https://doi.org/10.1038/nrg.2016.167.

Pijnappel, W W M Pim, Daniel Esch, Marijke P a Baltissen, Guangming Wu, Nikolai Mischerikow, Atze J Bergsma, Erik van der Wal, et al. 2013. “A Central Role for TFIID in the Pluripotent Transcription Circuitry.” *Nature* 495 (7442): 516–19. https://doi.org/10.1038/nature11970.

Plaschka, C., M. Hantsche, C. Dienemann, C. Burzinski, J. Plitzko, and P. Cramer. 2016. “Transcription Initiation Complex Structures Elucidate DNA Opening.” *Nature* 533 (7603). Nature Publishing Group: 353–58. https://doi.org/10.1038/nature17990.

Plet, A, D Eick, and J M Blanchard. 1995. “Elongation and Premature Termination of Transcripts Initiated from C-Fos and C-Myc Promoters Show Dissimilar Patterns.” *Oncogene* 10 (2): 319–28. http://www.ncbi.nlm.nih.gov/pubmed/7838531.

Pointud, J.-C., G. Mengus, S. Brancorsini, L. Monaco, M. Parvinen, P. Sassone-Corsi, and I. Davidson. 2003. “The Intracellular Localisation of TAF7L, a Paralogue of Transcription Factor TFIID Subunit TAF7, Is Developmentally Regulated during Male Germ-Cell Differentiation.” *Journal of Cell Science* 116 (9): 1847–58. https://doi.org/10.1242/jcs.00391.

Poon, D, Y Bai, A M Campbell, S Bjorklund, Y J Kim, S Zhou, R D Kornberg, and P A Weil. 1995. “Identification and Characterization of a TFIID-like Multiprotein Complex from Saccharomyces Cerevisiae.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (18): 8224–28. https://doi.org/10.1073/pnas.92.18.8224.

Poon, David, Allyson M. Campbell, Yu Bai, and P. Anthony Weil. 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.” *Journal of Biological Chemistry* 269 (37): 23135–40. http://www.ncbi.nlm.nih.gov/pubmed/8083216.

Powell, David W, Connie M Weaver, Jennifer L Jennings, K Jill McAfee, Yue He, P Anthony Weil, and Andrew J Link. 2004. “Cluster Analysis of Mass Spectrometry Data Reveals a Novel Component of SAGA.” *Molecular and Cellular Biology* 24 (16): 7249–59. https://doi.org/10.1128/MCB.24.16.7249-7259.2004.

Pray-Grant, M G, D Schieltz, S J McMahon, J M Wood, E L Kennedy, R G Cook, J L Workman, J R Yates, and P A Grant. 2002. “The Novel SLIK Histone Acetyltransferase Complex Functions in the Yeast Retrograde Response Pathway.” *Molecular and Cellular Biology* 22 (24): 8774–86. https://doi.org/10.1128/mcb.22.24.8774-8786.2002.

Presnyak, Vladimir, Najwa Alhusaini, Ying Hsin Chen, Sophie Martin, Nathan Morris, Nicholas Kline, Sara Olson, et al. 2015. "Codon Optimality Is a Major Determinant of MRNA Stability." *Cell* 160 (6). Cell Press: 1111–24. https://doi.org/10.1016/j.cell.2015.02.029.

Pugh, B. F., and R. Tjian. 1991. "Transcription from a TATA-Less Promoter Requires a Multisubunit TFIID Complex." *Genes and Development* 5 (11): 1935–45. https://doi.org/10.1101/gad.5.11.1935.

Pugh, B Franklin, and Robert Tjian. 1990. "Mechanism of Transcriptional Activation by Sp1: Evidence for Coactivators." *Cell* 61 (7): 1187–97. https://doi.org/10.1016/0092-8674(90)90683-6.

Qi, D., J. Larsson, and M. Mannervik. 2004. "Drosophila Ada2b Is Required for Viability and Normal Histone H3 Acetylation." *Molecular and Cellular Biology* 24 (18): 8080–89. https://doi.org/10.1128/MCB.24.18.8080-8089.2004.

Qureshi, Sohail A., and Stephen P. Jackson. 1998. "Sequence-Specific DNA Binding by the S. Shibatae TFIIB Homolog, TFB, and Its Effect on Promoter Strength." *Molecular Cell* 1 (3). Elsevier: 389–400. https://doi.org/10.1016/S1097-2765(00)80039-8.

Rabenstein, M D, S Zhou, J T Lis, and R Tjian. 1999. "TATA Box-Binding Protein (TBP)-Related Factor 2 (TRF2), a Third Member of the TBP Family." *Proceedings of the National Academy of Sciences of the United States of America* 96 (9). National Academy of Sciences: 4791–96. https://doi.org/10.1073/pnas.96.9.4791.

Rachez, Christophe, Bryan D. Lemon, Zalman Suldan, Virginia Bromleigh, Matthew Gamble, Anders M. Näär, Hediye Erdjument-Bromage, Paul Tempst, and Leonard P. Freedman. 1999. "Ligand-Dependent Transcription Activation by Nuclear Receptors Requires the DRIP Complex." *Nature* 398 (6730). Nature Publishing Group: 824–28. https://doi.org/10.1038/19783.

Radhakrishnan, Aditya, Ying Hsin Chen, Sophie Martin, Najwa Alhusaini, Rachel Green, and Jeff Coller. 2016. "The DEAD-Box Protein Dhh1p Couples MRNA Decay and Translation by Monitoring Codon Optimality." *Cell* 167 (1). Cell Press: 122–132.e9. https://doi.org/10.1016/j.cell.2016.08.053.

Rädle, Bernd, Andrzej J. Rutkowski, Zsolt Ruzsics, Caroline C. Friedel, Ulrich H. Koszinowski, and Lars Dölken. 2013. "Metabolic Labeling of Newly Transcribed RNA for High Resolution Gene Expression Profiling of RNA Synthesis, Processing and Decay in Cell Culture." *Journal of Visualized Experiments*, no. 78 (August). MyJoVE Corporation. https://doi.org/10.3791/50195.

Ranish, J. A., and S. Hahn. 1991. "The Yeast General Transcription Factor TFIIA Is Composed of Two Polypeptide Subunits." *Journal of Biological Chemistry* 266 (29): 19320–27. http://www.ncbi.nlm.nih.gov/pubmed/1918049.

Ranish, J A, W S Lane, and S Hahn. 1992. "Isolation of Two Genes That Encode Subunits of the Yeast Transcription Factor IIA." *Science* 255 (5048): 1127–29. http://www.hubmed.org/display.cgi?uids=1546313.

Reese, J C, L Apone, S S Walker, L A Griffin, and M R Green. 1994. “Yeast TAFIIS in a Multisubunit Complex Required for Activated Transcription.” *Nature* 371 (6497): 523–27. https://doi.org/10.1038/371523a0.

Reinberg, D., M. Horikoshi, and R. G. Roeder. 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.” *Journal of Biological Chemistry* 262 (7): 3322–30. https://doi.org/10.1073/pnas.87.23.9158.

Renner, Dan B., Yuki Yamaguchi, Tadashi Wada, Hiroshi Handa, and David H. Price. 2001. “A Highly Purified RNA Polymerase II Elongation Control System.” *Journal of Biological Chemistry* 276 (45): 42601–9. https://doi.org/10.1074/jbc.M104967200.

Riley, Maurisa F, Matthew S Bochter, Kanu Wahi, Gerard J Nuovo, and Susan E Cole. 2013. “Mir-125a-5p-Mediated Regulation of Lfng Is Essential for the Avian Segmentation Clock.” *Developmental Cell* 24 (5). Elsevier: 554–61. https://doi.org/10.1016/j.devcel.2013.01.024.

Riss, Anne, Elisabeth Scheer, Mathilde Joint, Simon Trowitzsch, Imre Berger, and László Tora. 2015. “Subunits of ADA-Two-A-Containing (ATAC) or Spt-Ada-Gcn5-Acetyltrasferase (SAGA) Coactivator Complexes Enhance the Acetyltransferase Activity of GCN5.” *Journal of Biological Chemistry* 290 (48): 28997–9. https://doi.org/10.1074/jbc.M115.668533.

Rodríguez-Navarro, Susana, Tamás Fischer, Ming Juan Luo, Oreto Antúnez, Susanne Brettschneider, Johannes Lechner, Jose E. Pérez-Ortín, Robin Reed, and Ed Hurt. 2004. “Sus1, a Functional Component of the SAGA Histone Acetylase Complex and the Nuclear Pore-Associated MRNA Export Machinery.” *Cell* 116 (1): 75–86. https://doi.org/10.1016/S0092-8674(03)01025-0.

Roeder, Robert G., and William J. Rutter. 1969. “Multiple Forms of DNA-Dependent RNA Polymerase in Eukaryotic Organisms.” *Nature* 224 (5216). Nature Publishing Group: 234–37. https://doi.org/10.1038/224234a0.

Rojo-Niersbach, Eileen, Takako Furukawa, and Naoko Tanese. 1999. “Genetic Dissection of HTAF(II)130 Defines a Hydrophobic Surface, Required for Interaction with Glutamine-Rich Activators.” *Journal of Biological Chemistry* 274 (47): 33778–84. https://doi.org/10.1074/jbc.274.47.33778.

Rougvie, a E, and J T Lis. 1990. “Postinitiation Transcriptional Control in Drosophila Melanogaster.” *Molecular and Cellular Biology* 10 (11): 6041–45. https://doi.org/10.1128/MCB.10.11.6041.Updated.

Ruppert, S, E H Wang, and Robert Tjian. 1993. “Cloning and Expression of Human TAFII250: A TBP-Associated Factor Implicated in Cell-Cycle Regulation.” *Nature* 362 (6416): 175–79. https://doi.org/10.1038/362175a0.

Ryu, Soojin, Sharleen Zhou, Andreas G. Ladurner, and Robert Tjian. 1999. “The Transcriptional Cofactor Complex CRSP Is Required for Activity of the Enhancer-Binding Protein Sp1.” *Nature* 397 (6718). Nature Publishing Group: 446–50.

https://doi.org/10.1038/17141.

Saga, Yumiko, Naomi Hata, Haruhiko Koseki, and Makoto M. Taketo. 1997. "Mesp2: A Novel Mouse Gene Expressed in the Presegmented Mesoderm and Essential for Segmentation Initiation." *Genes and Development* 11 (14): 1827–39. https://doi.org/10.1101/gad.11.14.1827.

Sakai, Yasuo, Chikara Meno, Hideta Fujii, Jinsuke Nishino, Hidetaka Shiratori, Yukio Saijoh, Janet Rossant, and Hiroshi Hamada. 2001. "The Retinoic Acid-Inactivating Enzyme CYP26 Is Essential for Establishing an Uneven Distribution of Retinoic Acid along the Anterio-Posterior Axis within the Mouse Embryo." *Genes and Development* 15 (2): 213–25. https://doi.org/10.1101/gad.851501.

Saleh, Ayman, Verna Lang, Robert Cook, and Christopher J. Brandl. 1997. "Identification of Native Complexes Containing the Yeast Coactivator/Repressor Proteins NGG1/ADA3 and ADA2." *Journal of Biological Chemistry* 272 (9). American Society for Biochemistry and Molecular Biology: 5571–78. https://doi.org/10.1074/jbc.272.9.5571.

Saluja, D, M F Vassallo, and N Tanese. 1998. "Distinct Subdomains of Human TAFII130 Are Required for Interactions with Glutamine-Rich Transcriptional Activators." *Molecular and Cellular Biology* 18 (10): 5734–43. https://doi.org/10.1128/MCB.18.10.5734.

Samara, Nadine L., Ajit B. Datta, Christopher E. Berndsen, Xiangbin Zhang, Tingting Yao, Robert E. Cohen, and Cynthia Wolberger. 2010. "Structural Insights into the Assembly and Function of the SAGA Deubiquitinating Module." *Science* 328 (5981): 1025–29. https://doi.org/10.1126/science.1190049.

Samuels, M, A Fire, and P A Sharp. 1982. "Separation and Characterization of Factors Mediating Accurate Transcription by RNA Polymerase II." *J. Biol. Chem.* 257 (23): 14419–27. http://www.jbc.org/content/257/23/14419.abstract.

Sanders, Steven L, Krassimira A Garbett, and P Anthony Weil. 2002. "Molecular Characterization of Saccharomyces Cerevisiae TFIID." *Molecular and Cellular Biology* 22 (16). American Society for Microbiology: 6000–6013. https://doi.org/10.1128/MCB.22.16.6000-6013.2002.

Sanders, Steven L, Edward R Klebanow, and P Anthony Weil. 1999. "TAF25p, a Non-Histone-like Subunit of TFIID and SAGA Complexes, Is Essential for Total MRNA Gene Transcription in Vivo." *Journal of Biological Chemistry* 274 (27). American Society for Biochemistry and Molecular Biology: 18847–50. https://doi.org/10.1074/jbc.274.27.18847.

Sasai, Yoshiki, Ryoichiro Kageyama, Yoshiaki Tagawa, Ryuichi Shigemoto, and Shigetada Nakanishi. 1992. "Two Mammalian Helix-Loop-Helix Factors Structurally Related to Drosophila Hairy and Enhancer of Split." *Genes and Development* 6 (12 PART B). Cold Spring Harbor Laboratory Press: 2620–34. https://doi.org/10.1101/gad.6.12b.2620.

Sawadogo, M, and R G Roeder. 1985. "Factors Involved in Specific Transcription by Human RNA Polymerase II: Analysis by a Rapid and Quantitative in Vitro Assay." *Proceedings of the National Academy of Sciences of the United States of America* 82 (13): 4394–98. https://doi.org/3925456.

