



**HAL**  
open science

## Analyses structure fonction du module de déubiquitination du complexe SAGA

Jacques Bonnet

► **To cite this version:**

Jacques Bonnet. Analyses structure fonction du module de déubiquitination du complexe SAGA. Biochimie, Biologie Moléculaire. Université de Strasbourg, 2012. Français. NNT : 2012STRAJ114 . tel-00870969

**HAL Id: tel-00870969**

**<https://theses.hal.science/tel-00870969>**

Submitted on 8 Oct 2013

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

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

*École Doctorale des Sciences de la Vie et de la Santé*

IGBMC - CNRS UMR 7104 - Inserm U 964

**THÈSE** présentée par :

**Jacques BONNET**

soutenue le : 19 mars 2012

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

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

**Analyses structure fonction du module  
de déubiquitination du complexe SAGA**

**THÈSE dirigée par :**

**M. DEVYS Didier**

MCU-PH, Université de Strasbourg

**RAPPORTEURS :**

**Mme STUTZ Françoise**

**M. KHOCHBIN Saadi**

Directeur de recherche, Université de Genève

Directeur de recherche, Institut Albert Bonniot, Grenoble

---

**AUTRES MEMBRES DU JURY :**

**Mme TORRES-PADILLA Maria-Elena**

Chargée de recherche, IGBMC

*« Que cherchez vous ? »*

Jn 1, 38

*« Tout ce qui est peut ne pas être. »*

Hume,  
Enquête sur l'entendement humain

*« Quand tu étais sous le figuier, je t'ai vu. »*

Jn 1, 48

A mes parents et à Marie-Paule,  
A toutes les personnes qui m'ont aidé à grandir...

## Remerciements

Je tiens à remercier le docteur Laszlo Tora de m'avoir permis de réaliser ma thèse au sein de son équipe ainsi que le docteur Didier Devys de m'avoir encadré durant la préparation de cette thèse.

Merci aux docteurs Bruno Kieffer et Marc Timmers pour les collaborations que nous avons mené ensemble.

Je voudrais également dire un immense merci à tous les membres de l'équipe Tora que j'ai côtoyé durant ces quatre ans et demi de thèse. Merci à tous les collègues avec lesquels j'ai pu réaliser des expériences ou avec lesquels j'ai pu discuter d'idées relatives à la science... ou à la vie d'une manière plus générale.

Je voudrais enfin remercier les docteurs Françoise Stutz, Maria-Elena Torres-Padilla et Saadi Khochbin d'avoir bien voulu évaluer ce travail.

## Avant propos

Chez les eucaryotes, parmi les acteurs impliqués dans la régulation de l'expression des gènes, il existe des co-activateurs transcriptionnels qui sont recrutés au niveau de promoteurs particuliers par des activateurs reconnaissant des séquences spécifiques d'ADN et dont la fonction est de faciliter l'initiation de la transcription par l'ARN Polymérase II. Un des co-activateurs qui a été le plus largement étudié est le complexe SAGA qui est conservé au cours de l'évolution de la levure à l'homme.

Chez la levure *S. cerevisiae*, deux modules fonctionnels du complexe SAGA sont directement impliqués dans la régulation des mécanismes de pré-initiation de la transcription. L'un de ces modules contient Gcn5, une enzyme qui possède une activité d'acétylation des histones H3 et qui participe à la décompaction de la chromatine au niveau des promoteurs. L'autre module peut interagir avec la protéine TBP et réguler le recrutement de celle-ci au niveau de certains promoteurs. Chez les eucaryotes supérieurs, seule l'activité d'acétylation des histones portée par GCN5 semble conservée et est impliquée dans la régulation de l'expression de gènes.

Le complexe SAGA possède une seconde activité enzymatique de déubiquitination (DUB) de l'histone H2B qui est portée par Ubp8 ou USP22, respectivement chez la levure et chez l'homme. Cette déubiquitinase interagit de façon très forte avec trois autres sous-unités de SAGA pour former un module structural et fonctionnel appelé module de déubiquitination (DUBm). Cependant, dans les conditions normales de croissance cellulaire, il semble que la fonction principale de l'activité DUB de SAGA ne joue pas un rôle majeur dans la régulation de l'expression des gènes.

L'activité DUB de SAGA permet par contre de maintenir un niveau relativement bas de H2B ubiquitiné dans la cellule. Même si le rôle exact joué par cette modification post-traductionnelle de l'histone H2B n'est pas encore clairement compris, d'importantes avancées permettant de préciser la fonction de cette marque ont été réalisées lors de ces dernières années. Par exemple, il a été clairement démontré que l'ubiquitination de H2B était un processus co-transcriptionnel et que cette marque était déposée sur le corps des gènes transcrits. Chez la levure, il est aussi connu que l'ubiquitination de H2B est requise pour permettre la di et la tri méthylation de la lysine 4 de l'histone H3, une marque retrouvée sur les promoteurs des gènes transcrits.

D'autre part, l'ubiquitination de H2B est également nécessaire pour permettre la di et la tri méthylation de la lysine 79 de l'histone H3. Cette dernière marque est importante pour maintenir le « silencing » des télomères chez la levure *S. cerevisiae*.

Au moment où j'ai débuté ma thèse, nous avons observé que la protéine recombinante USP22 seule ne présentait pas d'activité de déubiquitination alors que le complexe SAGA purifié à partir de cellules HeLa pouvait déubiquitiner H2B. De façon intéressante, il a été montré que l'activité catalytique de plusieurs autres enzymes de la famille de USP22 (la famille des ubiquitine spécifique protéases) était régulée de façon allostérique grâce à des interactions protéiques réalisées par ces enzymes.

Mon travail de thèse a donc été centré sur l'étude des mécanismes de régulation de l'activité DUB de SAGA. Grâce à des approches structure fonction, nous avons notamment montré que la formation du DUBm humain est nécessaire pour stimuler l'activité DUB de USP22. D'autre part, nous avons montré qu'un domaine contenant un doigt de zinc atypique et situé au sein d'une des sous-unités du module de déubiquitination pouvait interagir avec un nucléosome. De plus, nous avons montré que cette interaction participe à la régulation de l'activité DUB de SAGA en permettant sans doute de positionner le DUBm de façon optimale par rapport à son substrat.

Pour continuer ce travail, nous essayons maintenant de déterminer l'ensemble des régions du génome où SAGA exerce son activité DUB afin de mieux caractériser la fonction du DUBm.

# Table des Matières

<b>Remerciements .....</b>	<b>2</b>
<b>Avant propos.....</b>	<b>3</b>
<b>Table des Matières .....</b>	<b>5</b>
<b>Liste des Figures et Tableaux.....</b>	<b>8</b>
<b>Partie 1 : mécanismes d'initiations de la transcription .....</b>	<b>10</b>
I. Le promoteur de base.....	10
II. Les facteurs généraux de la transcription et la formation du complexe de préinitiation ..	11
III. Généralités sur la régulation de la transcription.....	12
IV. L'organisation de la chromatine et la notion de code des histones.....	13
V. Les coactivateurs de la transcription .....	15
<b>Partie 2 : SAGA, un complexe co-activateur de la transcription 17</b>	
<b>A. Généralités sur le complexe SAGA.....</b>	<b>17</b>
I. Un complexe multifonctionnel .....	17
II. Rôle dans la régulation de l'expression des gènes .....	21
1. Chez la levure .....	21
2. Rôle dans le développement des métazoaires .....	22
III. Recrutement aux promoteurs de gènes cibles .....	23
1. Localisation et mode d'action.....	23
2. Recrutement sur la chromatine.....	24
<b>B. Identification de ses sous-unités et caractérisation du complexe .....</b>	<b>26</b>
I. Cribles transcriptionnels <i>SPT</i> et des <i>ADA</i> .....	26
1. Gènes <i>SPT</i> de la classe TBP .....	27
2. Le crible ADA : identification d'adaptateurs.....	30
3. Chevauchement des cribles <i>SPT</i> et <i>ADA</i> .....	35
4. Identification du complexe SAGA .....	35
5. Cribles génétiques et modularité fonctionnelle .....	36
II. Caractérisation biochimique chez la levure .....	38
1. Analyse protéomique.....	38
2. Sus1 identifiée par un crible génétique .....	41
3. Identification du module de déubiquitination .....	41
III. Identification et caractérisation de SAGA chez l'Homme .....	44
1. Purification de complexes homologues .....	44
2. Identification de SGF29 et de SPT20 .....	48
3. Identification de ATXN7 .....	49
4. Identification d'un module de déubiquitination .....	51
IV. Evolution du complexe SAGA.....	53
1. Conservation au cours de l'évolution.....	53
2. Lien avec TBP chez les métazoaires ? .....	53
3. Duplication de gènes codant des sous-unités de SAGA.....	54
4. Autres complexes contenant des sous-unités de SAGA .....	55

<b>C. Modularité structurale et fonctions de SAGA .....</b>	<b>57</b>
I. Structure tridimensionnelle de SAGA.....	57
1. Structure modulaire .....	57
2. Localisation des sous-unités .....	57
3. Cœur structural de SAGA.....	58
4. Localisation des fonctions co-activatrices de SAGA .....	60
II. Analyse protéomique et architecture de SAGA.....	61
1. Approche expérimentale .....	61
2. Les 4 modules structuraux .....	61
3. Rôle structural et fonctionnel de Spt20 .....	61
4. Organisation du cœur structural .....	63
5. Vers une meilleure définition architecturale de SAGA.....	64
III. Analyse des transcriptomes et modules fonctionnels : .....	65
1. Chez <i>S. cerevisiae</i> .....	65
2. Chez <i>S. pombe</i> .....	66
IV. Analyses structurales et fonctionnelles de modules de SAGA.....	68
1. Le HAT module .....	68
2. Le DUB module .....	70
3. Tra1 et organisation structurale de SAGA .....	71
V. Données additionnelles pour définir la structure du complexe SAGA .....	73
1. Modifications post-traductionnelles des sous-unités de SAGA .....	73
2. Stœchiométrie des différentes sous-unités.....	74
VII Conclusions.....	76
 <b>Partie 3 : Les Déubiquitinases de la famille USP .....</b>	 <b>77</b>
<b>A. Le monde de l'ubiquitine.....</b>	<b>77</b>
I. Machinerie de mise en place .....	77
II. Une multitude de fonctions exercées par l'ubiquitine.....	80
<b>B. Mécanismes de régulation des USP.....</b>	<b>82</b>
I. Cinq familles de déubiquitinases.....	82
II. Caractéristiques générales des USP.....	83
1. Diversification au cours de l'évolution.....	83
2. Association à des complexes protéiques .....	83
3. Présence de domaines additionnels .....	83
4. Description structurale du domaine USP .....	84
5. Equilibre entre formes actives et inactives .....	88
6. Divers modes de régulations .....	89
III. Différentes régulations par les domaines Ubl .....	89
1. Régulation de Ubp6/USP14 par le protéasome .....	90
2. Activation allostérique de USP7 par ses domaines Ubl.....	93
3. Inhibition de USP4 par un domaine Ubl .....	97
IV. Activations allostériques réalisées par des ZnF-UBP .....	99
1. Régulation de Ubp8 par le DUBm de SAGA .....	99
2. USP5 reconnaît spécifiquement des chaînes de polyubiquitine libres .....	102
 <b>Partie 4 : l'ubiquitination de H2B .....</b>	 <b>104</b>
 <b>Partie 5 : Résultats .....</b>	 <b>105</b>
I. Analyses structure fonction des domaines SCA7 .....	105
II. Régulation de l'activité de USP22 par les sous-unités du DUBm .....	109



<b>Partie 6 : Discussion et Perspectives.....</b>	<b>112</b>
<b>A. Etude genome-wide de l'activité DUB de SAGA.....</b>	<b>112</b>
<b>B. Analyses structure fonction du DUBm.....</b>	<b>119</b>
I. Analyses structure fonction des domaines SCA7 .....	119
II. Régulation enzymatique de l'activité DUB de SAGA par les sous-unités du DUBm.....	122
<b>Références .....</b>	<b>126</b>

## Liste des Figures et Tableaux

**Figure 1.** Schéma la formation séquentielle du complexe de préinitiation de la transcription.

**Figure 2.** Structure atomique d'un nucléosome résolue par cristallographie.

**Figure 3.** Représentation schématique des différents niveaux de compaction de l'ADN.

**Tableau 1.** Composition du complexe SAGA dans différents organismes modèles.

**Figure 4.** Représentation schématique du complexe SAGA chez la levure *S. cerevisiae* et chez l'homme.

**Figure 5.** Le crible génétique SPT.

**Figure 6.** Démonstration *in vitro* de l'existence d'adaptateurs.

**Figure 7.** Modularité du complexe SAGA définie par les cribles génétiques.

**Figure 8.** Identification de sous-unités du complexe SAGA par des approches de protéomique.

**Figure 9.** Formation du module de déubiquitination de SAGA chez la levure *S. cerevisiae*.

**Figure 10.** GCN5 et PCAF sont des protéines homologues.

**Figure 11.** Structure tridimensionnelle de SAGA obtenue par microscopie électronique.

**Figure 12.** Modularité du complexe SAGA définie par une approche de protéomique.

**Figure 13.** Analyse des transcriptomes et définition de modules fonctionnels.

**Figure 14.** Structure de domaines du HATm de SAGA.

**Figure 15.** Organisation en domaines des protéines Tra1 et TRRAP.

**Figure 16.** Proximité spatiale de certaines sous-unités de SAGA avec Gcn5.

**Figure 17.** La molécule d'ubiquitine.

**Figure 18.** La machinerie d'ubiquitination.

**Figure 19.** Conformation adoptée par les différents types de chaînes de polyubiquitine.

**Figure 20.** Structure du domaine catalytique de différentes USP.

**Figure 21.** Vue détaillée du site catalytique de USP7.

**Figure 22.** Interaction de l'ubiquitine avec USP7.

**Figure 23.** Liaison de l'ubiquitine par deux endo-DUB.

**Figure 24.** Localisation de Ubp6/USP14 au sein de la particule 19S du protéasome.

**Figure 25.** Mécanisme d'activation de USP7 par ses domaines Ubl.

**Figure 26.** Mécanisme d'inhibition de USP4 par son domaine Ubl interne.

**Figure 27.** Structure du module de déubiquitination de SAGA chez *S. cerevisiae*.

**Figure 28.** Modèle de la liaison du DUBm avec un nucléosome ubiquitiné sur H2B.

**Figure 29.** Interaction du domaine ZnF-UBP de USP5 avec l'ubiquitine.

**Figure 30.** Domaines de liaison à l'ubiquitine de USP5.

**Figure 31.** La famille des protéines ATXN7.

**Figure 32.** SAGA exerce son activité DUB sur la région transcrite de gènes exprimés.

# **Introduction**

# **Partie 1 : mécanismes d'initiations de la transcription**

Avant d'étudier en détail le complexe co-activateur de la transcription SAGA, nous allons rappeler les évènements et les partenaires importants pour la transcription par l'ARN Polymérase II (Pol II) chez les eucaryotes. Nous nous limiterons à l'étude du processus précédant la synthèse de la première liaison phosphodiester ou étape d'initiation de la transcription. La Pol II va transcrire en ARN les gènes codant pour un ARN messenger, c'est-à-dire environ 20000 gènes chez l'homme, ainsi que certains ARN non codants, mais nous ne nous intéresserons qu'à la transcription des ARN messagers.

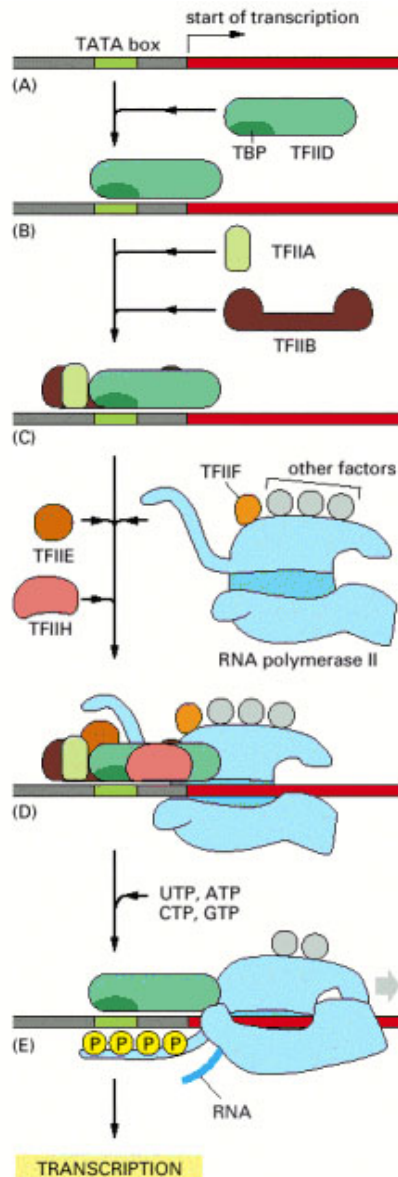
## **I. Le promoteur de base**

Chez les eucaryotes, devant chaque séquence codante, on retrouve un promoteur, appelé promoteur de base. Contrairement à ce qui se passe chez les bactéries, la Pol II n'est pas capable d'interagir directement avec ce promoteur de base. D'autres facteurs protéiques sont nécessaires pour recruter cette enzyme (Venters & Pugh, 2009).

Les études réalisées sur les promoteurs de base eucaryotes ont mis en évidence sept éléments conservés au niveau de ces promoteurs. Ces séquences d'ADN vont être reconnues par diverses protéines ce qui va permettre la formation d'un complexe fonctionnel de préinitiation de la transcription contenant la Pol II. Parmi ces éléments conservés, on retrouve la TATA Box, reconnue par la TATA Binding Protein (TBP) présente dans le complexe TFIID (Transcription Factor II D), un facteur général de la transcription. Il est important de noter que l'on ne trouve pas chacun des sept éléments au niveau de chaque promoteur de base. Au contraire, au niveau d'un promoteur de base, on retrouve différentes combinaisons d'éléments promoteurs. Cette diversité doit jouer un rôle sur l'efficacité de formation du complexe de préinitiation, sur le recrutement de différents facteurs transcriptionnels et donc sur la régulation de l'expression des gènes (Thomas & Chiang, 2006).

## II. Les facteurs généraux de la transcription et la formation du complexe de préinitiation

Le promoteur de base va servir de plateforme pour l'assemblage du complexe de



**Figure 1. Schéma la formation séquentielle du complexe de préinitiation de la transcription.**

Dans ce modèle, les facteurs généraux de la transcription se fixent séquentiellement pour former le complexe de préinitiation (PIC). Grâce à un changement de conformation le PIC passe d'un état fermé à un état ouvert et une phosphorylation de la queue C-terminale de la grande sous-unité de l'ARN polymérase II permet de commencer l'élongation.

préinitiation de la transcription. En plus de la Pol II, différents facteurs protéiques sont nécessaires pour la formation de ce complexe : les facteurs généraux de la transcription ou GTF. On en distingue 7 ; TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH. Ce sont des complexes multiprotéiques qui renferment de 1 à 15 protéines. Par exemple, TFIID renferme TBP ainsi que 14 TBP Associated Factors (TAF). L'association des GTF et le recrutement de la Pol II sur le promoteur de base va définir le site d'initiation ainsi que le sens de la transcription. La formation de ce complexe est suffisante pour permettre un niveau de transcription de base. Deux modèles, que nous allons brièvement présenter, ont été proposés pour décrire la mise en place du complexe de préinitiation de la transcription (Thomas & Chiang, 2006).

### L'assemblage séquentiel

Dans ce premier modèle, TFIID se fixe sur l'ADN au niveau du promoteur de base. Cette interaction est renforcée par la venue de TFIIA puis celle de TFIIB. La Pol II est ensuite recrutée dans le complexe avec TFIIF. Lorsque ces

facteurs forment un complexe stabilisé, TFIIE et enfin TFIIH peuvent à leur tour venir pour terminer la formation du complexe de préinitiation de la transcription (Figure 1) (Thomas & Chiang, 2006).

#### Le modèle holoenzyme

Ce modèle a vu le jour suite à la purification d'un complexe contenant l'ARN PolII, certains GTF ainsi que d'autres facteurs impliqués dans la régulation de la transcription comme le médiateur, un complexe co-activateur, et SWI/SNF, un complexe impliqué dans le remodelage de la chromatine. Dans ce cas, TFIID est toujours le facteur de reconnaissance et de liaison au promoteur de base. L'interaction est stabilisée par TFIIA et ce qui permet l'arrivée de l'holoenzyme préformée.

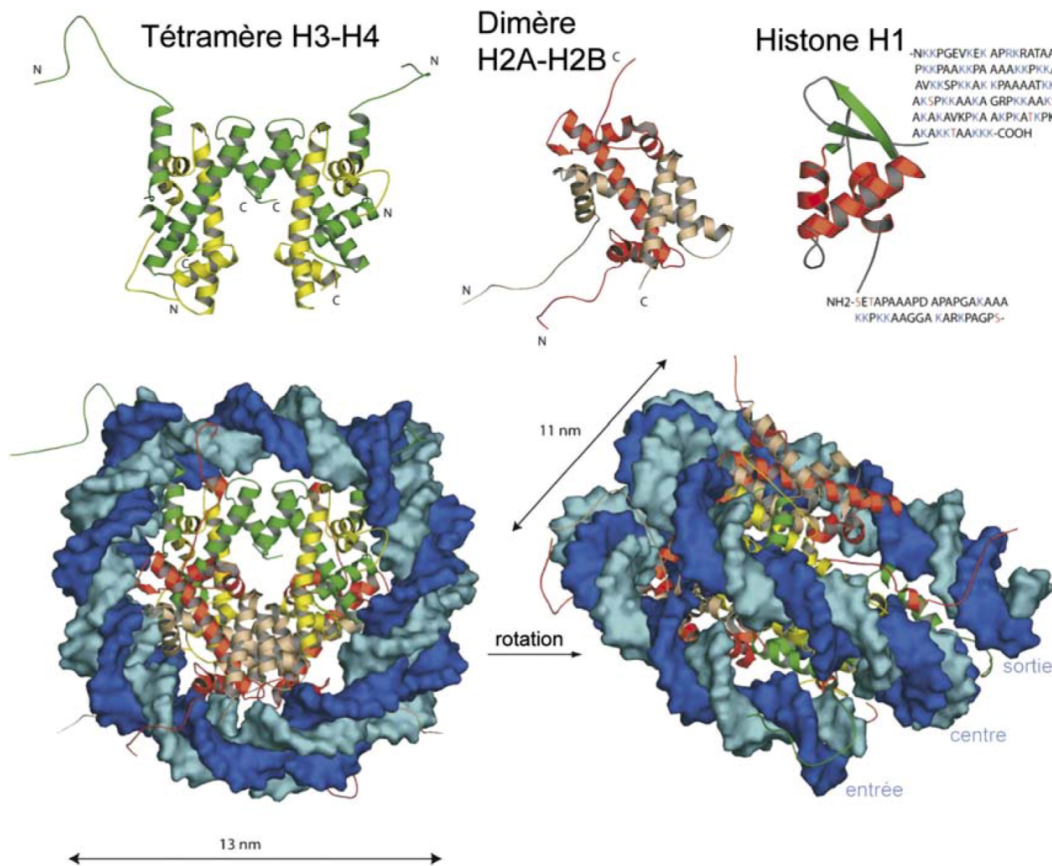
Il existe des arguments en faveur des deux modèles et à l'heure actuelle, il n'est pas possible de trancher en faveur de l'un ou de l'autre (Thomas & Chiang, 2006).