Schilbach, S., M. Hantsche, D. Tegunov, C. Dienemann, C. Wigge, H. Urlaub, and P. Cramer. 2017. “Structures of Transcription Pre-Initiation Complex with TFIIH and Mediator.” *Nature* 551 (7679). Nature Publishing Group: 204–9. https://doi.org/10.1038/nature24282.

Schmidt, Corina, Bodo Christ, Ketan Patel, and Beate Brand-Saberi. 1998. “Experimental Induction of BMP-4 Expression Leads to Apoptosis in the Paraxial and Lateral Plate Mesoderm.” *Developmental Biology* 202 (2): 253–63. https://doi.org/10.1006/dbio.1998.9011.

Schmidt, M C, C C Kao, R Pei, and a J Berk. 1989. “Yeast TATA-Box Transcription Factor Gene.” *Proceedings of the National Academy of Sciences of the United States of America* 86 (20). National Academy of Sciences: 7785–89. https://doi.org/10.1073/pnas.86.20.7785.

Schwalb, Björn, Margaux Michel, Benedikt Zacher, Katja Frü Hauf, Carina Demel, Achim Tresch, Julien Gagneur, and Patrick Cramer. 2016. “TT-Seq Maps the Human Transient Transcriptome.” *Science* 352 (6290). American Association for the Advancement of Science: 1225–28. https://doi.org/10.1126/science.aad9841.

Schweikhard, V., C. Meng, K. Murakami, C. D. Kaplan, R. D. Kornberg, and S. M. Block. 2014. “Transcription Factors TFIIF and TFIIS Promote Transcript Elongation by RNA Polymerase II by Synergistic and Independent Mechanisms.” *Proceedings of the National Academy of Sciences* 111 (18): 6642–47. https://doi.org/10.1073/pnas.1405181111.

Sekiguchi, T, E Noguchi, T Hayashida, T Nakashima, H Toyoshima, T Nishimoto, and T Hunter. 1996. “D-Type Cyclin Expression Is Decreased and P21 and P27 CDK Inhibitor Expression Is Increased When TsBN462 CCG1/TAFII250 Mutant Cells Arrest in G1 at the Restrictive Temperature.” *Genes to Cells : Devoted to Molecular & Cellular Mechanisms* 1 (1993). Blackwell Science Ltd: 687–705. https://doi.org/10.1046/j.1365-2443.1996.00259.x.

Serizawa, Hiroaki, Joan Weliky Conaway, and Ronald C Conaway. 1994. “An Oligomeric Form of the Large Subunit of Transcription Factor (TF) IIE Activates Phosphorylation of the RNA Polymerase II Carboxyl-Terminal Domain by TFIIH.” *Journal of Biological Chemistry* 269 (32): 20750–56. http://www.ncbi.nlm.nih.gov/pubmed/8051177.

Serizay, Jacques, and Julie Ahringer. 2018. “Genome Organization at Different Scales: Nature, Formation and Function.” *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2018.03.009.

Sermwittayawong, Decha, and Song Tan. 2006. “SAGA Binds TBP via Its Spt8 Subunit in Competition with DNA: Implications for TBP Recruitment.” *EMBO Journal* 25 (16). European Molecular Biology Organization: 3791–3800. https://doi.org/10.1038/sj.emboj.7601265.

Setiaputra, Dheva, James D Ross, Shan Lu, Derrick T Cheng, Meng Qiu Dong, and Calvin K Yip. 2015. “Conformational Flexibility and Subunit Arrangement of the Modular Yeast Spt-Ada-Gcn5 Acetyltransferase Complex.” *Journal of Biological Chemistry* 290 (16). American Society for Biochemistry and Molecular Biology: 10057–70. https://doi.org/10.1074/jbc.M114.624684.

Sharov, Grigory, Karine Voltz, Alexandre Durand, Olga Kolesnikova, Gabor Papai, Alexander G Myasnikov, Annick Dejaegere, Adam Ben Shem, and Patrick Schultz. 2017. "Structure of the Transcription Activator Target Tra1 within the Chromatin Modifying Complex SAGA." *Nature Communications* 8 (1). Nature Publishing Group: 1556. https://doi.org/10.1038/s41467-017-01564-7.

Shimogori, T. 2004. "Embryonic Signaling Centers Expressing BMP, WNT and FGF Proteins Interact to Pattern the Cerebral Cortex." *Development* 131 (22). The Company of Biologists Ltd: 5639–47. https://doi.org/10.1242/dev.01428.

Solow, Steven, Moreh Salunek, Robert Ryan, and Paul M. Lieberman. 2001. "TAFII 250 Phosphorylates Human Transcription Factor IIA on Serine Residues Important for TBP Binding and Transcription Activity." *Journal of Biological Chemistry* 276 (19): 15886–92. https://doi.org/10.1074/jbc.M009385200.

Sopta, M., R. W. Carthew, and J. Greenblatt. 1985. "Isolation of Three Proteins That Bind to Mammalian RNA Polymerase II." *Journal of Biological Chemistry* 260 (18): 10353–60. http://www.ncbi.nlm.nih.gov/pubmed/3860504.

Soriano, P. 1999. "Generalized LacZ Expression with the ROSA26 Cre Reporter Strain." *Nature Genetics* 21 (1): 70–71. https://doi.org/10.1038/5007.

Soutoglou, Evi, Màté a Demény, Elisabeth Scheer, Giulia Fienga, Paolo Sassone-Corsi, and Làszlò Tora. 2005. "The Nuclear Import of TAF10 Is Regulated by One of Its Three Histone Fold Domain-Containing Interaction Partners." *Molecular and Cellular Biology* 25 (10): 4092–4104. https://doi.org/10.1128/MCB.25.10.4092-4104.2005.

Soutourina, Julie. 2017. "Transcription Regulation by the Mediator Complex." *Nature Reviews Molecular Cell Biology*, December. Nature Publishing Group. https://doi.org/10.1038/nrm.2017.115.

Soutourina, Julie, Sandra Wydau, Yves Ambroise, Claire Boschiero, and Michel Werner. 2011. "Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo." *Science* 331 (6023): 1451–54. https://doi.org/10.1126/science.1200188.

Spedale, Gianpiero, H. Th Marc Timmers, and W. W.M.Pim Pijnappel. 2012. "ATAC-King the Complexity of SAGA during Evolution." *Genes and Development* 26 (6). Cold Spring Harbor Laboratory Press: 527–41. https://doi.org/10.1101/gad.184705.111.

Spitz, François, and Eileen E. M. Furlong. 2012. "Transcription Factors: From Enhancer Binding to Developmental Control." *Nature Reviews Genetics* 13 (9). Nature Publishing Group: 613–26. https://doi.org/10.1038/nrg3207.

Stargell, L a, and K Struhl. 1995. "The TBP-TFIIA Interaction in the Response to Acidic Activators in Vivo." *Science (New York, N.Y.)* 269 (5220): 75–78. https://doi.org/10.1126/science.7604282.

Stargell, Laurie A., Zarmik Moqtaderi, David R. Dorris, Ryan C. Ogg, and Kevin Struhl. 2000. "TFIIA Has Activator-Dependent and Core Promoter Functions in Vivo." *Journal of Biological Chemistry* 275 (17): 12374–80. https://doi.org/10.1074/jbc.275.17.12374.

Starr, D. Barry, Barbara C. Hoopes, and Diane K. Hawley. 1995. "DNA Bending Is an

Important Component of Site-Specific Recognition by the TATA Binding Protein." *Journal of Molecular Biology* 250 (4): 434–46. https://doi.org/10.1006/jmbi.1995.0388.

Stauber, Michael, Christine Laclef, Annalisa Vezzaro, Mahalia E. Page, and David Ish-Horowicz. 2012. "Modifying Transcript Lengths of Cycling Mouse Segmentation Genes." *Mechanisms of Development* 129 (1–4): 61–72. https://doi.org/10.1016/j.mod.2012.01.006.

Sterner, David E., Patrick A. Grant, Shannon M. Roberts, Laura J. Duggan, Rimma Belotserkovskaya, Lisa A. Pacella, Fred Winston, Jerry L. Workman, and Shelley L. Berger. 1999. "Functional Organization of the Yeast SAGA Complex: Distinct Components Involved in Structural Integrity, Nucleosome Acetylation, and TATA-Binding Protein Interaction." *Molecular and Cellular Biology* 19 (1). American Society for Microbiology: 86–98. https://doi.org/10.1128/MCB.19.1.86.

Sterner, David E, Rimma Belotserkovskaya, and Shelley L Berger. 2002. "SALSA, a Variant of Yeast SAGA, Contains Truncated Spt7, Which Correlates with Activated Transcription." *Proceedings of the National Academy of Sciences of the United States of America* 99 (18). National Academy of Sciences: 11622–27. https://doi.org/10.1073/pnas.182021199.

Strobl, L J, and D Eick. 1992. "Hold Back of RNA Polymerase II at the Transcription Start Site Mediates Down-Regulation of c-Myc in Vivo." *The EMBO Journal* 11 (9). European Molecular Biology Organization: 3307–14. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=556865&tool=pmcentrez&rendertype=abstract.

Suganuma, Tamaki, José L. Gutiérrez, Bing Li, Laurence Florens, Selene K. Swanson, Michael P. Washburn, Susan M. Abmayr, and Jerry L. Workman. 2008. "ATAC Is a Double Histone Acetyltransferase Complex That Stimulates Nucleosome Sliding." *Nature Structural and Molecular Biology* 15 (4): 364–72. https://doi.org/10.1038/nsmb.1397.

Suganuma, Tamaki, Arcady Mushegian, Selene K. Swanson, Susan M. Abmayr, Laurence Florens, Michael P. Washburn, and Jerry L. Workman. 2010. "The ATAC Acetyltransferase Complex Coordinates MAP Kinases to Regulate JNK Target Genes." *Cell* 142 (5). Cell Press: 726–36. https://doi.org/10.1016/j.cell.2010.07.045.

Suka, Noriyuki, Yuko Suka, Andrew A. Carmen, Jiansheng Wu, and Michael Grunstein. 2001. "Highly Specific Antibodies Determine Histone Acetylation Site Usage in Yeast Heterochromatin and Euchromatin." *Molecular Cell* 8 (2): 473–79. https://doi.org/10.1016/S1097-2765(01)00301-X.

Sumimoto, H, Y Ohkuma, A Hoffmann, S Shimasaki, M Horikoshi, and R G Roeder. 1991. "Structural Motifs and Potential Sigma Factor Homologies in the Large Subunit of Human General Transcription Factor TFIIE." *Nature* 354 (6352). Nature Publishing Group: 398–401. https://doi.org/10.1038/354401a0.

Sun, Mai, Björn Schwalb, Nicole Pirkl, Kerstin C. Maier, Arne Schenk, Henrik Failmezger, Achim Tresch, and Patrick Cramer. 2013. "Global Analysis of Eukaryotic MRNA Degradation Reveals Xrn1-Dependent Buffering of Transcript Levels." *Molecular Cell* 52 (1). Cell Press: 52–62. https://doi.org/10.1016/j.molcel.2013.09.010.

Sun, Mai, Björn Schwalb, Daniel Schulz, Nicole Pirkl, Stefanie Etzold, Laurent Larivière, Kerstin C Maier, Martin Seizl, Achim Tresch, and Patrick Cramer. 2012. "Comparative Dynamic Transcriptome Analysis (CDTA) Reveals Mutual Feedback between MRNA Synthesis and Degradation." *Genome Research* 22 (7). Cold Spring Harbor Laboratory Press: 1350–59. https://doi.org/10.1101/gr.130161.111.

Sun, Xin, Francesca V. Mariani, and Gail R. Martin. 2002. "Functions of FGF Signalling from the Apical Ectodermal Ridge in Limb Development." *Nature* 418 (6897): 501–8. https://doi.org/10.1038/nature00902.

Takashima, Y., T. Ohtsuka, A. Gonzalez, H. Miyachi, and R. Kageyama. 2011. "Intronic Delay Is Essential for Oscillatory Expression in the Segmentation Clock." *Proceedings of the National Academy of Sciences* 108 (8). National Academy of Sciences: 3300–3305. https://doi.org/10.1073/pnas.1014418108.

Takebayashi, K., Y. Sasai, Y. Sakai, T. Watanabe, S. Nakanishi, and R. Kageyama. 1994. "Structure, Chromosomal Locus, and Promoter Analysis of the Gene Encoding the Mouse Helix-Loop-Helix Factor HES-1. Negative Autoregulation through the Multiple N Box Elements." *Journal of Biological Chemistry* 269 (7): 5150–56. http://www.ncbi.nlm.nih.gov/pubmed/7906273.

Tanese, N, B F Pugh, and R Tjian. 1991. "Coactivators for a Proline-Rich Activator Purified from the Multisubunit Human TFIID Complex." *Genes and Development* 5 (12 A): 2212–24. https://doi.org/10.1101/gad.5.12a.2212.

Tanese, N, D Saluja, M F Vassallo, J L Chen, and A Admon. 1996. "Molecular Cloning and Analysis of Two Subunits of the Human TFIID Complex: HTAFII130 and HTAFII100." *Proceedings of the National Academy of Sciences of the United States of America* 93 (November): 13611–13616. https://doi.org/10.1073/pnas.93.24.13611.

Tatarakis, Antonis, Thanasis Margaritis, Celia Pilar Martinez-Jimenez, Antigone Kouskouti, William S Mohan, Anna Haroniti, Dimitris Kafetzopoulos, Làszlò Tora, and Iannis Talianidis. 2008. "Dominant and Redundant Functions of TFIID Involved in the Regulation of Hepatic Genes." *Molecular Cell* 31 (4): 531–43. https://doi.org/10.1016/j.molcel.2008.07.013.

Tessarz, Peter, and Tony Kouzarides. 2014. "Histone Core Modifications Regulating Nucleosome Structure and Dynamics." *Nature Reviews Molecular Cell Biology* 15 (11): 703–8. https://doi.org/10.1038/nrm3890.

Thomas, Mary C., and Cheng Ming Chiang. 2006. "The General Transcription Machinery and General Cofactors." *Critical Reviews in Biochemistry and Molecular Biology*. https://doi.org/10.1080/10409230600648736.

Thompson, C M, and R a Young. 1995. "General Requirement for RNA Polymerase II Holoenzymes in Vivo." *Proceedings of the National Academy of Sciences of the United States of America* 92 (10): 4587–90. https://doi.org/10.1073/pnas.92.10.4587.

Timmers, H. Th Marc, and Phillip A. Sharp. 1991. "The Mammalian TFIID Protein Is Present in Two Functionally Distinct Complexes." *Genes and Development* 5 (11): 1946–56. https://doi.org/10.1101/gad.5.11.1946.

Timmers, H T, R E Meyers, and P a Sharp. 1992. “Composition of Transcription Factor B-TFIID.” *Proceedings of the National Academy of Sciences of the United States of America* 89 (September): 8140–44. https://doi.org/10.1073/pnas.89.17.8140.

Tirode, Franck, Didier Busso, Frédéric Coin, and Jean-Marc Egly. 1999. “Reconstitution of the Transcription Factor TFIIH: Assignment of Functions for the Three Enzymatic Subunits, XPB, XPD, and Cdk7.” *Molecular Cell* 3 (1). Cell Press: 87–95. https://doi.org/10.1016/S1097-2765(00)80177-X.

Tora, Làszlò. 2002. “A Unified Nomenclature for TATA Box Binding Protein (TBP)-Associated Factors (TAFs) Involved in RNA Polymerase II Transcription.” *Genes and Development* 16 (6). Cold Spring Harbor Laboratory Press: 673–75. https://doi.org/10.1101/gad.976402.

Toshiro, Hayashida, Sekiguchi Takeshi, Noguchi Eishi, Sunamoto Hidetoshi, Ohba Tomoyuki, and Nishimoto Takeharu. 1994. “The CCG1/TAFII250 Gene Is Mutated in Thermosensitive G1 Mutants of the BHK21 Cell Line Derived from Golden Hamster.” *Gene* 141 (2): 267–70. https://doi.org/10.1016/0378-1119(94)90583-5.

Trcek, Tatjana, Daniel R. Larson, Alberto Moldón, Charles C. Query, and Robert H. Singer. 2011. “Single-Molecule MRNA Decay Measurements Reveal Promoter-Regulated MRNA Stability in Yeast.” *Cell* 147 (7). Cell Press: 1484–97. https://doi.org/10.1016/j.cell.2011.11.051.

Trowitzsch, Simon, Cristina Viola, Elisabeth Scheer, Sascha Conic, Virginie Chavant, Marjorie Fournier, Gabor Papai, et al. 2015a. “Cytoplasmic TAF2-TAF8-TAF10 Complex Provides Evidence for Nuclear Holo-TFIID Assembly from Preformed Submodules.” *Nature Communications* 6 (January). Nature Publishing Group: 6011. https://doi.org/10.1038/ncomms7011.

———. 2015b. “Cytoplasmic TAF2–TAF8–TAF10 Complex Provides Evidence for Nuclear Holo–TFIID Assembly from Preformed Submodules.” *Nature Communications* 6 (1). Nature Publishing Group: 6011. https://doi.org/10.1038/ncomms7011.

Tudor, Matthew, Peter J. Murray, Christina Onufryk, Rudolf Jaenisch, and Richard A. Young. 1999. “Ubiquitous Expression and Embryonic Requirement for RNA Polymerase II Coactivator Subunit Srb7 in Mice.” *Genes and Development* 13 (18). Cold Spring Harbor Laboratory Press: 2365–68. https://doi.org/10.1101/gad.13.18.2365.

Tzouanacou, Elena, Amélie Wegener, Filip J. Wymeersch, Valerie Wilson, and Jean François Nicolas. 2009. “Redefining the Progression of Lineage Segregations during Mammalian Embryogenesis by Clonal Analysis.” *Developmental Cell* 17 (3): 365–76. https://doi.org/10.1016/j.devcel.2009.08.002.

Újvári, Andrea, Mahadeb Pal, and Donal S. Luse. 2011. “The Functions of TFIIF during Initiation and Transcript Elongation Are Differentially Affected by Phosphorylation by Casein Kinase 2.” *Journal of Biological Chemistry* 286 (26): 23160–67. https://doi.org/10.1074/jbc.M110.205658.

Veloso, Artur, Killeen S. Kirkconnell, Brian Magnuson, Benjamin Biewen, Michelle T. Paulsen, Thomas E. Wilson, and Mats Ljungman. 2014. “Rate of Elongation by RNA

Polymerase II Is Associated with Specific Gene Features and Epigenetic Modifications." *Genome Research* 24 (6). Cold Spring Harbor Laboratory Press: 896–905. https://doi.org/10.1101/gr.171405.113.

Ventura, Andrea, David G. Kirsch, Margaret E. McLaughlin, David A. Tuveson, Jan Grimm, Laura Lintault, Jamie Newman, Elizabeth E. Reczek, Ralph Weissleder, and Tyler Jacks. 2007. "Restoration of P53 Function Leads to Tumour Regression in Vivo." *Nature* 445 (7128): 661–65. https://doi.org/10.1038/nature05541.

Vermeulen, Michiel, H. Christian Eberl, Filomena Matarese, Hendrik Marks, Sergei Denissov, Falk Butter, Kenneth K. Lee, et al. 2010. "Quantitative Interaction Proteomics and Genome-Wide Profiling of Epigenetic Histone Marks and Their Readers." *Cell* 142 (6): 967–80. https://doi.org/10.1016/j.cell.2010.08.020.

Vermeulen, Michiel, Klaas W. Mulder, Sergei Denissov, W. W M Pim Pijnappel, Frederik M A van Schaik, Radhika A. Varier, Marijke P A Baltissen, Henk G. Stunnenberg, Matthias Mann, and H. Th Marc Timmers. 2007. "Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4." *Cell* 131 (1). Elsevier: 58–69. https://doi.org/10.1016/j.cell.2007.08.016.

Vermot, Julien, Jabier Gallego Llamas, Valérie Fraulob, Karen Niederreither, Pierre Chambon, and Pascal Dollé. 2005. "Retinoic Acid Controls the Bilateral Symmetry of Somite Formation in the Mouse Embryo." *Science* 308 (5721). American Association for the Advancement of Science: 563–66. https://doi.org/10.1126/science.1108363.

Verrijzer, C Peter, Kyoko Yokomori, J L Chen, and Robert Tjian. 1994. "Drosophila TAFII150: Similarity to Yeast Gene TSM-1 and Specific Binding to Core Promoter DNA." *Science (New York, N.Y.)* 264 (5161): 933–41. http://www.ncbi.nlm.nih.gov/pubmed/8178153.

Vignali, M, A H Hassan, K E Neely, and J L Workman. 2000. "ATP-Dependent Chromatin-Remodeling Complexes." *Molecular and Cellular Biology* 20 (6). American Society for Microbiology (ASM): 1899–1910. http://www.ncbi.nlm.nih.gov/pubmed/10688638.

Vilhais-Neto, Gonçalo C., Marjorie Fournier, Jean Luc Plassat, Mihaela E. Sardiu, Anita Saraf, Jean Marie Garnier, Mitsuji Maruhashi, Laurence Florens, Michael P. Washburn, and Olivier Pourquié. 2017. "The WHHERE Coactivator Complex Is Required for Retinoic Acid-Dependent Regulation of Embryonic Symmetry." *Nature Communications* 8 (1). Nature Publishing Group: 728. https://doi.org/10.1038/s41467-017-00593-6.

Vo ngoc, Long, Yuan-Liang Wang, George A. Kassavetis, and James T. Kadonaga. 2017. "The Punctilious RNA Polymerase II Core Promoter." *Genes & Development* 31 (13): 1289–1301. https://doi.org/10.1101/gad.303149.117.

Voss, a K, T Thomas, P Petrou, K Anastassiadis, H Schöler, and P Gruss. 2000. "Taube Nuss Is a Novel Gene Essential for the Survival of Pluripotent Cells of Early Mouse Embryos." *Development (Cambridge, England)* 127 (24): 5449–61. http://www.ncbi.nlm.nih.gov/pubmed/11076765.

Walker, Amy K., Joel H. Rothman, Yang Shi, and T. Keith Blackwell. 2001. "Distinct Requirements for C.Elegans TAFIIs in Early Embryonic Transcription." *EMBO Journal* 20 (18): 5269–79. https://doi.org/10.1093/emboj/20.18.5269.

Walker, Scott S., Wu Cheng Shen, Joseph C. Reese, Lynne M. Apone, and Michael R. Green. 1997. "Yeast TAF(II)145 Required for Transcription of G1/S Cyclin Genes and Regulated by the Cellular Growth State." *Cell* 90 (4): 607–14. https://doi.org/10.1016/S0092-8674(00)80522-X.

Wang, E H, and R Tjian. 1994. "Promoter-Selective Transcriptional Defect in Cell Cycle Mutant Ts13 Rescued by HTAFII250." *Science (New York, N.Y.)* 263 (5148). American Association for the Advancement of Science: 811–14. https://doi.org/10.1126/science.8303298.

Wang, P Jeremy, and David C Page. 2002. "Functional Substitution for TAF(II)250 by a Retroposed Homolog That Is Expressed in Human Spermatogenesis." *Human Molecular Genetics* 11 (19). Oxford University Press: 2341–46. https://doi.org/10.1093/hmg/11.19.2341.

Wang, Yuan Liang, Francesco Faiola, Muyu Xu, Songqin Pan, and Ernest Martinez. 2008. "Human ATAC Is a GCN5/PCAF-Containing Acetylase Complex with a Novel NC2-like Histone Fold Module That Interacts with the TATA-Binding Protein." *Journal of Biological Chemistry* 283 (49): 33808–15. https://doi.org/10.1074/jbc.M806936200.

Warfield, Linda, Srinivas Ramachandran, Tiago Baptista, Didier Devys, Laszlo Tora, and Steven Hahn. 2017. "Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIID." *Molecular Cell* 68 (1). Cell Press: 118–129.e5. https://doi.org/10.1016/j.molcel.2017.08.014.

Warrier, Sunita, Samer Nuwayhid, Julia A. Sabatino, Kelsey F. Sugrue, and Irene E. Zohn. 2017. "Supt20 Is Required for Development of the Axial Skeleton." *Developmental Biology* 421 (2): 245–57. https://doi.org/10.1016/j.ydbio.2016.11.009.

Watanabe, Tomomichi, Kazuhiro Hayashi, Aki Tanaka, Tadashi Furumoto, Fumio Hanaoka, and Yoshiaki Ohkuma. 2003. "The Carboxy Terminus of the Small Subunit of TFIIE Regulates the Transition from Transcription Initiation to Elongation by RNA Polymerase II." *Molecular and Cellular Biology* 23 (8). American Society for Microbiology (ASM): 2914–26. https://doi.org/10.1128/MCB.23.8.2914-2926.2003.

Weake, Vikki M., Kenneth K. Lee, Sebastián Guelman, Chia Hui Lin, Christopher Seidel, Susan M. Abmayr, and Jerry L. Workman. 2008. "SAGA-Mediated H2B Deubiquitination Controls the Development of Neuronal Connectivity in the Drosophila Visual System." *EMBO Journal* 27 (2). EMBO Press: 394–405. https://doi.org/10.1038/sj.emboj.7601966.

Weake, Vikki M., Selene K. Swanson, Arcady Mushegian, Laurence Florens, Michael P. Washburn, Susan M. Abmayr, and Jerry L. Workman. 2009. "A Novel Histone Fold Domain-Containing Protein That Replaces TAF6 in Drosophila SAGA Is Required for SAGA-Dependent Gene Expression." *Genes and Development* 23 (24). Cold Spring Harbor Laboratory Press: 2818–23. https://doi.org/10.1101/gad.1846409.

Weideman, Christian A., Robert C. Netter, Lawrence R. Benjamin, John J. McAllister, Lumelle A. Schmiedekamp, Robert A. Coleman, and B. Franklin Pugh. 1997. "Dynamic Interplay of TFIIA, TBP and TATA DNA." *Journal of Molecular Biology* 271 (1): 61–75. https://doi.org/10.1006/jmbi.1997.1152.

Weil, P. A., Donal S. Luse, Jacqueline Segall, and Robert G. Roeder. 1979. “Selective and Accurate Initiation of Transcription at the Ad2 Major Late Promotor in a Soluble System Dependent on Purified Rna Polymerase Ii and Dna.” *Cell* 18 (2): 469–84. https://doi.org/10.1016/0092-8674(79)90065-5.

Weinzierl, R O, S Ruppert, B D Dynlacht, N Tanese, and R Tjian. 1993. “Cloning and Expression of Drosophila TAFII60 and Human TAFII70 Reveal Conserved Interactions with Other Subunits of TFIID.” *The EMBO Journal* 12 (13): 5303–9. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=413796&tool=pmcentrez&rendertype=abstract.