#### La synthèse de la première liaison phosphodiester

Une fois le complexe formé, TFIIH, grâce à son activité hélicase ATP dépendante va permettre de passer d'un complexe fermé à un complexe ouvert. La Pol II peut alors catalyser la formation de la première liaison phosphodiester alors que la sérine 5 du domaine C-terminal de la grande sous-unité de la Pol II est phosphorylé par TFIIH (Thomas & Chiang, 2006).

### **III. Généralités sur la régulation de la transcription**

Pour réguler le taux de transcription dans un type cellulaire ou à un stade donné du développement, des régulateurs de la transcription (activateurs ou répresseurs) peuvent se lier spécifiquement à l'ADN dans la région du promoteur de base et/ou au niveau de séquences de régulation plus éloignées (des « enhancers » ou « silencers »). Le plus souvent, l'information apportée par les régulateurs est transmise à la machinerie transcriptionnelle de base par des co-régulateurs (co-activateurs et co-répresseurs). Ceci peut se faire par une interaction directe entre les co-régulateurs et la machinerie de base ou de façon indirecte grâce à une modification de la structure de la chromatine (Martinez, 2002).



**Figure 2. Structure atomique d'un nucléosome résolue par cristallographie.**

Chaque brin d'ADN est représenté par un bleu différent. L'ADN fait 1,7 tour en s'enroulant autour du nucléosome. En haut à gauche est représenté le tétramère H3-H4, avec H3 en vert et H4 en jaune. En haut au centre est représenté le dimère H2A-H2B, avec H2A en rouge et H2B en rose. Enfin, en haut à droite est représenté l'histone H1 ; les séquences en N-terminal et en C-terminal, non structurées, sont indiquées. La figure est adaptée de : (Khorasanizadeh, 2004).

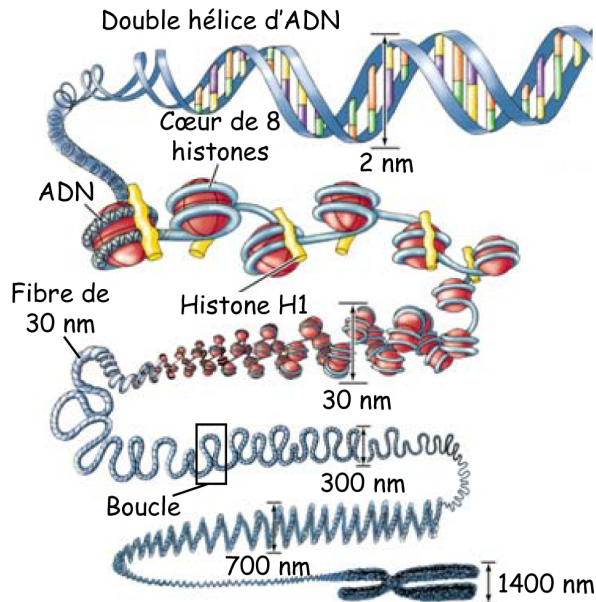
#### IV. L'organisation de la chromatine et la notion de code des histones

##### La chromatine

Chez les eucaryotes, l'ADN est localisé dans le noyau et il interagit avec des protéines histones pour former la chromatine. Le premier niveau d'organisation de la chromatine est le nucléosome. Celui-ci est constitué d'un octamère d'histones et de 146 pb d'ADN. L'octamère comprend deux copies des histones H2A, H2B, H3 et H4 (Figure 2). Les histones sont de petites protéines basiques formées d'un domaine globulaire et d'une région plus flexible et plus chargée en N-terminal. Les histones forment des hétérodimères spécifiques grâce à leurs domaines de type histone (HFD) qui s'assemblent entre eux pour former un octamère. Dans les génomes d'eucaryotes, on



retrouve cette structure tous les  $200 \pm 40$  pb. L'ADN qui se trouve entre deux octamères d'histone est appelé ADN Linker. L'histone H1 quant à lui va permettre de stabiliser les niveaux d'organisation structuraux supérieurs, le nucléosome étant le premier niveau d'organisation de la chromatine. On observe ensuite le collier de perles, la fibre de 30 nm, celle de 300 nm jusqu'au chromosome mitotique (Figure 3).



**Figure 3. Représentation schématique des différents niveaux de compaction de l'ADN.**

Les différents niveaux de compaction de l'ADN sont représentés de la forme la moins condensée, l'ADN libre, en haut, à la forme la plus condensée, le chromosome en métaphase de la mitose, en bas. Le diamètre de la fibre chromatinienne est indiqué à chaque niveau. La figure est adaptée de : (Hansen, 2002).

De façon générale, on distingue deux types de chromatine. La première, l'hétérochromatine correspond à des régions qui ne sont pas transcrites. Elles sont très condensées pendant l'interphase. Certaines sont des zones non codantes comme les centromères ou les télomères (hétérochromatine constitutive) ou des zones non transcrites dans la cellule considérée à un temps donné (hétérochromatine conditionnelle). Le second type de chromatine est l'euchromatine, la chromatine ouverte dans laquelle les gènes peuvent être transcrits (Thomas & Chiang, 2006).

### Le code histone

La structure de la chromatine joue un rôle important dans la régulation de l'expression des gènes. Les queues N-terminales des histones peuvent être modifiées par des enzymes spécifiques. Ces modifications covalentes et réversibles peuvent être notamment des acétylations, des méthylation, des phosphorylations des ADP-ribosylations ou des ubiquitinations et elles ont lieu à des positions spécifiques.

L'hypothèse du code histone propose que les différentes modifications de la chromatine soient reconnues par des protéines régulatrices spécifiques et qu'elles véhiculeraient donc une information particulière. De plus, ces modifications sont interdépendantes. Par ailleurs, il est probable que la structure de la chromatine dépende largement de la concentration locale, du type, et de la position précise des différentes modifications. Par exemple, une hyperacétylation des histones rend la chromatine plus accessible pour la transcription.

La « lecture » de ce code va finalement conduire à l'activation ou à la répression d'un gène et d'une façon plus globale influencer des phénomènes comme la prolifération cellulaire ou la différenciation par exemple. Des états épigénétiques vont être ainsi créés qui vont être transmis aux cellules filles par la méiose ou la mitose. Le code histone apporte donc une information complémentaire à l'information génétique portée par l'ADN (Jenuwein & Allis, 2001). Les enzymes qui vont modifier les queues d'histones jouent donc un rôle crucial dans la régulation de la transcription. Par la suite, nous allons étudier un complexe coactivateur de la transcription qui possède une activité d'acétylation des histones H3 et une activité de déubiquitination de l'histone H2B.

### **V. Les coactivateurs de la transcription**

Ces dernières années, des études biochimiques et génétiques ont mis en évidence de nombreux corégulateurs de la transcription. Ces corégulateurs (co-activateurs et co-répresseurs) sont recrutés la plupart du temps au niveau de promoteurs grâce à des interactions spécifiques avec les domaines d'activation ou de répression des régulateurs de la transcription. Ils peuvent aussi dans certains cas contacter la machinerie transcriptionnelle de base et jouent un rôle prépondérant dans la régulation de la transcription. Il existe deux grandes catégories de corégulateurs : les cofacteurs de remodelage de la chromatine et les adaptateurs. Les adaptateurs sont des cofacteurs qui vont faire le lien entre les régulateurs et la machinerie de transcription de base. Le médiateur est, par exemple, un complexe adaptateur qui joue un rôle crucial dans la régulation de la transcription d'un très grand nombre de gènes. Il va faire, entre autres, le lien entre des activateurs et la grande sous-unité de la Pol II (Martinez, 2002).

Durant ma thèse, j'ai travaillé sur un complexe co-activateur entraînant des modifications de la chromatine, le complexe SAGA. Ces complexes protéiques possèdent une activité enzymatique qui va modifier la structure de la chromatine dont la nature est

intrinsèquement répressive et rendre l'ADN accessible à la machinerie transcriptionnelle. Il existe deux types de complexe comme nous allons le voir ; les complexes de remodelage ATP dépendants et les complexes de modification post-traductionnelle des histones (Martinez, 2002).

#### Les complexes de remodelage ATP dépendants

Ils vont altérer l'association des histones à l'ADN. On peut citer par exemple le complexe SWI-SNF dont on trouve des homologues chez la levure *S. cerevisiae*, la drosophile et l'Homme. Ces complexes sont formés par de nombreuses sous-unités dont une possède une activité ATPase. Trois modes de remodelage de la chromatine ont été mis en évidence : un mouvement de translation de certains octamères d'histones, l'élimination d'un octamère d'histones, ou la modification de la distance entre deux nucléosomes (Venters & Pugh, 2009).

#### Les complexes de modification post-traductionnelle des histones

Ces complexes vont, entre autres, participer à la mise en place du code des histones. Nous allons maintenant présenter en détail le complexe SAGA, qui est considéré comme un modèle de co-activateur modifiant de façon post-traductionnelle les histones grâce à ses activités d'acétylation et de déubiquitination des histones (cf. Partie sur le complexe SAGA).

## **Partie 2 : SAGA, un complexe co-activateur de la transcription**

### **A. Généralités sur le complexe SAGA**

Identifié dans un premier temps chez la levure *S. cerevisiae*, SAGA est un complexe macromoléculaire de près de 2 MDa qui comporte 19 sous-unités dans cet organisme et il apparaît que la composition du complexe SAGA est très conservée au cours de l'évolution de la levure à l'homme (Tableau 1 et Figure 4). En me basant sur les publications originales et sur plusieurs revues récentes (Koutelou et al, 2010; Rodriguez-Navarro, 2009; Samara & Wolberger, 2011; Weake & Workman, 2011), je vais brièvement présenter les diverses fonctions associées au complexe SAGA. J'ai ensuite choisi d'analyser un grand nombre d'approches expérimentales, qui de l'identification des sous-unités du complexe par des cribles génétiques ou des approches de protéomique jusqu'à des études structure fonction très récentes, ont permis de montrer que SAGA est un complexe modulaire et qu'il existe un lien très étroit entre cette organisation structurale et la réalisation des nombreuses fonctions associées à SAGA.

#### **I. Un complexe multifonctionnel**

Le complexe SAGA est considéré comme un prototype de co-activateur avec une activité de modification post-traductionnelle des histones. SAGA est en effet un co-activateur car il peut être recruté au niveau de promoteurs spécifiques par un grand nombre d'activateurs grâce à la sous-unité appelée Tra1 chez la levure. De plus, SAGA possède deux activités enzymatiques de modification de la chromatine : une activité d'acétylation des histones (HAT) et une activité de déubiquitination des histones. De façon intéressante, ces deux activités enzymatiques ne sont pas isolées au sein du complexe SAGA mais sont localisées au sein de modules structuraux qui régulent l'activité des deux enzymes appelées Gcn5/GCN5 et Ubp8/USP22 (Ubiquitin Specific Protease 8 et 22) respectivement chez la levure et l'homme (cf. C. IV. 1. et 2. et Partie Résultats). Chez l'homme, le module HAT (HATm) est composé par les protéines ADA2b, ADA3, SGF29 et GCN5 alors que ATXN7, ATXN7L3, ENY2 et USP22 forment le module de

	<i>S. cerevisiae</i> et <i>S. pombe</i>	<i>D. melanogaster</i>	<i>H. sapiens</i>
<b>Module HAT</b> (Acétylation de H3)	<b>Gcn5</b>	<b>Gcn5</b>	<b>GCN5/PCAF</b>
	Ada2	Ada2b	ADA2b
	Ada3	Ada3	ADA3
	Sgf29	Sgf29	SGF29
<b>Module DUB</b> (Déubiquitination de H2B)	<b>Ubp8</b>	<b>Nonstop</b>	<b>USP22</b>
	Sgf73	CG9866	ATXN7/-L1 /-L2
	Sgf11	Sgf11	ATXN7L3
	Sus1	E(y)2	ENY2
<b>Lien avec des activateurs</b>	Tra1	Tra1	TRRAP
<b>Régulation de TBP</b>	Spt3	Spt3	SPT3
	Spt8	-	-
<b>Cœur structural</b>	TAF5	WDA	TAF5L
	TAF6	SAF6	TAF6L
	TAF9	TAF9	TAF9/TAF9b
	TAF10	TAF10b	TAF10
	TAF12	TAF12	TAF12
	Ada1	Ada1	ADA1
	Spt7	Spt7	SPT7
	Spt20	Spt20	SPT20