Weiss, Samuel B., and Leonard Gladstone. 1959. “A Mammalian System for The Incorporation of Cytidine Triphosphate into Ribonucleic Acid.” *Journal of the American Chemical Society* 81 (15). American Chemical Society: 4118–19. https://doi.org/10.1021/ja01524a087.

Wertent, Sebastiaan, André Mitschler, Christophe Romier, Yann Gaël Gangloff, Sylvie Thuault, Irwin Davidson, and Dino Moras. 2002. “Crystal Structure of a Subcomplex of Human Transcription Factor TFIID Formed by TATA Binding Protein-Associated Factors HTAF4 (HTAFII135) and HTAF12 (HTAFII20).” *Journal of Biological Chemistry* 277 (47). American Society for Biochemistry and Molecular Biology: 45502–9. https://doi.org/10.1074/jbc.M206587200.

Westerling, T., E. Kuuluvainen, and T. P. Makela. 2007. “Cdk8 Is Essential for Preimplantation Mouse Development.” *Molecular and Cellular Biology* 27 (17): 6177–82. https://doi.org/10.1128/MCB.01302-06.

Wieczorek, E., M. Brand, X. Jacq, and L. Tora. 1998. “Function of TAF(II)-Containing Complex without TBP in Transcription by RNA Polymerase II.” *Nature* 393 (6681). Nature Publishing Group: 187–91. https://doi.org/10.1038/30283.

Wild, Thomas, and Patrick Cramer. 2012. “Biogenesis of Multisubunit RNA Polymerases.” *Trends in Biochemical Sciences* 37 (3): 99–105. https://doi.org/10.1016/j.tibs.2011.12.001.

Winfree, a.T. 1989. “From Clocks to Chaos: The Rhythms of Life.” *Mathematical Biosciences*. https://doi.org/10.1016/0025-5564(89)90037-0.

Woychik, N A, and R A Young. 1990. “RNA Polymerase II: Subunit Structure and Function.” *Trends Biochem Sci* 15 (9): 347–51. http://www.ncbi.nlm.nih.gov/pubmed/1700503.

Wright, K. J., M. T. Marr, and R. Tjian. 2006. “TAF4 Nucleates a Core Subcomplex of TFIID and Mediates Activated Transcription from a TATA-Less Promoter.” *Proceedings of the National Academy of Sciences* 103 (33): 12347–52. https://doi.org/10.1073/pnas.0605499103.

Wu, Pei-Yun Jenny, and Fred Winston. 2002. “Analysis of Spt7 Function in the Saccharomyces Cerevisiae SAGA Coactivator Complex.” *Molecular and Cellular Biology* 22 (15): 5367–79. https://doi.org/10.1128/MCB.22.15.5367-5379.2002.

Wu, Pei Yun Jenny, Christine Ruhlmann, Fred Winston, and Patrick Schultz. 2004. “Molecular

Architecture of the S. Cerevisiae SAGA Complex." *Molecular Cell* 15 (2). Elsevier: 199–208. https://doi.org/10.1016/j.molcel.2004.06.005.

Wu, Tao P., Tao Wang, Matthew G. Seetin, Yongquan Lai, Shijia Zhu, Kaixuan Lin, Yifei Liu, et al. 2016. "DNA Methylation on N6-Adenine in Mammalian Embryonic Stem Cells." *Nature* 532 (7599): 329–33. https://doi.org/10.1038/nature17640.

Xiao, LiJuan, MinJung Kim, and Jeff DeJong. 2006. "Developmental and Cell Type-Specific Regulation of Core Promoter Transcription Factors in Germ Cells of Frogs and Mice." *Gene Expression Patterns* 6 (4): 409–19. https://doi.org/10.1016/j.modgep.2005.09.005.

Xie, Xiaoling, Tetsuro Kokubo, Steven L. Cohen, Urooj A. Mirza, Alexander Hoffmann, Brian T. Chait, Robert G. Roeder, Yoshihiro Nakatani, and Stephen K. Burley. 1996. "Structural Similarity between TAFs and the Heterotetrameric Core of the Histone Octamer." *Nature* 380 (6572). Nature Publishing Group: 316–22. https://doi.org/10.1038/380316a0.

Xiong, Yue, Gregory J. Hannon, Hui Zhang, David Casso, Ryuji Kobayashi, and David Beach. 1993. "P21 Is a Universal Inhibitor of Cyclin Kinases." *Nature* 366 (6456): 701–4. https://doi.org/10.1038/366701a0.

Xu, Wanting, Diane G. Edmondson, Yvonne A. Evrard, Maki Wakamiya, Richard R. Behringer, and Sharon Y. Roth. 2000. "Loss of Gcn512 Leads to Increased Apoptosis and Mesodermal Defects during Mouse Development." *Nature Genetics* 26 (2). Nature Publishing Group: 229–32. https://doi.org/10.1038/79973.

Xu, Yali, Joseph P Milazzo, Tim D.D. Somerville, Yusuke Tarumoto, Yu-Han Huang, Elizabeth L Ostrander, John E Wilkinson, Grant A Challen, and Christopher R Vakoc. 2018. "A TFIID-SAGA Perturbation That Targets MYB and Suppresses Acute Myeloid Leukemia." *Cancer Cell* 33 (1). Elsevier: 13–28.e8. https://doi.org/10.1016/j.ccell.2017.12.002.

Yamaguchi, Yuki, Hirotaka Shibata, and Hiroshi Handa. 2013. "Transcription Elongation Factors DSIF and NELF: Promoter-Proximal Pausing and Beyond." *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829 (1): 98–104. https://doi.org/10.1016/j.bbagrm.2012.11.007.

Yamauchi, T., J Yamauchi, T Kuwata, T Tamura, T Yamashita, N Bae, H Westphal, K Ozato, and Y Nakatani. 2000. "Distinct but Overlapping Roles of Histone Acetylase PCAF and of the Closely Related PCAF-B/GCN5 in Mouse Embryogenesis." *Proceedings of the National Academy of Sciences* 97 (21). National Academy of Sciences: 11303–6. https://doi.org/10.1073/pnas.97.21.11303.

Yan, Q, R J Moreland, J W Conaway, and R C Conaway. 1999. "Dual Roles for Transcription Factor IIF in Promoter Escape by RNA Polymerase II." *The Journal of Biological Chemistry* 274 (50). American Society for Biochemistry and Molecular Biology: 35668–75. https://doi.org/10.1074/JBC.274.50.35668.

Yang, Chuhu, Eugene Bolotin, Tao Jiang, Frances M. Sladek, and Ernest Martinez. 2007. "Prevalence of the Initiator over the TATA Box in Human and Yeast Genes and Identification of DNA Motifs Enriched in Human TATA-Less Core Promoters." *Gene* 389 (1). NIH Public Access: 52–65. https://doi.org/10.1016/j.gene.2006.09.029.

Yang, Xiang Jiao, Vasily V. Ogryzko, Jun Ichi Nishikawa, Bruce H. Howard, and Yoshihiro Nakatani. 1996. “A P300/CPB-Associated Factor That Competes with the Adenoviral Oncoprotein E1A.” *Nature* 382 (6589). Nature Publishing Group: 319–24. https://doi.org/10.1038/382319a0.

Yokomori, K., A. Admon, J. A. Goodrich, J. L. Chen, and R. Tjian. 1993a. “Drosophila TFIIA-L Is Processed into Two Subunits That Are Associated with the TBP/TAF Complex.” *Genes and Development* 7 (11): 2235–45. https://doi.org/10.1101/gad.7.11.2235.

Yokomori, K, J L Chen, A Admon, S Zhou, and R Tjian. 1993b. “Molecular Cloning and Characterization of DTAFII30 Alpha and DTAFII30 Beta: Two Small Subunits of Drosophila TFIID.” *Genes & Development* 7 (12B): 2587–97. http://www.ncbi.nlm.nih.gov/pubmed/8276241.

Yokomori, Kyoko, Martin P. Zeidler, Jin Long Chen, C. Peter Verrijzer, Marek Mlodzik, and Robert Tjian. 1994. “Drosophila TFIIA Directs Cooperative DNA Binding with TBP and Mediates Transcriptional Activation.” *Genes and Development* 8 (19): 2313–23. https://doi.org/10.1101/gad.8.19.2313.

Young, Richard A. 1991. “RNA Polymerase II.” *Annual Review of Biochemistry* 60 (1): 689–715. https://doi.org/10.1146/annurev.bi.60.070191.003353.

Yudkovsky, Natalya, Jeffrey A. Ranish, and Steven Hahn. 2000. “A Transcription Reinitiation Intermediate That Is Stabilized by Activator.” *Nature* 408 (6809). Nature Publishing Group: 225–29. https://doi.org/10.1038/35041603.

Zawel, Leigh, K Prasanna Kumar, and Danny Reinberg. 1995. “Recycling of the General Transcription Factors during RNA Polymerase II Transcription.” *Genes & Development* 9 (12): 1479–90. https://doi.org/10.1101/gad.9.12.1479.

Zhang, Di, Tarja Leena Penttila, Patricia L. Morris, and Robert G. Roeder. 2001. “Cell- and Stage-Specific High-Level Expression of TBP-Related Factor 2 (TRF2) during Mouse Spermatogenesis.” *Mechanisms of Development* 106 (1–2): 203–5. https://doi.org/10.1016/S0925-4773(01)00439-7.

Zhang, Guoqiang, Hua Huang, Di Liu, Ying Cheng, Xiaoling Liu, Wenxin Zhang, Ruichuan Yin, et al. 2015. “N6-Methyladenine DNA Modification in Drosophila.” *Cell* 161 (4): 893–906. https://doi.org/10.1016/j.cell.2015.04.018.

Zhang, Xiao Yong, Maya Varthi, Stephen M. Sykes, Charles Phillips, Claude Warzecha, Wenting Zhu, Anastasia Wyce, Alan W. Thorne, Shelley L. Berger, and Steven B. McMahon. 2008. “The Putative Cancer Stem Cell Marker USP22 Is a Subunit of the Human SAGA Complex Required for Activated Transcription and Cell-Cycle Progression.” *Molecular Cell* 29 (1): 102–11. https://doi.org/10.1016/j.molcel.2007.12.015.

Zhang, Zhen, Jason R. O’Rourke, Michael T. McManus, Mark Lewandoski, Brian D. Harfe, and Xin Sun. 2011. “The MicroRNA-Processing Enzyme Dicer Is Dispensable for Somite Segmentation but Essential for Limb Bud Positioning.” *Developmental Biology* 351 (2). Academic Press: 254–65. https://doi.org/10.1016/j.ydbio.2011.01.005.

Zhao, Yue, Guillaume Lang, Saya Ito, Jacques Bonnet, Eric Metzger, Shun Sawatsubashi, Eriko Suzuki, et al. 2008. "A TFTC/STAGA Module Mediates Histone H2A and H2B Deubiquitination, Coactivates Nuclear Receptors, and Counteracts Heterochromatin Silencing." *Molecular Cell* 29 (1): 92–101. https://doi.org/10.1016/j.molcel.2007.12.011.

Zhou, H., I. Grubisic, K. Zheng, Y. He, P. J. Wang, T. Kaplan, and R. Tjian. 2013a. "Taf7l Cooperates with Trf2 to Regulate Spermiogenesis." *Proceedings of the National Academy of Sciences* 110 (42): 16886–91. https://doi.org/10.1073/pnas.1317034110.

Zhou, H., T. Kaplan, Y. Li, I. Grubisic, Z. Zhang, P.J. Wang, M.B. Eisen, and R. Tjian. 2013b. "Dual Functions of TAF7L in Adipocyte Differentiation." *ELife* 2013 (2): e00170. https://doi.org/10.7554/eLife.00170.

Zhou, Haiying, Bo Wan, Ivan Grubisic, Tommy Kaplan, and Robert Tjian. 2014. "TAF7L Modulates Brown Adipose Tissue Formation." *ELife* 2014 (3). eLife Sciences Publications Ltd: 1–11. https://doi.org/10.7554/eLife.02811.

Zhou, J, J Zwicker, P Szymanski, M Levine, and R Tjian. 1998. "TAFII Mutations Disrupt Dorsal Activation in the Drosophila Embryo." *Proceedings of the National Academy of Sciences of the United States of America* 95 (23): 13483–88. https://doi.org/10.1073/pnas.95.23.13483.

Zhou, Qiang, Paul M. Lieberman, Thomas G. Boyer, and Arnold J. Berk. 1992. "Holo-TFIID Supports Transcriptional Stimulation by Diverse Activators and from a TATA-Less Promoter." *Genes and Development* 6 (10): 1964–74. https://doi.org/10.1101/gad.6.10.1964.

Zohn, Irene E., Yingqiu Li, Edward Y. Skolnik, Kathryn V. Anderson, Jiahuai Han, and Lee Niswander. 2006. "P38 and a P38-Interacting Protein Are Critical for Downregulation of E-Cadherin during Mouse Gastrulation." *Cell* 125 (5): 957–69. https://doi.org/10.1016/j.cell.2006.03.048.

Zsindely, Nóra, Tibor Pankotai, Zsuzsanna Újfaludi, Dániel Lakatos, Orbán Komonyi, László Bodai, László Tora, and Imre M. Boros. 2009. "The Loss of Histone H3 Lysine 9 Acetylation Due to DSAGA-Specific DAda2b Mutation Influences the Expression of Only a Small Subset of Genes." *Nucleic Acids Research* 37 (20). Oxford University Press: 6665–80. https://doi.org/10.1093/nar/gkp722.

Zybailov, Boris, Amber L. Mosley, Mihaela E. Sardiu, Michael K. Coleman, Laurence Florens, and Michael P. Washburn. 2006. "Statistical Analysis of Membrane Proteome Expression Changes in Saccharomyces Cerevisiae." *Journal of Proteome Research* 5 (9): 2339–47. https://doi.org/10.1021/pr060161n.

# ANNEXES

**Annexe I**

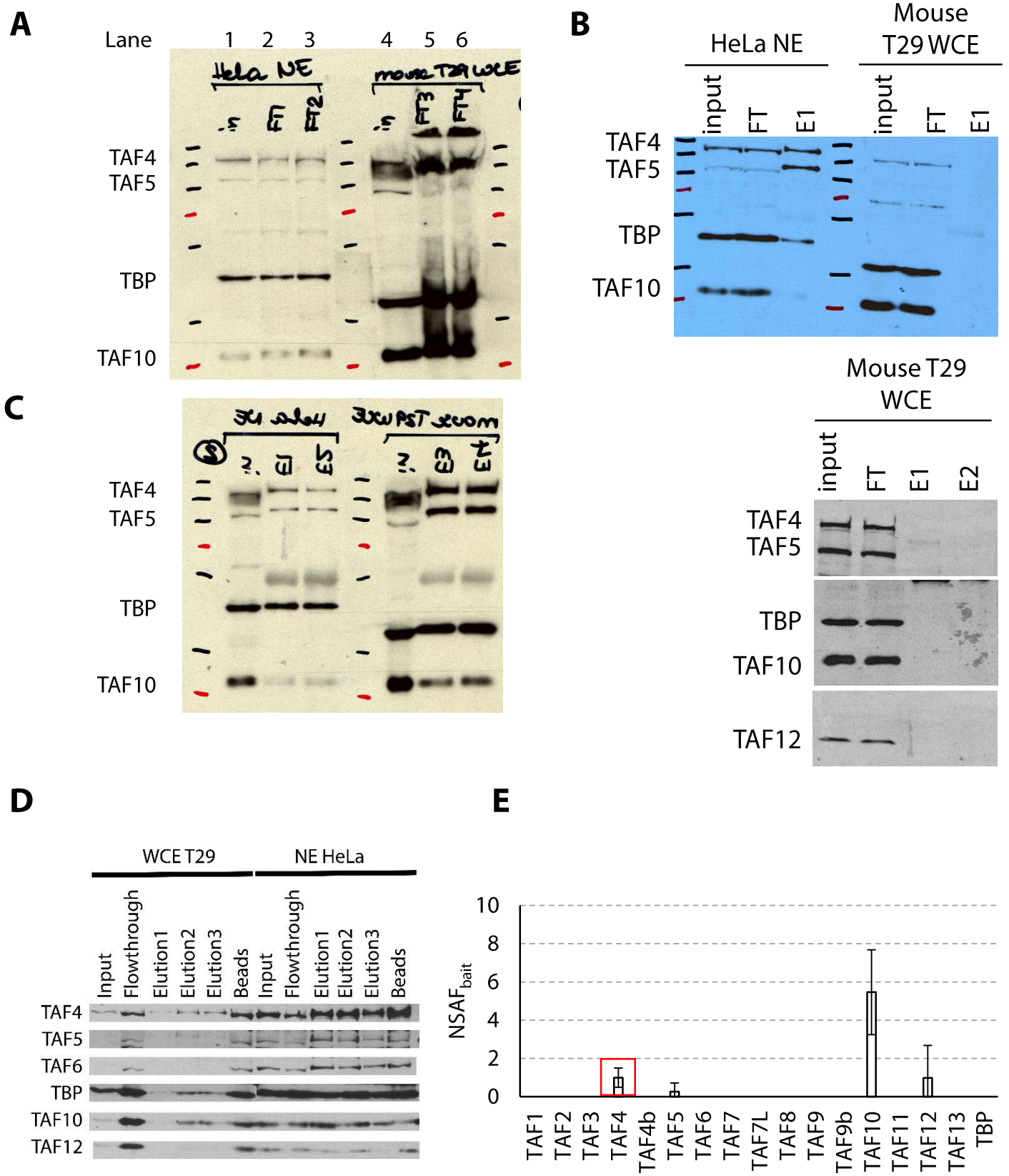
```
library(ggplot2)
#ouvrir tableau en sélectionnant une feuille en particulier
IPID <- readline(prompt="Which IP do you want to analyze? ")
tableauIPTAFX <- read.csv(IPID, header=TRUE)
#supprimer Keratin et Ig mais vérifier ce qui est supprimé
tableaucleanIPTAFX <- tableauIPTAFX[-grep("Keratin,",
tableauIPTAFX$Description),]
tableaucleanIPTAFX2 <- tableaucleanIPTAFX[-grep("^Ig ",
tableaucleanIPTAFX$Description),]
#générer tableaux de ce qui a été enlevé de l'analyse
refcheck <- tableauIPTAFX$Description
Keratincheck <- tableaucleanIPTAFX$Description
Igcheck <- tableaucleanIPTAFX2$Description
cleancheckkeratin <- setdiff(refcheck, Keratincheck)
cleancheckkeratinfinal <- as.data.frame(cleancheckkeratin)
cleancheckig <- setdiff(refcheck, Igcheck)
cleancheckigfinal <- as.data.frame(cleancheckig)
#remplacer NA par 0 pour éviter d'obtenir NA dans les calculs
tableaucleanIPTAFXwoNA <- tableaucleanIPTAFX2
tableaucleanIPTAFXwoNA[is.na(tableaucleanIPTAFXwoNA)] <- 0
#calcul moyenne MOCK
x <-
(tableaucleanIPTAFXwoNA[c('X..PSM.D2','X..PSM.E2','X..PSM.F2')
])
tableaucleanIPTAFXwoNA$meanMOCK <- rowMeans(x, na.rm = TRUE)
#calcul SAF triplicats
tableaucleanIPTAFXwoNA$SAF1 <- apply
(tableaucleanIPTAFXwoNA[,c('X..AAs','X..PSM.A2')],1,
function(x) {(x[2]/x[1])})
tableaucleanIPTAFXwoNA$SAF2 <- apply
(tableaucleanIPTAFXwoNA[,c('X..AAs','X..PSM.B2')],1,
function(x) {(x[2]/x[1])})
tableaucleanIPTAFXwoNA$SAF3 <- apply
(tableaucleanIPTAFXwoNA[,c('X..AAs','X..PSM.C2')],1,
function(x) {(x[2]/x[1])})
#calcul SAF mean MOCK
tableaucleanIPTAFXwoNA$SAFMOCK <- apply
(tableaucleanIPTAFXwoNA[,c('X..AAs','meanMOCK')],1,
function(x) {(x[2]/x[1])})
#calcul SAFs corrected
tableaucleanIPTAFXwoNA$SAF1corrected <-
apply(tableaucleanIPTAFXwoNA[c('SAFMOCK', 'SAF1')],1,
function(x) {x[2]-x[1]})
tableaucleanIPTAFXwoNA$SAF1corrected[tableaucleanIPTAFXwoNA$SA
F1corrected < 0] <- 0
tableaucleanIPTAFXwoNA$SAF2corrected <-
apply(tableaucleanIPTAFXwoNA[c('SAFMOCK', 'SAF2')],1,
function(x) {x[2]-x[1]})
```

```
tableaucleanIPTAFXwoNA$SAF2corrected[tableaucleanIPTAFXwoNA$SA
F2corrected < 0] <- 0
tableaucleanIPTAFXwoNA$SAF3corrected <-
apply(tableaucleanIPTAFXwoNA[c('SAFMOCK', 'SAF3')],1,
function(x) {x[2]-x[1]})
tableaucleanIPTAFXwoNA$SAF3corrected[tableaucleanIPTAFXwoNA$SA
F3corrected < 0] <- 0
#calculs somme SAFs corrected
tableaucleanIPTAFXwoNA$somme1 <- apply(tableaucleanIPTAFXwoNA
[c('SAF1corrected')],2, function(x) {sum(x)})
tableaucleanIPTAFXwoNA$somme2 <- apply(tableaucleanIPTAFXwoNA
[c('SAF2corrected')],2, function(x) {sum(x)})
tableaucleanIPTAFXwoNA$somme3 <- apply(tableaucleanIPTAFXwoNA
[c('SAF3corrected')],2, function(x) {sum(x)})
#calcul NSAF à partir des SAFs corrected
tableaucleanIPTAFXwoNA$NSAF1 <- apply
(tableaucleanIPTAFXwoNA[,c('somme1','SAF1corrected')],1,
function(x) {(x[2]/x[1])*100})
tableaucleanIPTAFXwoNA$NSAF2 <- apply
(tableaucleanIPTAFXwoNA[,c('somme2','SAF2corrected')],1,
function(x) {(x[2]/x[1])*100})
tableaucleanIPTAFXwoNA$NSAF3 <- apply
(tableaucleanIPTAFXwoNA[,c('somme3','SAF3corrected')],1,
function(x) {(x[2]/x[1])*100})
#moyenne NSAF
y <- (tableaucleanIPTAFXwoNA[,c('NSAF1','NSAF2', 'NSAF3')])
tableaucleanIPTAFXwoNA$meanNSAF <- rowMeans(y, na.rm = TRUE)
#calculer SD des NSAFs
tableaucleanIPTAFXwoNA$SDNSAF <- apply(tableaucleanIPTAFXwoNA
[c('NSAF1','NSAF2','NSAF3')],1, function(x) {sd(x)})
#normalisation NSAF par rapport au NSAF bait (1)
row.names(tableaucleanIPTAFXwoNA)<-
tableaucleanIPTAFXwoNA$Accession
baitID <- readline(prompt="give the Accession # of the bait
protein ")
normalization <- tableaucleanIPTAFXwoNA[baitID,'NSAF1']
tableaucleanIPTAFXwoNA$NSAFbaitref1 <- normalization
normalization <- tableaucleanIPTAFXwoNA[baitID,'NSAF2']
tableaucleanIPTAFXwoNA$NSAFbaitref2 <- normalization
normalization <- tableaucleanIPTAFXwoNA[baitID,'NSAF3']
tableaucleanIPTAFXwoNA$NSAFbaitref3 <- normalization
#moyenne NSAFbaitref
z <- (tableaucleanIPTAFXwoNA[,c('NSAFbaitref1','NSAFbaitref2',
'NSAFbaitref3')])
tableaucleanIPTAFXwoNA$meanNSAFbaitref <- rowMeans(z, na.rm =
TRUE)
#normalisation NSAF par la moyenne des NSAFbaitref (2)
tableaucleanIPTAFXwoNA$NSAF1normbait <- apply
(tableaucleanIPTAFXwoNA[,c('meanNSAFbaitref','NSAF1')],1,
function(x) {(x[2]/x[1])})
```