**Tableau 1. Composition du complexe SAGA dans différents organismes modèles.**

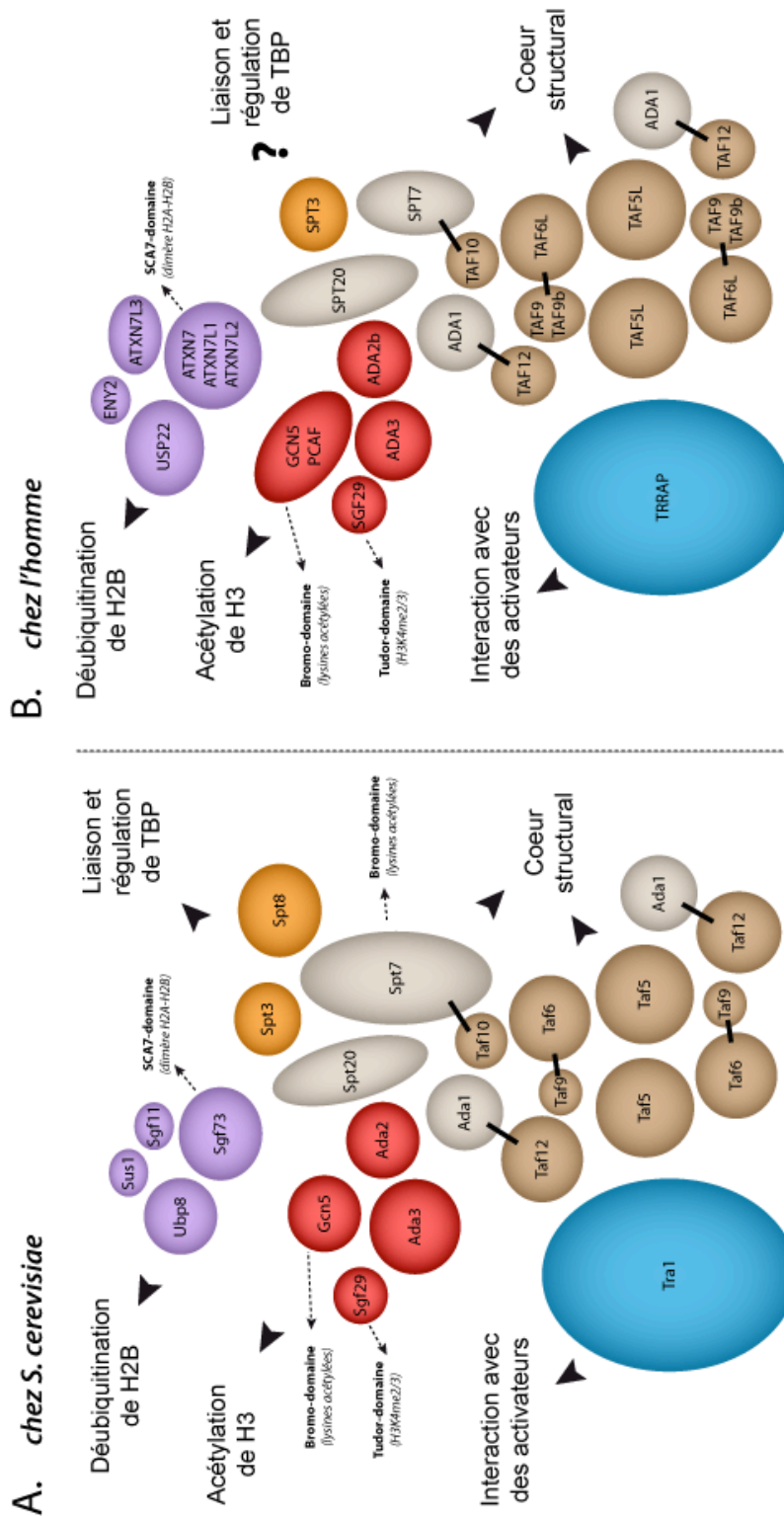
Le contenu en sous-unités des différents modules fonctionnels de SAGA (colonne de gauche) est indiqué chez les levures *S. cerevisiae* et *S. pombe*, la drosophile et l'homme. Les deux lignes correspondant aux activités enzymatiques du complexe sont en gras. Les protéines en noir sont spécifiques du complexe SAGA alors que les protéines en bleu sont partagées avec d'autres complexes. Chez l'homme, certaines sous-unités (ex : GCN5) possèdent des paralogues, qui ont été identifiés au sein du complexe et qui sont donc indiqués dans le tableau. Notons enfin que la fonction de « Régulation de TBP » a été uniquement caractérisée chez *S. cerevisiae* et qu'aucun homologue de Spt8 n'a été identifié chez les eucaryotes supérieurs.

déubiquitination (DUBm) (Tableau 1 et Figure 4). SAGA acétyle plusieurs lysines de la queue N-terminale de l'histone H3 et il est supposé que l'acétylation des histones permet de décompacter la chromatine autour du promoteur ainsi que de recruter des facteurs pouvant reconnaître cette marque ce qui faciliterait la formation d'un complexe de préinitiation de la transcription (Nagy & Tora, 2007). Nous reviendrons brièvement sur les fonctions associées à l'activité HAT de SAGA dans la partie consacrée à l'analyse structure fonction du HATm (cf. C. IV. 1.). L'activité DUB de SAGA est principalement dirigée contre la mono-ubiquitination de H2B (H2Bub) et a été moins étudiée que l'activité HAT. Néanmoins, des études réalisées chez la levure, la drosophile et l'homme montrent que cette activité est requise pour permettre la transcription optimale d'un certain nombre de gènes SAGA dépendants (Henry et al, 2003; Weake et al, 2008). Chez la drosophile, il a aussi été suggéré que l'activité DUB de SAGA permettait d'empêcher la propagation de l'hétérochromatine (Zhao et al, 2008). Dans une prochaine partie, nous reparlerons des fonctions qui sont associées à l'ubiquitination et à la désubiquitination de H2B (cf. Partie sur l'ubiquitination de H2B).

En plus de ces deux modules enzymatiques, SAGA possède un cœur structural composé par plusieurs protéines TAF (TATA-binding protein (TBP)-associated factors) qui sont partagées avec le facteur général de la transcription TFIID et par plusieurs sous-unités spécifiques de SAGA comme Spt7, Spt20 et Ada1 chez *S. cerevisiae* (Figure 4) (Grant et al, 1998a; Sterner et al, 1999). Enfin, chez *S. cerevisiae*, Spt3 et Spt8 interagissent de façon directe avec des membres de la machinerie transcriptionnelle de base comme TBP et régulent la distribution de cette protéine au niveau de certains promoteurs (Figure 4) (Eisenmann et al, 1992; Laprade et al, 2007; Larschan & Winston, 2001; Mohibullah & Hahn, 2008). Il a notamment été montré que le recrutement de TBP et la formation du complexe de préinitiation de la transcription au promoteur du gène *HIS4* dépendait uniquement du complexe SAGA et pas de TFIID (Fishburn et al, 2005). Cette activité de SAGA est également requise pour la mobilisation du rétrotransposon Ty chez *S. cerevisiae* ce qui a permis d'identifier les différentes sous-unités de ce module par le crible génétique SPT (cf. B. I. 1.).

Il a également été suggéré que SAGA participait au couplage entre la transcription et l'export des ARNm. Tout d'abord, chez la levure *S. cerevisiae*, la sous-unité Sus1 du DUBm de SAGA est également retrouvée au sein du complexe TREX2 (Transcription Export 2) qui est associé avec le NPC (Nuclear Pore Complex) et régule l'export des

## SAGA, un complexe multifonctionnel et modulaire



**Figure 4 : Représentation schématique du complexe SAGA chez la levure *S. cerevisiae* (A.) et chez l'homme (B.).**

Chaque sous-unité du complexe SAGA est représentée par un rond ou un oval dont la taille est à peu près proportionnelle à la masse de la protéine et dont la couleur correspond au module fonctionnel auquel appartient la protéine. Les sous-unités du HAT module sont ainsi en rouge, celles du Spt-TBP module en violet, celles du Spt-TBP module en orange, Tra1 est en bleu, les sous-unités structurales spécifiques de SAGA sont en gris et les protéines Taf en marron. La fonction principale de chaque module est indiquée derrière une tête de flèche. La présence du point d'interrogation au niveau de SPT3 dans la partie B illustre que les connaissances actuelles ne permettent pas de confirmer ou d'exclure avec certitude le lien fonctionnel avec TBP chez l'homme. Les paires de type histone sont mises en évidence par un trait noir reliant les deux sous-unités. Les domaines de liaison de la chromatine sont indiqués par une flèche en pointillé et la modification post-traductionnelle ou la région du nucléosome reconnue est spécifiée entre parenthèses. Lorsque une sous-unité du complexe SAGA de levure possède plusieurs homologues chez l'homme, l'ensemble de ces paralogues humains est indiqué (ex : GCN5 et PCAF).

ARNm (Rodriguez-Navarro et al, 2004). De façon intéressante, la délétion de Sus1 chez *S. cerevisiae* a pour conséquence un défaut dans l'export des ARNm et ce phénotype est accentué par la délétion additionnelle de Sgf11, une autre sous-unité du DUBm de SAGA (Kohler et al, 2006). D'autres études ont indiqué que les sous-unités Ada2 (HATm) et

Sgf73 (DUBm) de SAGA régulaient le couplage transcription – export et participaient au positionnement de gènes actifs à proximité de pores nucléaires pour faciliter leur export, un mécanisme appelé « gene gating » (Cabal et al, 2006; Köhler et al, 2008). Chez la drosophile, il a également été montré que e(y)2, la protéine homologue de Sus1, était à la fois présente dans SAGA et dans AMEX (Anchoring and mRNA export), l'homologue de TREX2 (Kurshakova et al, 2007). De plus, les résultats de cette étude montrent que e(y)2 interagit avec le NPC et suggèrent que SAGA et AMEX participent à l'export des ARNm par l'ancrage d'un certain nombre de gènes transcrits à la périphérie nucléaire (Kurshakova et al, 2007).

Enfin plusieurs études proposent que SAGA soit impliqué dans le processus de réparation de l'ADN chez l'homme (Brand et al, 2001; Martinez et al, 2001; Robert et al, 2006) et que SAGA joue un rôle dans le développement de certains cancers, par exemple en régulant des gènes dépendant de l'oncogène cMyc. Pour une revue voir (Koutelou et al, 2010).

## **II. Rôle dans la régulation de l'expression des gènes**

### **1. Chez la levure**

Chez la levure *S. cerevisiae*, aucune sous-unité spécifique du complexe SAGA n'est essentielle. Les seules sous-unités essentielles présentes dans SAGA sont également partagées avec d'autres complexes. C'est le cas des protéines TAF qui sont essentielles car elles sont partagées avec le facteur général de la transcription TFIID qui est indispensable pour la transcription d'une très grande majorité de gènes (Huisinga & Pugh, 2004). D'autre part, Tra1 est essentielle mais la protéine est aussi présente dans le complexe HAT NuA4. Comme Esa1, l'enzyme HAT de ce complexe NuA4, est essentielle et que Gcn5 ne l'est pas, il est donc fort probable que Tra1 soit essentielle à cause de sa présence dans le complexe NuA4 plutôt que dans le complexe SAGA (Allard et al, 1999; Clarke et al, 1999; Smith et al, 1998).