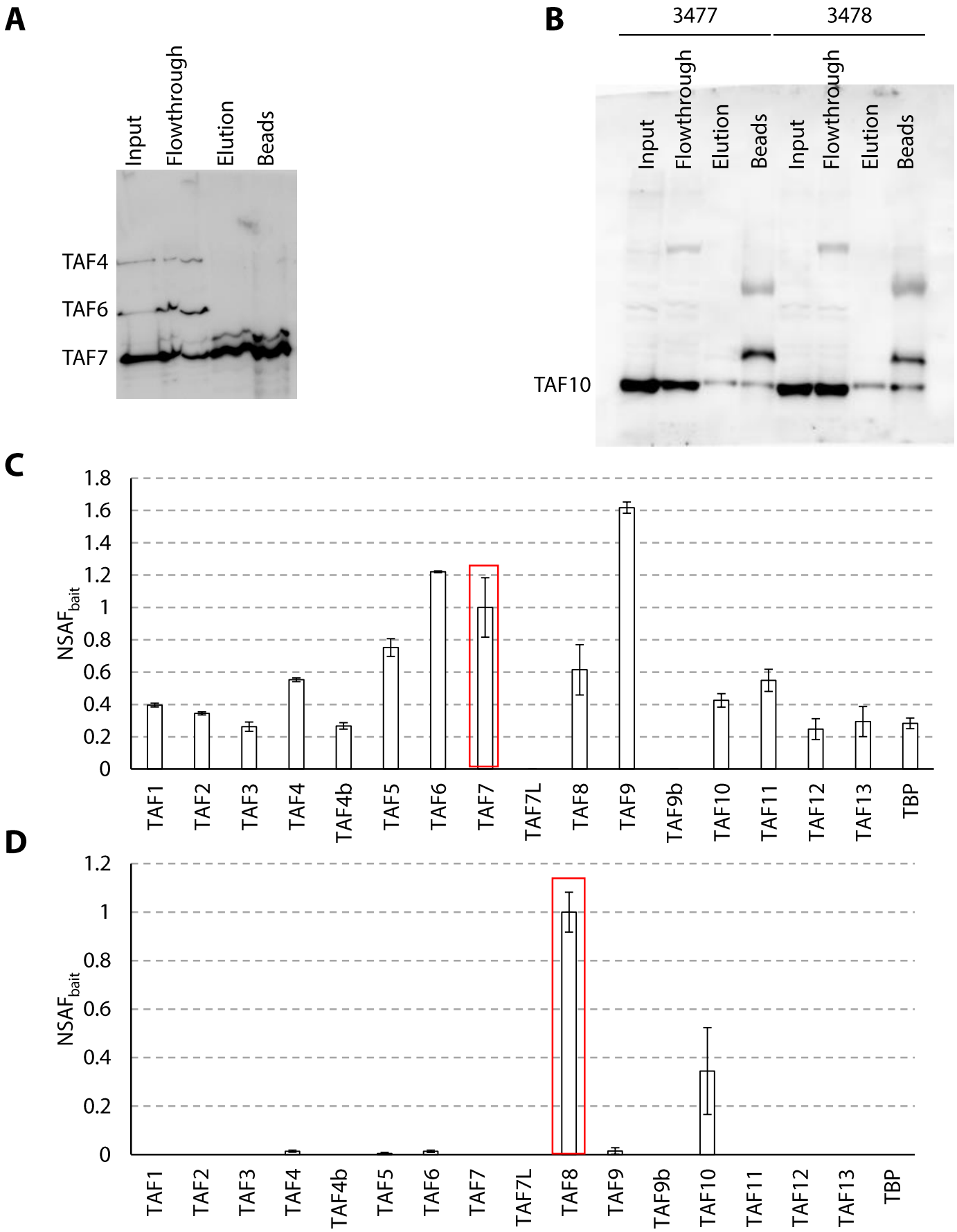
```
tableaucleanIPTAFXwoNA$NSAF2normbait <- apply
(tableaucleanIPTAFXwoNA[,c('meanNSAFbaitref','NSAF2')],1,
function(x) {(x[2]/x[1])})
tableaucleanIPTAFXwoNA$NSAF3normbait <- apply
(tableaucleanIPTAFXwoNA[,c('meanNSAFbaitref','NSAF3')],1,
function(x) {(x[2]/x[1])})
#moyenne NSAF bait
meanNSAFbait <-
(tableaucleanIPTAFXwoNA[,c('NSAF1normbait','NSAF2normbait',
'NSAF3normbait')])
tableaucleanIPTAFXwoNA$meanNSAFbait <- rowMeans(meanNSAFbait,
na.rm = TRUE)
#calculer SD des NSAFs bait
tableaucleanIPTAFXwoNA$SDNSAFbait <-
apply(tableaucleanIPTAFXwoNA
[c('NSAF1normbait','NSAF2normbait','NSAF3normbait')],1,
function(x) {sd(x)})
#fusionner tableau contenant la liste des sous-unités avec le
tableau d'analyse
sublist <- read.csv("donnees/sublist.csv", header=TRUE,
sep=",")
fusion <- merge.data.frame(tableaucleanIPTAFXwoNA, sublist,
by="Accession", all.y=TRUE)
#produire tableau final avec réarrangement des colonnes
tableauIPTAFXreorder <- fusion[,c('Accession','ID','meanNSAF',
'SDNSAF')]
#trier par ordre alphabétique
sortedtableau <-
tableauIPTAFXreorder[order(tableauIPTAFXreorder$ID),]
sortedtableau[is.na(sortedtableau)] <- 0
filename <- readline(prompt="Write the name of the file
(X.csv)")
write.csv(sortedtableau,filename, row.names=FALSE)
#générer tableau RNA Pol I et Pol III
sublist_TBP_associated <-
read.csv("donnees/TBP_associated_list.csv", header=TRUE,
sep=",")
fusion_TBP_associated <-
merge.data.frame(tableaucleanIPTAFXwoNA,
sublist_TBP_associated, by="Accession", all.y=TRUE)
#produire tableau final avec réarrangement des colonnes
tableauIPTAFXreorder_TBP_associated <-
fusion_TBP_associated[,c('Accession','ID','meanNSAFbait',
'SDNSAFbait')]
#trier par ordre alphabétique
sortedtableau_TBP_associated <-
tableauIPTAFXreorder_TBP_associated[order(tableauIPTAFXreorder
_TBP_associated$ID),]
sortedtableau_TBP_associated[is.na(sortedtableau_TBP_associate
d)] <- 0
```

```
filename <- readline(prompt="Write the name of the file for
TBP associated complexes (X.csv)")
write.csv(sortedtableau_TBP_associated,filename,
row.names=FALSE)
#générer tableau TFIID
sublistTFIID <- read.csv("donnees/TFIID_list.csv",
header=TRUE, sep=",")
fusion_TFIID <- merge.data.frame(tableaucleanIPTAFXwoNA,
sublistTFIID, by="Accession", all.y=TRUE)
#trier par ordre alphabétique
sublistTFIID <- read.csv("donnees/TFIID_list_reorder.csv",
header=TRUE, sep=",")
tableauIPTAFXreorder_TFIID_numbers <-
merge.data.frame(fusion_TFIID,sublistTFIID, by="ID", all.y =
TRUE)
sortedtableau_TFIID <-
tableauIPTAFXreorder_TFIID_numbers[order(tableauIPTAFXreorder_
TFIID_numbers$number),]
sortedtableau_TFIID[is.na(sortedtableau_TFIID)] <- 0
#produire tableau final avec réarrangement des colonnes
tableauIPTAFXreorder_TFIID_numbers_final <-
sortedtableau_TFIID[,c('Accession.x','ID','meanNSAFbait',
'SDNSAFbait')]
filename <- readline(prompt="Write the name of the file for
TFIID complex (X.csv)")
write.csv(tableauIPTAFXreorder_TFIID_numbers_final,filename,
row.names=FALSE)
#générer tableau TIP60
sublistTIP60 <- read.csv("donnees/TIP60_list.csv",
header=TRUE, sep=",")
fusion_TIP60 <- merge.data.frame(tableaucleanIPTAFXwoNA,
sublistTIP60, by="Accession", all.y=TRUE)
#produire tableau final avec réarrangement des colonnes
tableauIPTAFXreorder_TIP60 <-
fusion_TIP60[,c('Accession','ID','meanNSAFbait',
'SDNSAFbait')]
#trier par ordre alphabétique
sortedtableau_TIP60 <-
tableauIPTAFXreorder_TIP60[order(tableauIPTAFXreorder_TIP60$ID
),]
sortedtableau_TIP60[is.na(sortedtableau_TIP60)] <- 0
filename <- readline(prompt="Write the name of the file for
TIP60 complex (X.csv)")
write.csv(sortedtableau_TIP60,filename, row.names=FALSE)
#générer tableau ATAC
sublistATAC <- read.csv("donnees/ATAC_list.csv", header=TRUE,
sep=",")
fusion_ATAC <- merge.data.frame(tableaucleanIPTAFXwoNA,
sublistATAC, by="Accession", all.y=TRUE)
#produire tableau final avec réarrangement des colonnes
```

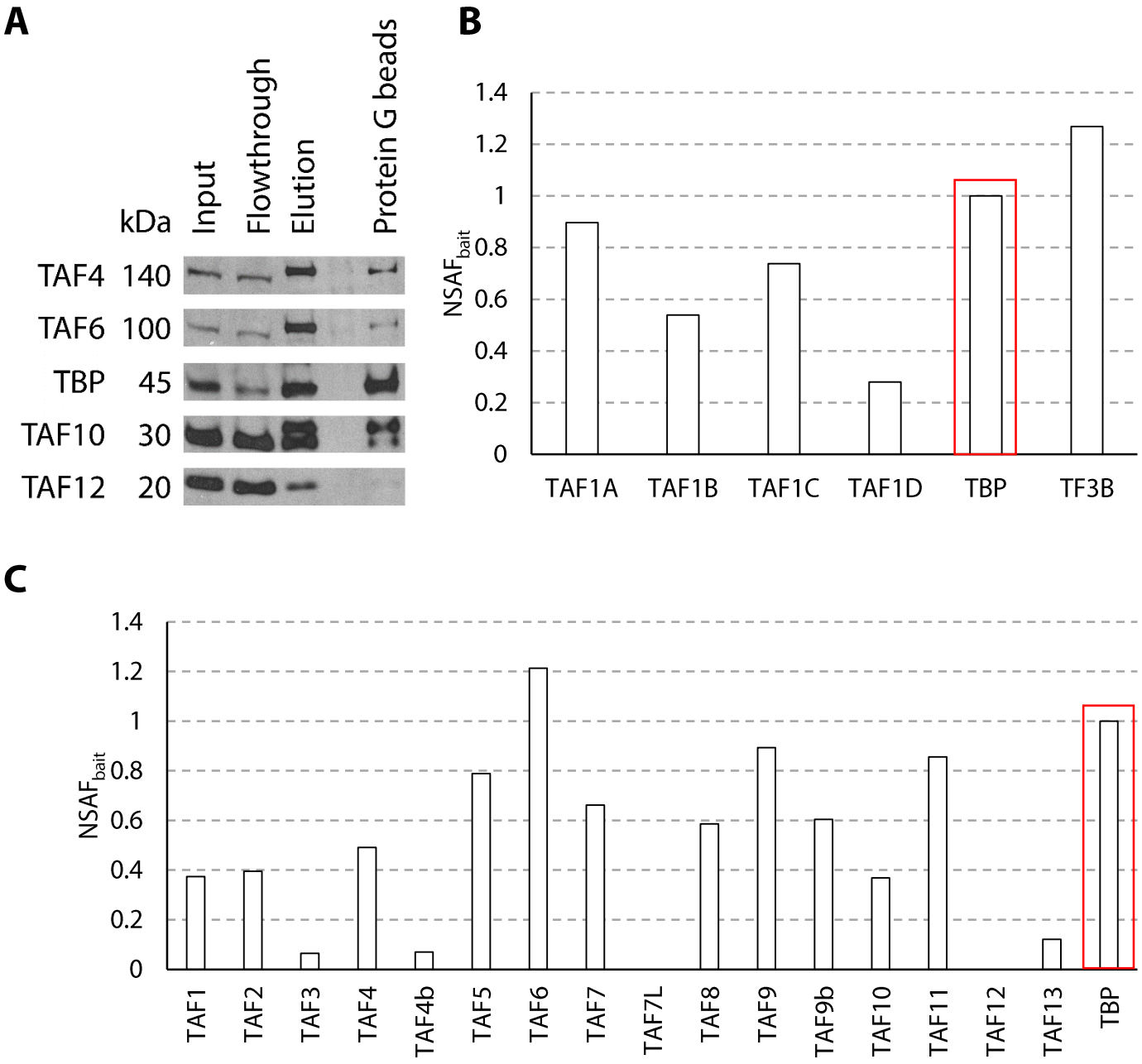
```
tableauIPTAFXreorder_ATAC <-
fusion_ATAC[,c('Accession','ID','meanNSAFbait', 'SDNSAFbait')]
#trier par ordre alphabétique
sortedtableau_ATAC <-
tableauIPTAFXreorder_ATAC[order(tableauIPTAFXreorder_ATAC$ID),
]
sortedtableau_ATAC[is.na(sortedtableau_ATAC)] <- 0
filename <- readline(prompt="Write the name of the file for
ATAC complex (X.csv)")
write.csv(sortedtableau_ATAC,filename, row.names=FALSE)
#générer tableau SAGA
sublistSAGA <- read.csv("donnees/SAGA_list.csv", header=TRUE,
sep=",")
fusion_SAGA <- merge.data.frame(tableaucleanIPTAFXwoNA,
sublistSAGA, by="Accession", all.y=TRUE)
#produire tableau final avec réarrangement des colonnes
tableauIPTAFXreorder_SAGA <-
fusion_SAGA[,c('Accession','ID','meanNSAFbait', 'SDNSAFbait')]
#trier par ordre alphabétique
sortedtableau_SAGA <-
tableauIPTAFXreorder_SAGA[order(tableauIPTAFXreorder_SAGA$ID),
]
sortedtableau_SAGA[is.na(sortedtableau_SAGA)] <- 0
filename <- readline(prompt="Write the name of the file for
SAGA complex (X.csv)")
write.csv(sortedtableau_SAGA,filename, row.names=FALSE)
```