Malgré l'absence de sous-unité essentielle spécifique de SAGA, ce complexe est requis pour une expression optimale d'un très grand nombre de gènes comme le montrent les analyses transcriptionnelles réalisées dans des levures *S. cerevisiae* ou *S. pombe* présentant une délétion de nombreuses sous-unités de SAGA (Helmlinger et al, 2011; Lenstra et al, 2011). Chez *S. cerevisiae*, une autre analyse transcriptionnelle réalisée à partir de plusieurs mutants de SAGA et de TFIID a suggéré que la transcription



de 10% des gènes dépend principalement de SAGA alors que 90% des gènes seraient principalement régulés par TFIID (Huisinga & Pugh, 2004). Les 10% de gènes sous la dépendance de SAGA sont notamment impliqués dans une réponse à divers stress et semblent hautement régulés (Huisinga & Pugh, 2004). Cependant, il convient de rester prudent avec l'interprétation d'un tel modèle. En effet, parmi les gènes dont l'expression requiert TFIID, la transcription optimale de beaucoup d'entre eux nécessite néanmoins la présence du complexe SAGA (Huisinga & Pugh, 2004).

De plus, chez *S. cerevisiae*, de nombreux gènes non exprimés dans les conditions optimales sont régulés par le complexe SAGA. Par exemple, un mutant *gcn5Δ* ou un mutant *spt3Δ* ne va pas pousser correctement en présence de galactose comme seule source de sucre. En effet la délétion de l'enzyme HAT ou celle de Spt3, qui régule la distribution de TBP, provoque une diminution drastique de l'expression de plusieurs gènes *GAL* dont les produits sont nécessaires au métabolisme du galactose (Bian et al, 2011; Dudley et al, 1999; Sterner et al, 1999). Pour conclure, on peut donc dire que sans être absolument requis à la vie des levures dans les conditions optimales, SAGA semble néanmoins nécessaire pour une bonne adaptation des levures à leur environnement chez *S. cerevisiae* et ceci est également le cas chez *S. pombe* (Helmlinger et al, 2011; Sterner et al, 1999).

## 2. Rôle dans le développement des métazoaires

Des mutations au sein de sous-unités de SAGA provoquent des anomalies au cours du développement de souris ou de mouche *Drosophila melanogaster* (Koutelou et al, 2010).

Chez la drosophile, des mutations affectant chacune des deux activités enzymatiques de SAGA ont des conséquences sur le développement embryonnaire. Par exemple, des mutants de sous-unités du DUBm causent une altération de la mise en place de l'axone des photorécepteurs (Weake et al, 2008). De plus, un mutant de Ada2b, une sous-unité du HATm de SAGA, meurt au cours du développement embryonnaire (Pankotai et al, 2005). D'autre part, il semblerait que SAGA puisse exercer une partie de ses fonctions de co-activateur indépendamment de ses activités enzymatiques. En effet, un mutant d'une sous-unité structurale, SAF6 (SAGA associated factor 6, un paralogue de TAF6 spécifique de SAGA (Tableau 1) meurt au cours du développement embryonnaire. Or, en analysant des larves de ces mutants, il a été montré que l'expression de plusieurs gènes régulés par SAGA était diminuée mais que les niveaux globaux de H2Bub et

d'acétylation de différentes lysines de H3 dont K9 et K14 n'étaient pas modifiés (Weake et al, 2009).

Chez la souris, des mutations de SPT20, une sous-unité structurale de SAGA, conduisent à des défauts de la gastrulation (Zohn et al, 2006). De plus, des souris dans lesquelles *Gcn5* est absente (*Gcn5*  $-/-$ ), meurent durant leur développement embryonnaire au jour 10,5 postcoitum à cause de défauts au niveau du mésoderme qui sont dus à une apoptose accrue dans ces tissus (Xu et al, 2000). Il faut cependant préciser que chez les métazoaires, GCN5 est retrouvée au sein de deux complexes et que sa fonction n'est donc pas uniquement liée à SAGA.

En conclusion, il apparaît que SAGA exerce des rôles cruciaux dans la régulation de l'expression des gènes chez les métazoaires, et qu'un complexe SAGA intègre est nécessaire pour le développement de ces organismes.

### **III. Recrutement aux promoteurs de gènes cibles**

#### **1. Localisation et mode d'action**

La première étape dans l'activation de la transcription par le complexe SAGA nécessite le recrutement de celui-ci au niveau des promoteurs de certains gènes cibles. Des expériences de ChIP réalisées chez la levure *S. cerevisiae* ont montré qu'après l'induction du gène modèle *GAL1*, l'activateur Gal4 recrute SAGA sur le promoteur de ce gène avant la PolII et la formation du PIC (Bhaumik & Green, 2001; Bryant & Ptashne, 2003; Larschan & Winston, 2001). Il a aussi été montré que la perte d'une sous-unité du HATm provoque une diminution importante de l'acétylation des histones au niveau du promoteur de *GAL1* ce qui réduit fortement la transcription du gène (Bian et al, 2011; Sterner et al, 1999). De plus, la délétion de *SPT3* empêche le recrutement de TBP au promoteur de *GAL1* ce qui bloque l'induction transcriptionnelle de ce gène (Dudley et al, 1999). Comme nous l'avons vu, les deux mutants *gcn5Δ* ou *spt3Δ* présentent un retard de croissance sur un milieu ne possédant que du galactose comme source de sucre. De plus, un double mutant *gcn5Δ spt3Δ* présente un phénotype encore plus sévère sur un tel milieu ce qui indique que l'activité HAT de SAGA et celle de régulation de TBP agissent de façon complémentaire pour réguler l'expression du gène *GAL1* (Sterner et al, 1999). Cependant, comme nous le verrons par la suite, il n'est pas possible de généraliser cette observation à tous les promoteurs régulés par SAGA. Enfin, des études plus récentes indiquent que l'activité DUB de SAGA est aussi nécessaire pour une

induction transcriptionnelle optimale de *GAL1* (Henry et al, 2003). L'effet de la délétion de l'ubiquitine protéase Ubp8 sur la transcription de *GAL1* est plus modeste que celui observé en l'absence de Gcn5 ou de Spt3. Cependant, de façon intéressante, l'effet transcriptionnel sur *GAL1* ainsi que le phénotype de retard de croissance sur un milieu galactose est plus fort dans une souche *gcn5Δ ubp8Δ* que dans une souche *gcn5Δ* (Henry et al, 2003; Köhler et al, 2008; Lee et al, 2005). On peut donc conclure que l'ensemble des activités de SAGA participe à la régulation de la transcription du gène *GAL1*.

D'autres expériences de ChIP ont démontré que le complexe SAGA était enrichi au niveau du promoteur de très nombreux gènes dans diverses espèces de levures mais aussi chez les eucaryotes supérieurs (Helmlinger et al, 2011; Lang et al, 2011; Sellam et al, 2009; Venters et al, 2011; Weake et al, 2011). Par exemple, la sous-unité Ada2 du HATm de SAGA est localisée sur environ 200 promoteurs de gènes impliqués dans une réponse à différents stress ainsi que de gènes de voies métaboliques chez le pathogène *Candida albicans*. De plus, dans des cellules humaines, les sous-unités du complexe SAGA SGF29, TRRAP, SPT3, SPT20, ATXN7L3 et ADA3 ont été localisées par ChIP sur les promoteurs de gènes permettant la réponse à un stress du réticulum endoplasmique, de gènes associés à la régulation de la transcription, de la réplication et du développement (Nagy et al, 2009; Vermeulen et al, 2010) (cf. Partie Résultats). Une étude récente a aussi montré que dans des cellules humaines, la sous-unité SPT20 de SAGA était associée au promoteur de gènes de ménage dont l'expression est constitutive (Krebs et al, 2011).

## 2. Recrutement sur la chromatine

De nombreuses études montrent que chez *S. cerevisiae*, Tra1 joue un rôle crucial dans le recrutement de SAGA au niveau de promoteurs spécifiques en interagissant de façon directe avec des activateurs de la transcription (Figure 4) (Bhaumik et al, 2004; Brown et al, 2001; Fishburn et al, 2005; Reeves & Hahn, 2005). Chez les mammifères, il a été montré que l'orthologue de Tra1, TRRAP (Transformation/transcription domain-associated protein) interagit directement avec des activateurs de la transcription comme les facteurs oncogéniques c-Myc et E2F (McMahon et al, 1998). On peut alors se demander si l'interaction de Tra1/TRRAP avec des activateurs est suffisante pour recruter SAGA sur l'ensemble de ses gènes cibles. Un travail récent réalisé chez la levure *S. pombe* montre que cela n'est pas le cas. En effet, dans cet organisme, il existe deux paralogues Tra1 et Tra2, le premier étant non essentiel et spécifique de SAGA et le second étant spécifique du complexe HAT NuA4. De façon surprenante, des expériences

de CHIP dans une souche *tra1Δ* montrent que le recrutement du complexe SAGA est perdu sur le promoteur de certains gènes cibles mais qu'il n'est pas altéré sur le promoteur d'autres gènes (Helmlinger et al, 2011).

D'autres sous-unités de SAGA devraient donc aussi être impliquées dans des interactions avec des activateurs transcriptionnels. Chez la levure, des études biochimiques suggèrent que, en plus de Tra1, TAF12 pourrait interagir directement avec des activateurs (Fishburn et al, 2005; Reeves & Hahn, 2005). Chez *S. cerevisiae*, l'interaction de Spt3 et Spt8 avec TBP pourrait également faciliter le recrutement de SAGA au niveau de certains promoteurs (Larschan & Winston, 2001; Mohibullah & Hahn, 2008). Enfin, plusieurs sous-unités de SAGA possèdent un domaine de liaison à la chromatine et ce type d'interactions pourrait également stabiliser le recrutement du complexe SAGA sur ses gènes cibles (Bian et al, 2011; Hassan et al, 2002; Vermeulen et al, 2010). Par exemple, chez *S. cerevisiae*, le Bromo-domaine de Spt7 et celui de Gcn5, reconnaissent des lysines acétylées et sont importants pour la rétention de SAGA sur certains promoteurs en l'absence d'activateurs de la transcription (Hassan et al, 2002). Récemment, il a également été démontré que le double Tudor-domaine de SGF29 peut interagir spécifiquement avec la di ou la tri méthylation de la lysine K4 de l'histone H3 et que cette interaction est nécessaire au recrutement optimale du complexe SAGA sur l'ensemble de ses cibles chez la levure *S. cerevisiae* comme le suggèrent des expériences de CHIP (Bian et al, 2011; Vermeulen et al, 2010). Dans la partie résultat, nous verrons également qu'un domaine d'une sous-unité du DUBm interagit avec la face latérale du nucléosome au niveau du dimère H2A-H2B et que cette interaction pourrait faciliter le recrutement du DUBm vers son substrat (Figure 4) (cf. Partie Résultats).

En conclusion, on peut dire que des activateurs de la transcription ainsi que des modifications de la chromatine permettent à SAGA d'être présent sur des promoteurs. Cependant, on ne peut pas exclure que SAGA exerce une partie de ses fonctions à d'autres locus génomiques. En effet, les expériences de CHIP ne détectent le complexe SAGA qu'au niveau de promoteurs, là où SAGA est stabilisé par des activateurs. Comme nous le verrons dans la partie Résultats, SAGA doit aussi agir sur la région transcrite des gènes sans y être détecté par des expériences CHIP à cause d'un temps de résidence trop court (cf. Partie Résultats).

## **B. Identification de ses sous-unités et caractérisation du complexe**

SAGA n'est pas simplement un complexe d'acétylation des histones mais ses nombreuses sous-unités lui permettent d'exercer un grand nombre de fonctions. La composition en sous-unités du complexe SAGA est aujourd'hui connue et admise par tous. Chez la levure *S. cerevisiae*, SAGA contient 19 sous-unités ayant été caractérisées de diverses manières :

- par des cribles génétiques basés sur des effets transcriptionnels
- en utilisant des approches de protéomique permettant :
  - o l'identification de sous-unités codées par des gènes essentiels
  - o la caractérisation de protéines de fonction inconnue
- par la recherche d'un phénotype de létalité synthétique avec un facteur d'export des ARNm.

De la caractérisation du gène *SPT3* chez la levure *S. cerevisiae* en 1984, à l'identification de la protéine Sus1, 20 années se sont écoulées. Nous allons voir que les différentes stratégies utilisées pour identifier des sous-unités de SAGA mettent en évidence la nature modulaire et multi fonctionnelle du complexe SAGA. Cette partie est certes historique mais n'est pas une simple chronologie d'identifications de sous-unités. J'ai aussi choisi de travailler sur cet aspect pour décrire les expériences qui sont à la base des connaissances actuelles sur le complexe SAGA que nous avons discutées dans le chapitre précédent. Enfin, j'ai essayé de remettre en lumière certains résultats et hypothèses sur la fonction de SAGA quelques peu oubliés aujourd'hui.

### **I. Cribles transcriptionnels *SPT* et des *ADA***

Un crible génétique permet d'isoler des gènes importants pour un processus biologique. Dans les deux cas qui vont nous intéresser, les levures *S. cerevisiae* qui ont été choisies ne sont pas capable de pousser dans une condition donnée. Un agent mutagène a été utilisé pour provoquer des mutations capables de supprimer ce phénotype d'absence de croissance.

Nous allons voir en détail deux cribles qui ont permis de décrire deux familles de sous-unités de SAGA. Pour cela, je me suis intéressé à des papiers relativement anciens que j'ai essayé de lire de façon critique avec ce que nous savons du complexe SAGA aujourd'hui. J'ai été surpris de voir que plusieurs des premières expériences réalisées pour caractériser les gènes codant des sous-unités de SAGA ne sont plus discutées dans

les reviews actuelles. Ces résultats ont-ils été passés sous silence car la qualité des expériences n'était pas suffisamment bonne ou sont-ils simplement passés de mode ? C'est ce que j'ai essayé de mettre en évidence et de discuter.

### 1. Gènes *SPT* de la classe TBP

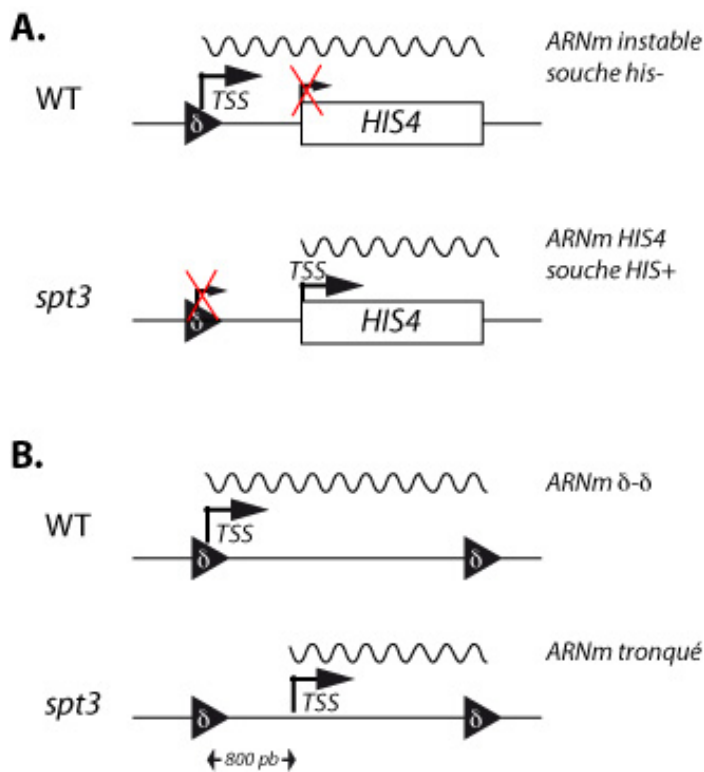
#### Description du crible

Le crible génétique SPT (Suppression of Ty) est basé sur l'utilisation du rétrotransposon Ty de la levure *S. cerevisiae*. Ce rétrotransposon est composé par deux régions. La partie centrale de 5,9 kb contient deux gènes codant des protéines similaires aux protéines virales Gag et Pol. Cette région est entourée par deux répétitions directes d'une séquence appelée  $\partial$ . L'insertion de l'élément Ty en amont d'un gène peut modifier son expression. Suivant le gène et le lieu de l'insertion, une grande variété de phénotypes allant de l'expression constitutive à la perte de l'expression du gène peut être observée. La souche utilisée pour ce crible génétique dérive d'une souche présentant un élément Ty en amont du gène *HIS4*. Dans cette souche, le rétrotransposon a été mobilisé et il ne reste plus qu'une seule séquence  $\partial$ . Cette insertion d'un élément  $\partial$  provoque la perte complète du produit du gène *HIS4* et la souche présente donc une auxotrophie pour l'histidine, c'est à dire une incapacité à synthétiser l'histidine (Figure 5 A.). Le but de ce crible génétique a donc été de caractériser des mutants pouvant pousser sur un milieu ne contenant pas d'histidine (Winston et al, 1984a).

#### Identification de SPT3

Cette stratégie a permis l'identification de plusieurs gènes dont *SPT3* (Suppression of Ty 3) dont la fonction a été étudiée en détail (Winston et al, 1984a). Un défaut de la transcription de l'élément Ty a pu être mis en évidence dans une souche mutante *spt3* (Figure 5 B.). En effet, dans une souche sauvage, l'élément Ty peut être transcrit, l'initiation de la transcription ayant lieu dans la première séquence  $\partial$  de Ty et se terminant dans la deuxième séquence  $\partial$ . Dans une souche *spt3*, la transcription  $\partial$ - $\partial$  disparaît et un transcrit commençant 800 pb en aval de la première région  $\partial$  est synthétisé (Figure 5 B.).

Une souche *spt3* ne pouvant plus initier la transcription depuis une séquence  $\partial$ , l'expression des gènes possédant un élément Ty dans leur région 5' n'est donc plus altérée dans cette souche (Winston et al, 1984b). Comme la souche *spt3* présente également un défaut dans la reproduction sexuée et dans la sporulation, les auteurs



**Figure 5. Le crible génétique SPT.**

**A.** Dans une souche sauvage (WT), l'insertion de la séquence  $\delta$  du rétrotransposon Ty et la transcription à partir de cet élément  $\delta$ , en amont du promoteur du gène *HIS4*, bloque la transcription correcte de ce gène et provoque une auxotrophie de la souche pour l'histidine. Chez un mutant *spt3*, la transcription depuis l'élément  $\delta$  est perdue et le gène *HIS4* est normalement exprimé.

**B.** Pour mobiliser le rétrotransposon Ty, une souche sauvage (WT) est capable de transcrire l'élément Ty depuis sa première séquence  $\delta$  jusqu'à sa seconde séquence  $\delta$ . En revanche, une souche mutante *spt3* ne peut plus initier la transcription depuis l'élément  $\delta$  et perd donc la capacité à mobiliser l'élément Ty.

proposent que *SPT3* exerce une fonction plus générale que la seule transcription des éléments Ty (Winston et al, 1984b).

#### Sous-groupe de gènes *SPT3*

Les mutations perte de fonction des différents gènes identifiés lors du premier crible génétique SPT permettent à la souche utilisée de perdre son auxotrophie pour l'histidine (Winston et al, 1984a). Par contre, seule une souche *spt3* perd la capacité d'initier la transcription depuis l'élément  $\delta$  et produit un transcrit de l'élément Ty raccourci de 800 pb (Winston et al, 1987). Un crible complémentaire a été réalisé pour identifier de

nouveaux gènes ayant les mêmes caractéristiques que *SPT3*. Deux gènes, *SPT7* et *SPT8*, ont pu être ainsi isolés. La plupart des phénotypes de la souche *spt3* sont retrouvés pour des mutants de ces deux nouveaux gènes comme par exemple un défaut de sporulation (Winston et al, 1987).

Il a également été montré que les produits des gènes *SPT7* et *SPT8* ne régulent pas la transcription de *SPT3* (Winston et al, 1987). Il semble donc que les produits de ces trois gènes agissent ensemble pour réguler notamment la transcription de l'élément Ty depuis sa séquence  $\delta$ . Les auteurs de l'étude imaginent plusieurs mécanismes pour expliquer cela. L'une des propositions est : (je cite) « les trois produits des gènes pourraient interagir pour former un complexe qui serait directement impliqué dans l'activation de la transcription depuis la séquence  $\delta$  » (Winston et al, 1987). C'est la

première mention d'un complexe putatif qui contiendrait au moins Spt3, Spt7 et Spt8 et rétrospectivement, une preuve génétique de la modularité fonctionnelle de SAGA. Il est également proposé qu'un de ces trois facteurs pourrait lier l'ADN et reconnaître spécifiquement la boîte TATA de la séquence  $\delta$  (Winston et al, 1987). En fait, même si cette dernière proposition n'est pas rigoureusement exacte, on sait aujourd'hui que ces trois protéines font partie d'un module fonctionnel de SAGA qui interagit avec TBP pour réguler l'initiation de la transcription chez la levure *S. cerevisiae* (Eisenmann et al, 1992; Laprade et al, 2007; Larschan & Winston, 2001; Mohibullah & Hahn, 2008) (cf. A. I.).

#### La fonction du sous-groupe *SPT3* est reliée à celle de TBP

3 mutations dans le gène *SPT15* ont également été isolées par le crible SPT. Les souches qui présentent des mutations faux-sens dans *SPT15* présentent un phénotype sévère ce qui suggère que le produit de ce gène est requis pour une fonction très importante voire essentielle dans la cellule. Le clonage du gène *SPT15*, son séquençage et une analyse génétique ont permis de montrer que *SPT15* code la protéine TBP (Eisenmann et al, 1989). En effet, TBP venait d'être identifiée indépendamment par une autre équipe qui avait purifié biochimiquement cette protéine sur la base de son activité dans des tests de transcription *in vitro* (Hahn et al, 1989).

Plusieurs autres gènes ont été isolés par le crible SPT dont des gènes codant les histones. Cependant, un sous-groupe de gènes *SPT*, appelé le groupe TBP, dans lequel on retrouve *SPT15*, *SPT3*, *SPT7* et *SPT8* a pu être défini. Comme nous l'avons vu, ces gènes partagent en effet de nombreuses propriétés. Des mutants de ces gènes affectent notamment l'expression d'un même ensemble de gènes. La fonction des produits de ces gènes est donc supposée être liée à celle de TBP et à l'initiation de la transcription.

F. Winston et coll. ont voulu poursuivre leur crible génétique afin d'identifier de nouveaux gènes impliqués dans la régulation de la fonction de TBP. Un nouveau crible a été réalisé en utilisant une souche haploïde dans laquelle une deuxième copie des gènes *SPT3* et *SPT15* avait été introduite pour éviter d'identifier des mutants de ces gènes. Ceci a permis d'identifier le gène *SPT20* qui présente un défaut dans la transcription de l'élément Ty comme tous les mutants de la classe TBP. Cependant, les mutants *spt20* présentent un retard de croissance sur milieu riche, sont auxotrophes pour l'inositol et ne poussent pas sur un milieu ne contenant que du galactose comme sucre. Dans ces deux derniers cas, ce phénotype s'explique par l'absence de la transcription de gènes impliqués dans ces voies métaboliques (*INO1*, *GAL1* et *GAL10* par exemple) comme cela



a été montré par Northern Blot (Roberts & Winston, 1996). Il est intéressant de noter que ces phénotypes particuliers des mutants *spt20* sont également retrouvés chez des mutants *spt7* et *spt15*. Au total, ces résultats sont en accord avec les données actuelles concernant la modularité de SAGA et ils indiquent un lien fonctionnel entre d'une part Spt3 et Spt8 qui sont uniquement impliquées dans la régulation de TBP et d'autre part Spt20 et Spt7 qui en plus de participer à la régulation de TBP, jouent un rôle dans l'intégrité structurale de SAGA.