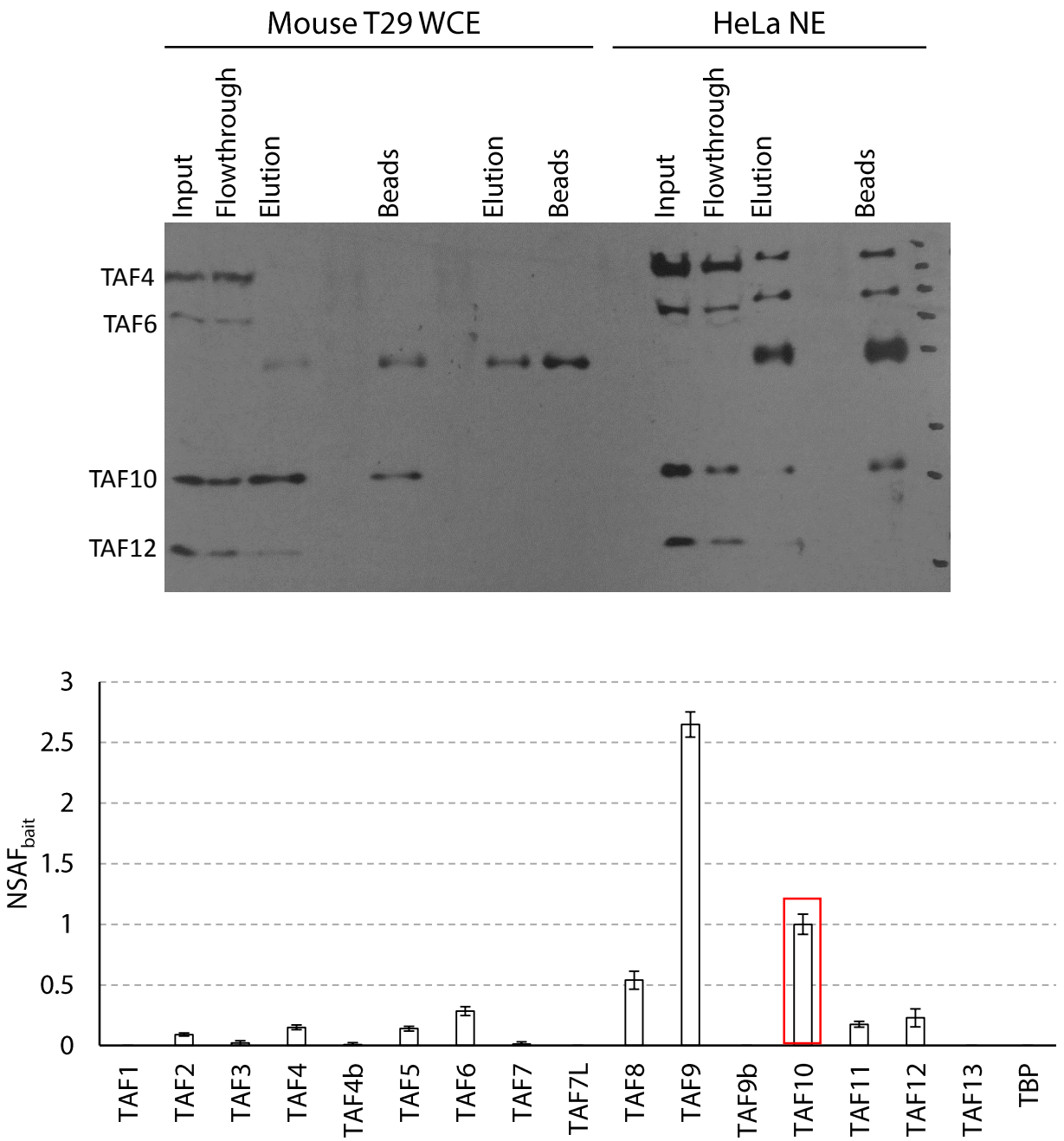

**Annexe II: Antibody validation for TAF1, TAF2, TAF3 and TAF4 immuno-precipitation**. (A) (B) (C) (D) Western-blot of (A) TAF1 (2439 fraction #7 (lane 2 & 5) and 2440 fraction #2 (lane 3 & 6) immuno-precipitation (30 sec exposure), (B) TAF2 immuno-precipitation, (C) TAF3 immuno-precipitation followed by two elutions (E1 & E2), (D) TAF4 (32TA2B9) immuno-precipitation followed by serial elution steps (E1, E2 & E3), (E) $NSAF_{bait}$ values for mouse TFIID complex subunits by mass spectrometry for TAF4 immuno-precipitation from HeLa nuclear extracts (positive control) or from mouse thymocyte whole cellular extracts. In;: input, FT: flowthrough, E: elution fractions. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3



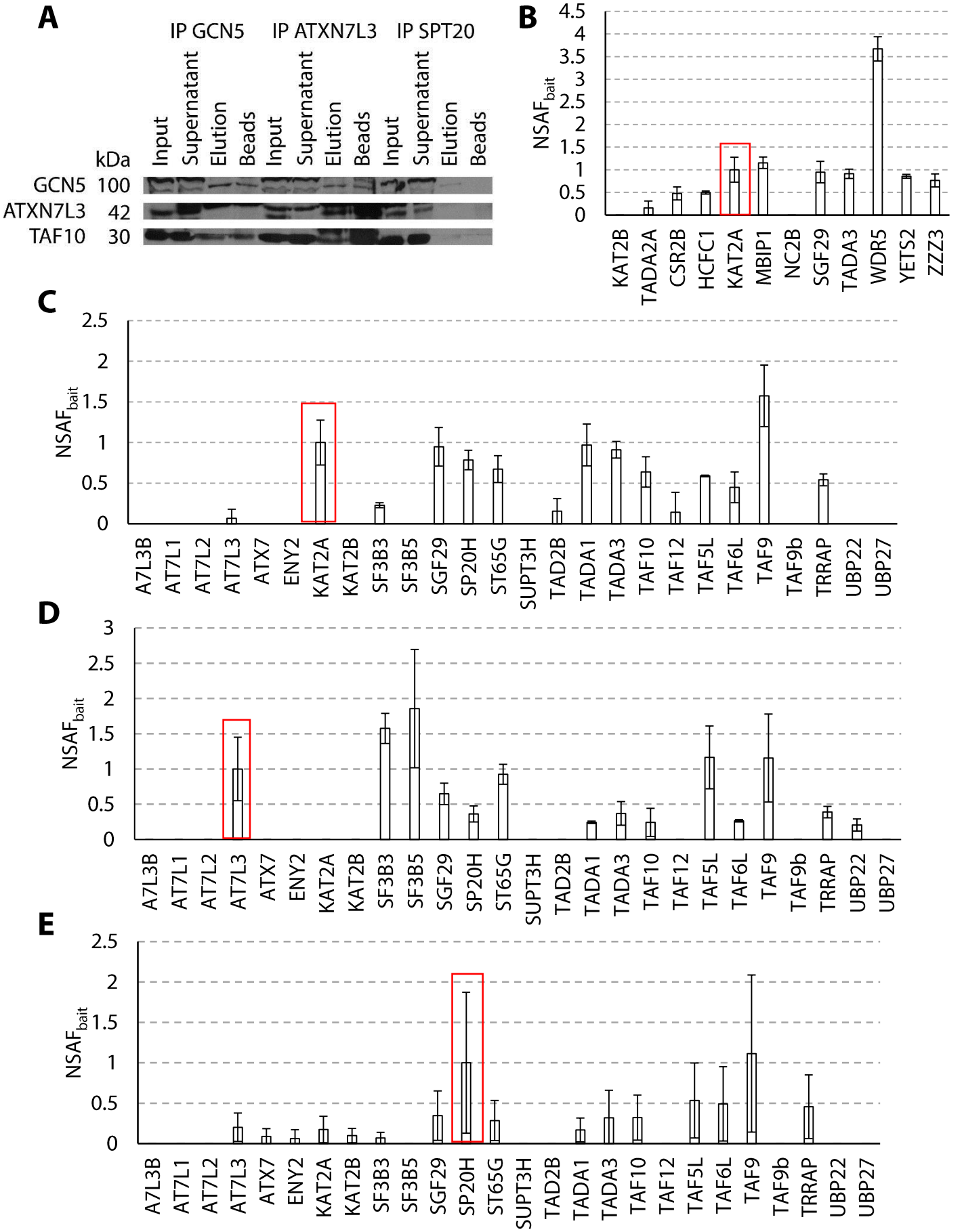

**Annexe III: Antibody validation for TAF7 and TAF8 immuno-precipitations**. (A) (B) Western-blot of (A) TAF7 (3475 fraction #2) immuno-precipitation from mouse thymocyte whole cellular extracts and (B) TAF8 (3477 and 3478) immuno-precipitation from mouse whole embryo between E8.5-E10.5. (C) (D) $NSAF_{bait}$ values for TFIID subunits from (C) TAF7 immuno-precipitation from mouse thymocyte whole cellular extracts and (D) TAF8 (3477) immuno-precipitation from mouse whole embryo cellular extracts. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3



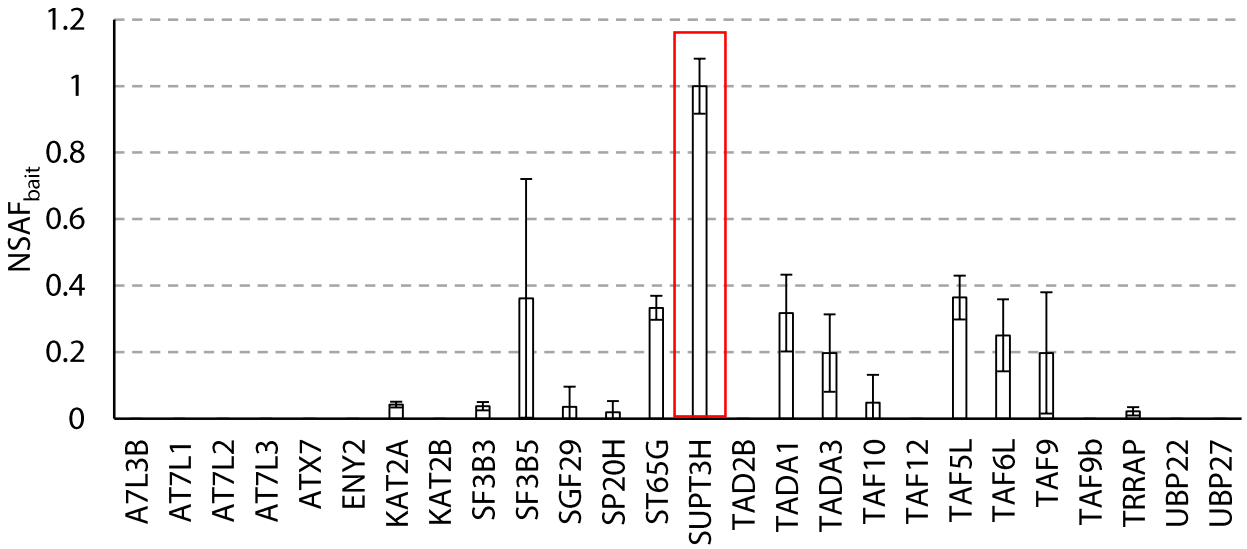

**Annexe IV: Antibody validation for TBP immuno-precipitation**. (A) Western-blot of TBP immuno-precipitations from mouse thymocyte whole cellular extracts. (B) (C) $NSAF_{bait}$ values for (B) SL1 subunits complex (TAF1A, TAF1B, TAF1C, TAF1D) and TF3B-TBP complex and (C) TFIID subunits of TBP immuno-precipitation from mouse thymocyte whole cellular extracts. The red rectangle indicate the bait. *n*=1



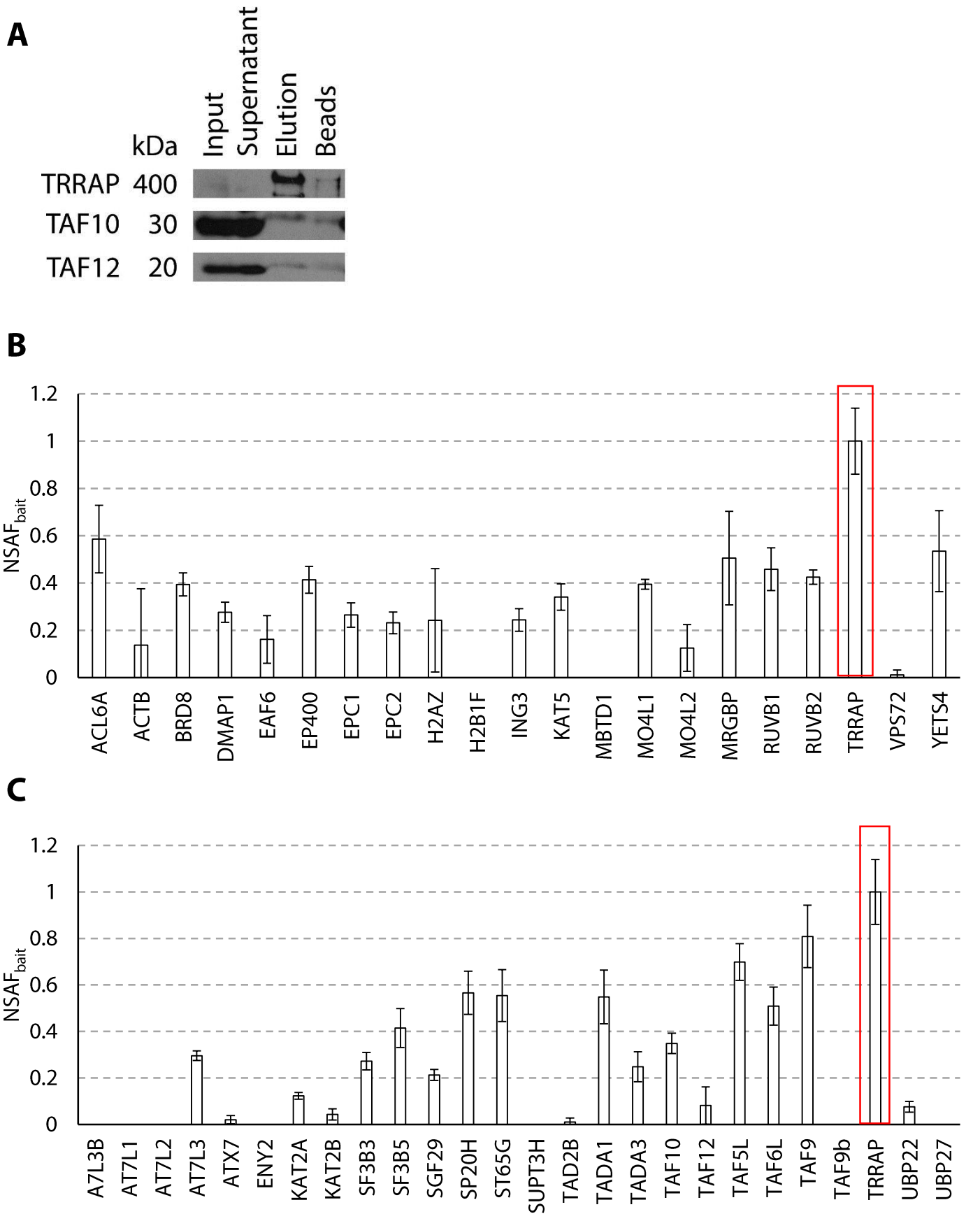

**Annexe V: Antibody validation for TAF10 immuno-precipitation**. (A) Western-blot of TAF10 immuno-precipitation from mouse thymocyte whole cellular extracts and HeLa nuclear extracts. (B) $NSAF_{bait}$ values for TFIID subunits from TAF10 immuno-precipitation from mouse thymocyte whole cellular extracts. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3