## **2. Le crible ADA : identification d'adaptateurs**

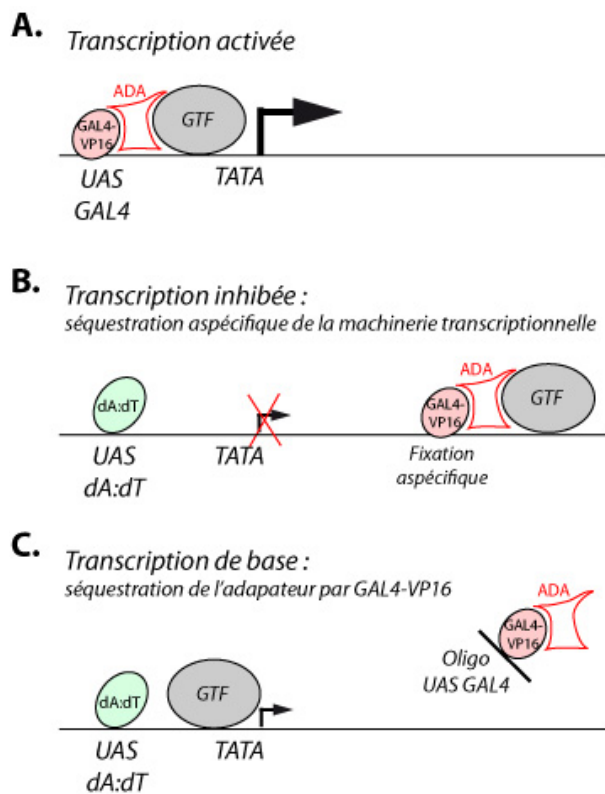
Le gène codant Spt20 a été identifié par une autre équipe de recherche qui a développé un crible génétique qui à première vue a très peu de similarités avec le crible SPT. Ce nouveau crible, que je vais maintenant décrire a permis l'identification d'un groupe de sous-unités de SAGA.

### L'activateur chimérique GAL4-VP16

Une question fondamentale pour comprendre les mécanismes de régulation de l'expression des gènes est de savoir comment des activateurs liés au niveau de séquence distales (des enhancers chez les eucaryotes supérieurs ou UAS (Upstream Activating Sequence) chez la levure) peuvent recruter la machinerie transcriptionnelle de base au niveau d'un promoteur donné. Au début des années 1990, le but de nombreuses équipes était d'identifier des protéines de type « adaptateurs » pouvant faire le lien entre activateurs et facteurs généraux de la transcription. Dans un premier temps, l'équipe de L. Guarente a utilisé l'activateur chimérique GAL4-VP16 pour démontrer l'existence de tels adaptateurs. Cette chimère possède le domaine de liaison à l'ADN de l'activateur GAL4. Le fort domaine d'activation acide provient quant à lui de la protéine VP16, une protéine du virus de l'herpès (Berger et al, 1990).

La fonction de cet activateur chimérique a été analysée par des tests de transcription *in vitro*. Sur une matrice d'ADN nu présentant l'UAS reconnue par GAL4, GAL4-VP16 permet une activation de la transcription en présence d'extraits nucléaires de levure (Figure 6 A.). Par contre, GAL4-VP16 inhibe la transcription en présence d'une UAS non reconnue par GAL4 (ex : UAS dA : dT) ou en présence d'un simple promoteur de base contenant une boîte TATA.

Pour expliquer ces observations, les auteurs proposent qu'en l'absence d'une séquence d'ADN spécifique de l'activateur chimérique GAL4-VP16, celui-ci se fixe de



**Figure 6. Démonstration *in vitro* de l'existence d'adaptateurs.**

**A.** En présence d'un extrait cellulaire de levure, l'activateur chimérique GAL4-VP16 peut activer la transcription à partir d'une matrice d'ADN nue possédant l'UAS (Upstream Activating Sequence) reconnue par GAL4. **B.** En l'absence de l'UAS reconnue par GAL4, l'activateur chimérique GAL4-VP16 se fixe de façon aspécifique sur la matrice d'ADN et y séquestre la machinerie transcriptionnelle de base. **C.** L'ajout d'un oligonucléotide dont la séquence est reconnue spécifiquement par GAL4-VP16 empêche l'activateur chimérique de lier la matrice d'ADN mais lui permet de titrer un adaptateur. Ceci empêche l'activation de la transcription par un autre activateur et seule une transcription de base peut avoir lieu.

manière aspécifique à l'ADN et y séquestre la machinerie transcriptionnelle de base ce qui empêche une transcription correcte (Figure 6 B.). En effet, l'inhibition de la transcription observée *in vitro* en l'absence de l'UAS reconnue par GAL4 est dépendante d'un domaine d'activation de VP16 intègre.

Pour mieux caractériser les effets de cet activateur chimérique, une autre expérience de transcription *in vitro* a été menée. En plus de l'extrait nucléaire et de la chimère GAL4-VP16, un oligonucléotide, dont la séquence est reconnue spécifiquement par GAL4, a été ajouté à la réaction. Dans ce cas, l'utilisation d'une UAS de type dA:dT (non reconnue par la chimère) ne permet pas l'activation de la transcription mais par contre, la transcription de base n'est plus inhibée (Figure 6 C.). On peut donc conclure que l'activateur chimérique,

qui ne peut plus se fixer à l'ADN, va quand même titrer un facteur permettant de faire le lien entre des activateurs et la machinerie transcriptionnelle de base. Cette expérience démontre donc l'existence d'adaptateurs utilisés par les activateurs pour interagir avec la machinerie transcriptionnelle de base dans le but d'activer la transcription (Berger et al, 1990).

### Description du crible

Dans un second temps, les propriétés de l'activateur chimérique GAL4-VP16 ont été mises à profit pour identifier des adaptateurs chez la levure *S. cerevisiae*. L'expression de

GAL4-VP16 est toxique au sein de cet organisme si les deux domaines de la protéine sont intègres. Le recrutement de la machinerie transcriptionnelle de base est sans doute dérégulé ce qui empêche une transcription des gènes correcte. L'idée du crible génétique, basée sur l'identification de mutants permettant de reverser la toxicité de GAL4-VP16, était donc de caractériser des mutations qui pourraient abolir l'interaction entre les adaptateurs et le domaine d'activation acide de VP16 (Berger et al, 1992).

#### Les gènes *ADA1*, *ADA2* et *ADA3*

Les trois premiers gènes à avoir été découverts par ce crible génétique sont *ADA1*, *ADA2* et *ADA3*. Dans un premier temps, le gène *ADA1* n'a pas été étudié davantage car ses propriétés semblaient différentes de celles des autres gènes *ADA*. Nous expliquerons brièvement pourquoi dans un prochain paragraphe (cf. B. I. 3.). Les caractéristiques fonctionnelles des gènes *ADA2* et *ADA3* ont par contre été étudiées en détail et sont relativement similaires. Par exemple, la *trans*-activation par les activateurs acides VP16 et GCN4 est perdue dans des souches mutantes *ada2* et *ada3*. Les mutants *ada2* et *ada3* sont viables mais présentent un défaut de croissance sur milieu minimum ainsi qu'un phénotype de sensibilité à la température. Enfin, une souche dans laquelle *ada2* et *ada3* sont mutés tous les deux ne présente pas d'aggravation du phénotype. Ceci suggère que les produits des deux gènes se trouvent dans une même voie fonctionnelle. L'ensemble de ces critères a permis de définir le groupe des gènes *ADA* (Berger et al, 1992; Pina et al, 1993).

Parmi les expériences réalisées pour caractériser le gène *ADA2*, une d'entre elles a retenu mon attention. Des tests de transcription *in vitro* sur des matrices d'ADN nu ont été effectués à partir d'extraits protéiques provenant de souches mutantes *ada2*. Contrairement à ce qui se passe avec un extrait sauvage, l'ajout d'activateurs à l'extrait *ada2* comme GAL4-VP16 ou GCN4 ne permet pas d'activer la transcription. Cette observation n'est pas faite avec tous les types d'activateurs mais elle semble suggérer le produit du gène *ADA2* exerce un rôle d'adaptateur physique entre certains activateurs et la machinerie transcriptionnelle de base (Berger et al, 1992).

Les mutants *ada2* et *ada3* présentent également un certain nombre de spécificités. L'une d'entre elles m'a intéressé, car elle amène les auteurs de l'étude à discuter d'un potentiel lien fonctionnel entre *ADA3* et les gènes de la famille *SPT*. En effet, les extraits préparés à partir de souches mutantes *ada3* ont un effet sur la transcription de base *in vitro*, ce qui n'est pas observé avec le mutant *ada2*. Le choix du

site d'initiation de la transcription est notamment modifié avec un extrait *ada3* par rapport à un extrait sauvage. Ceci suggère une potentielle relation fonctionnelle entre les protéines Ada3 et TBP. De plus, ce phénotype rappelle celui observé avec les mutants *spt3*, *spt7* et *spt8*. Un possible lien entre *ADA3* et la famille des gènes *SPT* a alors été suggéré sans être démontré (Pina et al, 1993).

#### *ADA4* code la protéine Gcn5

En utilisant le même crible génétique basé sur l'utilisation de l'activateur chimérique GAL4-VP16, un nouveau gène possédant les caractéristiques du groupe *ADA* a été identifié (Marcus et al, 1994). Dans un premier temps, il est appelé *ADA4*, mais suite à son clonage, les chercheurs de l'équipe de L. Guarente se rendent compte que ce gène a déjà été identifié dans un autre crible génétique et a été appelé *GCN5*. A l'époque, une des seules informations connues sur la protéine Gcn5 était sa séquence et donc le fait qu'elle possédait un Bromo-domaine. La fonction de ce domaine n'était pas caractérisée mais ce type de domaine avait été trouvé dans plusieurs facteurs impliqués dans la régulation de la transcription. Lors du clonage de *GCN5*, on a donc supposé que la fonction du produit de ce gène était impliquée dans la régulation de la transcription. Le fait que *GCN5* apparaisse dans le crible *ADA* en est la confirmation (Marcus et al, 1994).

A cette même époque, il a été montré que la protéine p55 du protozoaire cilié *Tetrahymena thermophila*, peut acétyler *in vitro* des histones à partir d'acétyl CoA (Brownell & Allis, 1995). L'identification du gène codant p55 chez *Tetrahymena* a permis de déterminer la séquence de cette protéine. Il en ressort que p55 présente une très forte conservation de séquence avec Gcn5 ce qui suggère que ces deux protéines sont homologues. De plus, dans les mêmes conditions expérimentales, une activité similaire à celle de p55 a été observée avec Gcn5. Ceci démontre donc que Gcn5 est une enzyme pouvant acétyler les histones (Brownell et al, 1996). Enfin, l'équipe de D. Allis a montré *in vitro* que la protéine recombinante Gcn5 acétylait de façon préférentielle la lysine 14 de l'histone H3 et la lysine 8 de l'histone H4 de levure (Kuo et al, 1996). Nous avons parlé de l'importance de l'activité HAT de SAGA dans les généralités sur le complexe SAGA et nous discuterons de la spécificité de substrat de Gcn5 dans un prochain paragraphe (cf. C. IV. 1.).

### Interactions physiques

La caractérisation de Gcn5 comme membre de la famille de gènes *ADA* permet de mettre en évidence la première interaction physique entre des sous-unités du futur complexe SAGA. En effet, Ada2 interagit de façon directe avec Gcn5 *in vitro* ainsi que dans un test de double hybride (Marcus et al, 1994). Dans cette même étude, les auteurs proposent que les produits des gènes de la famille *ADA* puissent exercer leur fonction au sein d'un complexe (Marcus et al, 1994). L'année d'après, un complexe recombinant contenant Gcn5, Ada2 et Ada3 est purifié. Une interaction directe entre la région N-terminale de Ada2 et la région centrale de Gcn5 est observée ainsi qu'entre la région centrale de Ada2 et la région C-terminale de Ada3 (Barlev et al, 1995; Horiuchi et al, 1995).