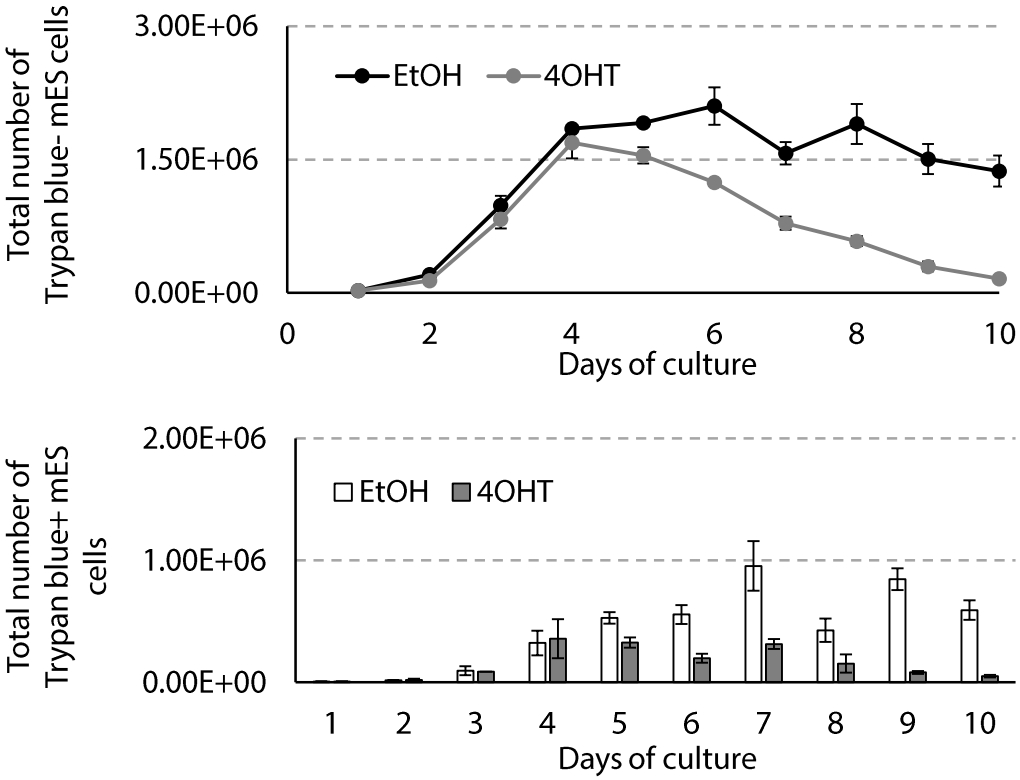

**Annexe VI: Antibody validation for GCN5, ATXN7L3, and SUPT20H immuno-precipitations**. (A) Western-blot of GCN5 (mouse KAT2A), ATXN7L3, SUPT20H immuno-precipitations (30 sec exposure). (B) (C) (D) (E) $NSAF_{bait}$ values for (B) ATAC complex subunits from GCN5 immuno-precipitation, (C) (D) (E) SAGA complex subunits from (C) GCN5, (D) ATXN7L3 and (E) SUPT20H immuno-precipitations from mouse whole cell extract from thymocytes (T29). The red rectangle indicate the bait. Error bars indicate s.d. *n*=3



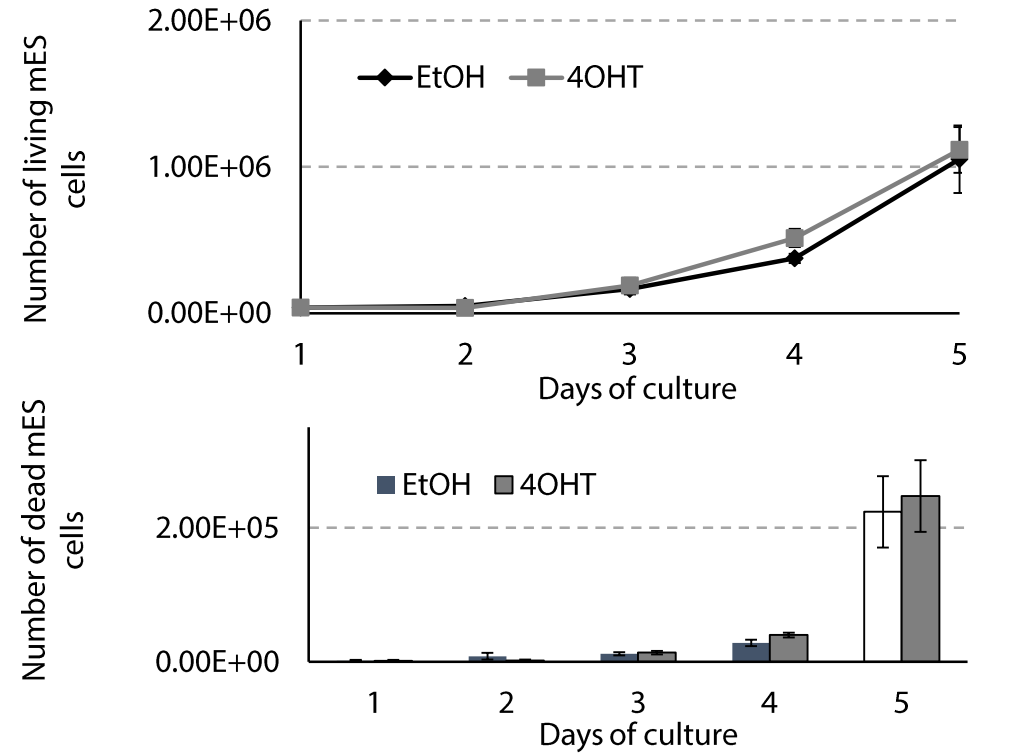

**Annexe VII: Antibody validation for SUPT3H immuno-precipitation**. $NSAF_{bait}$ values for SAGA complex subunits from SUPT3H immuno-precipitation from mouse whole cell extract from thymocytes (T29. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3



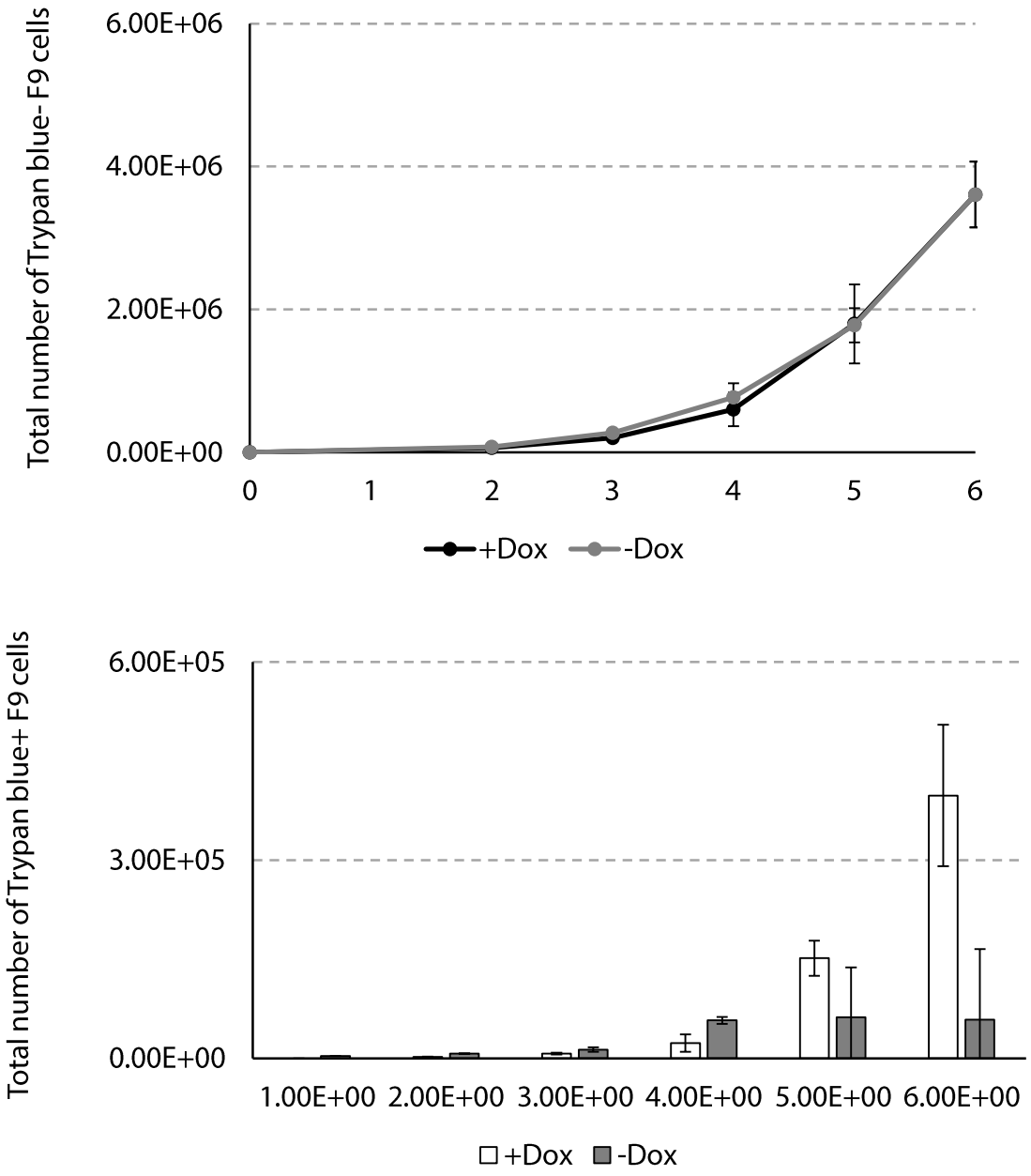

**Annexe VIII: Antibody validation for TRRAP immuno-precipitation**. (A) Western-blot of TRRAP immuno-precipitation. (B) (C) $NSAF_{bait}$ values for (B) TIP60/NuA4 complex subunits, (C) SAGA complex subunits from TRRAP immuno-precipitation from mouse whole cell extract from thymocytes (T29. The red rectangle indicate the bait. Error bars indicate s.d. *n*=3




**Annexe IX: TAF10 is required for normal mES cell growth.** *R26*$^{CreERT2/R}$*; Taf10*$^{flox/flox}$ mES cells were used. (A) Growth curve of mES cells treated with EtOH (vehicle) or 100 nM 4OHT at day 1, plotted*Tg(Mesogenin-YFP* as the number of Trypan- (living) mES cells. (B) Number of Trypan blue positive mES cells treated with EtOH (vehicle) or 100 nM 4OHT at day 1. Control and mutant conditions are indicated in white and gray, respectively. Error bars indicate s.d. *n*=3




**Annexe X: 4OHT treatment does not affect controls mES cellular growth and viability.** E14 mES cells haboring *Pax7-GFP* allele were used as a control for testing the effect of 4OHT. A) Growth curve of mES cells treated with EtOH (vehicle) or 100 nM 4OHT at day 1, plotted as the number of living mES cells. (B) Number of dead mES cells treated with EtOH (vehicle) or 100 nM 4OHT at day 1. Control and mutant conditions are indicated in white and gray, respectively. Error bars indicate s.d. *n*=3




**Annexe XI: Cellular growth is not affected by doxycycline in F9 wild-type cells.** (A) Growth curve of F9 wild-type cells treated with or without doxycyline at day 1, and plotted as the number of Trypan blue negative (living) F9 cells. (B) Number of Trypan blue positive (dead) wild-type F9 cells treated with or without doxycyline at day 1. Error bars indicate s.d. *n*=3

**Paul Bardot**

# Analyse de la composition et de la fonction de la machinerie basale de transcription au cours du développement et de la différenciation


## Résumé

TFIID et SAGA sont deux complexes importants pour la transcription, contenant la sous-unité TAF10. Nous avons analysé leur composition dans l'embryon murin et différents contextes cellulaires par immuno-précipitation et spectrométrie de masse. Les sous unités des complexes TFIID et SAGA ont été détectées en proportion différente selon le type cellulaire. Par filtration sur gel, des sous-complexes de TFIID ont aussi été détectés. En absence de TAF10, l'assemblage de TFIID et SAGA est fortement affecté mais la formation des somites n'est pas initialement affecté ni l'expression globale des gènes. L'analyse des niveaux d'ARN totaux et naissants dans les cellules ES murines suggèrent que TFIID et SAGA sont requis globalement pour l'initiation de la transcription, mais que la diminution de la synthèse des ARNm est compensée.


Mots-clés : TFIID ; SAGA ; Transcription ; Développement


## Résumé en anglais

TFIID and SAGA are two multi-subunit complexes which play important roles in transcription and that contain the TAF10 subunit. By immunoprecipitation followed by mass spectrometry, we analyzed the composition of TFIID and SAGA complexes in the embryo as well as in different cellular contexts. TFIID and SAGA complexes subunits were detected in different proportions depending on the cellular context. By gel filtration, we also detected distinct TFIID sub-complexes. In the absence of TAF10, TFIID and SAGA assembly is severely impaired but neither early somitogenesis nor global gene expression is affected. Steady-state and newly-transcribed mRNA analyses in mES cells suggest that TFIID and SAGA are generally required for transcription initiation. However, the decrease of mRNA synthesis is compensated.


Keywords: TFIID; SAGA; Transcription; Development