### *ADA5* code la protéine Spt20

Le crible ADA a également permis l'identification d'un nouveau gène appelé *ADA5* (Marcus et al, 1994). Ce gène partage les propriétés des autres gènes *ADA*. De plus, une interaction physique a pu être démontrée entre Ada5 et Ada3 *in vivo*. Cependant, des mutants du gène *ADA5* présentent également un phénotype plus sévère que les autres gènes *ADA*. La délétion du gène *ADA5* provoque notamment un retard de croissance sur un milieu riche. De plus, la quantité d'ARNm codant l'activateur chimérique GAL4-VP16 est diminuée en l'absence de *ADA5*. Comme le gène codant GAL4-VP16 est sous la dépendance du promoteur *ADH1*, les auteurs de l'étude ont voulu déterminer l'importance de *ADA5* dans l'initiation de la transcription depuis divers promoteurs. Pour cela, un gène rapporteur sous la dépendance de différents types d'UAS a été utilisé chez la levure. Comme attendu dans une souche *ada5*, une perte de la trans-activation est observée avec les UAS affectées par la délétion de *ADA2* ou *ADA3*. Par contre, l'activation de la transcription est aussi perdue avec plusieurs autres UAS comme *ADH1*, *HO* ou *INO1* dans une souche *ada5Δ* ce qui n'était pas le cas avec les autres gènes *ADA* (Marcus et al, 1996). Enfin, en comparant les séquences des gènes *ADA5* et *SPT20*, les chercheurs se sont rendus compte qu'il s'agissait en fait du même gène (Marcus et al, 1996; Roberts & Winston, 1996).

Lors du clonage du gène *ADA5*, L. Guarente et coll. ont observé qu'un fragment de *ADA5* codant les 437 premiers résidus de la protéine Ada5/Spt20 peut partiellement compléter les phénotypes observés avec une souche *ada5Δ*. Ce fragment permet notamment de supprimer le phénotype de retard de croissance sur milieu riche et atténuer celui observé sur milieu minimal. Par contre, ce mutant de *ada5* est toujours

résistant à l'expression de l'activateur chimérique GAL4-VP16. Ceci suggère donc que la région C-terminale de Ada5/Spt20 pourrait interagir avec le module d'acétylation de SAGA qui contient les produits des autres gènes *ADA*. L'observation d'un alignement multiple des protéines Ada5/Spt20 de plusieurs espèces de levure, permet de délimiter une seule région conservée qui pourrait remplir cette fonction (résidus 465 à 520 chez *S. cerevisiae*).

### 3. Chevauchement des cribles SPT et ADA

Comme nous l'avons vu précédemment, les mutants *ada1* présentaient des phénotypes différents de ceux observés avec les souches *ada2* et *ada3*. Dans un premiers temps, l'étude de *ADA1* n'avait pas été continuée (Berger et al, 1992), mais il apparaît en fait, qu'en plus des phénotypes propres aux gènes *ADA*, *ADA1* possède aussi les caractéristiques des gènes *SPT7* et *SPT20* (Horiuchi et al, 1997; Marcus et al, 1996).

Dans leurs objectifs initiaux, ces deux cribles génétiques paraissent totalement indépendants. Ils ont en effet permis d'identifier deux groupes de gènes qui sont chacun caractérisés par un certain nombre de propriétés. De façon plus intéressante, plusieurs gènes identifiés par l'un ou l'autre des cribles répondent aux critères définissant ces deux groupes. En effet, *ADA1*, *ADA5/SPT20*, *SPT7* et *SPT15* présentent à la fois le phénotype ADA et aussi le phénotype SPT. A l'inverse, *ADA2*, *ADA3* et *GCN5* présentent uniquement le phénotype ADA. Enfin, parmi les gènes *SPT* du sous-groupe TBP, seuls *SPT3* et *SPT8* ne présentent que le phénotype SPT.

D'un point de vue génétique, ces observations suggèrent donc un lien fonctionnel entre les gènes *SPT* de la classe TBP (*SPT3*, *SPT7*, *SPT8* et *SPT20*) et les gènes *ADA*. Pour expliquer ces observations génétiques, une des possibilités discutées dans l'article de F. Winston est que l'ensemble des produits de ces gènes exerce leur fonction au sein d'un complexe. Spt20, Spt7 et Ada1 pourraient alors faire le lien entre deux régions fonctionnelles, l'une contenant les produits des gènes *ADA* et l'autre les produits des gènes *SPT* (Roberts & Winston, 1996).

### 4. Identification du complexe SAGA

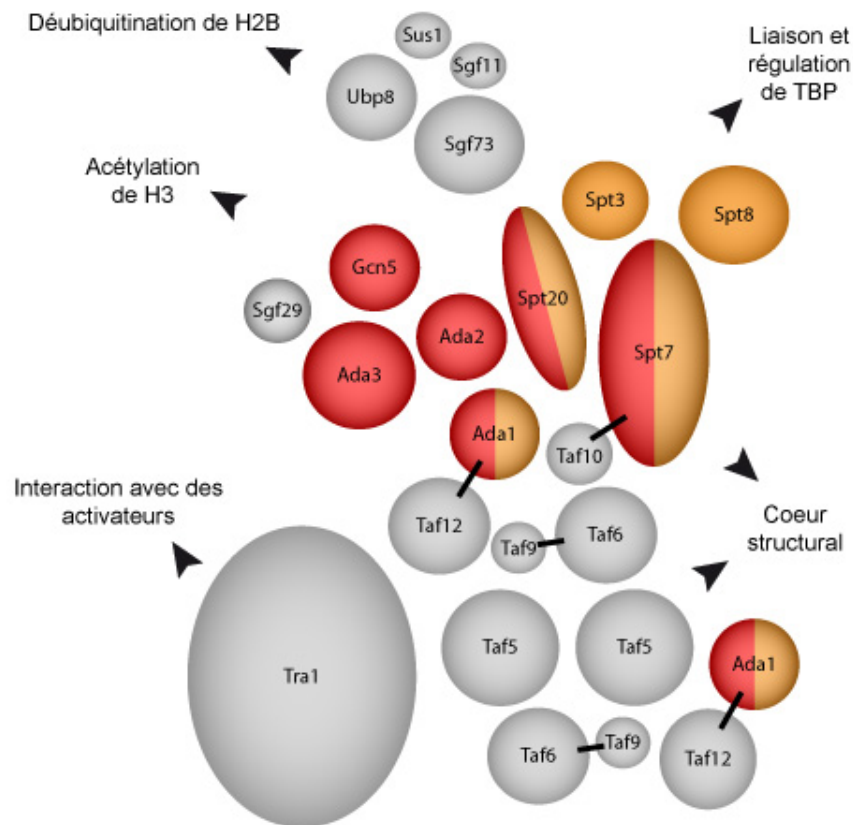
Afin de purifier un tel complexe contenant des protéines Spt et Ada, l'équipe de J. Workman a mis à profit le fait que Gcn5 possède une activité d'acétylation des histones. Ces chercheurs ont montré que la protéine recombinante Gcn5 n'est capable d'acétyler l'histone H3 que s'il est libre mais pas lorsqu'il est incorporé au sein d'un nucléosome.

Par contre, dans un extrait de levure, il existe des activités enzymatiques pouvant acétyler un nucléosome. Après fractionnement d'un extrait protéique de levure par une colonne échangeuse d'ions, ces différentes activités ont été séparées en plusieurs pics qui possèdent chacun une activité d'acétylation de nucléosomes. Par western blot, en utilisant des anticorps spécifiques, les auteurs ont alors voulu déterminer les protéines présentes dans chaque pic. Dans l'un d'eux, la présence de Ada2 et Gcn5 ainsi que des protéines Spt3, Spt7, et Spt20 a été observée. Une fractionnement biochimique plus fine a ensuite montré que l'ensemble de ces protéines fait partie d'un complexe macromoléculaire qui a une masse apparente d'environ 1,8 MDa. Ceci démontre bien qu'il existe un lien physique entre les protéines Ada et les protéines Spt. Ce complexe a donc été logiquement appelé SAGA pour Spt - Ada - Gcn5 Acétyl transférase (Grant et al, 1997).

### 5. Cribles génétiques et modularité fonctionnelle

La description précise et détaillée de ces deux cribles génétiques nous permet de comprendre quelle a été la première vision fonctionnelle d'un certain nombre des gènes codant des sous-unités de SAGA. De plus, de façon très significative, ces deux cribles nous donnent déjà une vision assez précise de l'organisation fonctionnelle et modulaire du complexe SAGA (Figure 7). En effet, les trois gènes *ADA1*, *SPT7* et *SPT20*, qui partagent les deux phénotypes ADA et SPT, sont impliqués dans l'architecture de SAGA et doivent participer au recrutement des autres modules fonctionnels au sein du complexe. Dans l'un de ces complexes, on retrouve les protéines Spt3 et Spt8 qui régulent la fonction de la protéine TBP chez la levure. On peut sans doute retenir de ce premier chapitre que ce module est requis pour la transcription du rétrotransposon Ty ainsi que pour la reproduction sexuée des levures. Ce ne sont peut-être pas des fonctions essentielles à court terme mais sur le long terme, cette fonction de SAGA doit être cruciale pour la vie des levures et leur adaptation à divers conditions environnementales. Enfin, nous verrons que les produits des gènes de la famille *ADA* interagissent entre eux pour former un module d'acétylation au sein de SAGA. A ce propos, je voudrais souligner ici une observation surprenante. La vision actuelle est que Ada2, Ada3, Gcn5 et Sgf29 [le gène codant cette protéine possède aussi les propriétés fonctionnelles du groupe ADA (Lee et al, 2011a)] permettent la formation du module d'acétylation des histones. Cependant, nous venons de voir que les gènes *ADA* ont été identifiés à la base comme des « adaptateurs » entre un activateur et des facteurs de la

## Partie 2 : SAGA, un complexe co-activateur de la transcription



**Figure 7. Modularité du complexe SAGA définie par les cribles génétiques.**

Chaque sous-unité du complexe SAGA de la levure *S. cerevisiae* est représentée par un rond ou un oval dont la taille est à peu près proportionnelle à la masse de la protéine. Les sous-unités dont la délétion cause un phénotype SPT sont colorées en orange, celles dont la délétion provoque un phénotype ADA sont en rouge et celles dont la délétion conduit aux deux phénotypes sont représentées en orange et en rouge. Les sous-unités indiquées en gris n'ont pas été identifiées lors des cribles génétiques. La fonction des modules de SAGA définis dans la Figure 4 est indiquée derrière une tête de flèche.

machinerie transcriptionnelle de base. Tout le phénotype ADA peut-il alors être expliqué par la seule activité d'acétylation portée par Gcn5 ? Pour y répondre, il serait intéressant d'analyser les propriétés d'une souche dans laquelle les résidus catalytiques de Gcn5 seraient mutés. Une réponse partielle a été apportée par l'équipe de S. Berger qui a analysé une série de mutants de résidus du domaine acétyl-transférase de Gcn5 (Wang et al, 1998). Pour chacun de ces mutants, l'activité d'acétylation des complexes contenant Gcn5 a été mesurée *in vitro* et des tests de trans-activation avec les activateurs chimériques LexA-Gcn4 et GAL4-VP6 ont été réalisés *in vivo*. Les résultats indiquent une très bonne corrélation entre un défaut d'acétylation *in vitro* et la perte de la trans-activation *in vivo* (Wang et al, 1998). Cependant, à ma connaissance, aucune étude ne démontre clairement qu'un mutant catalytique de Gcn5 présente l'ensemble des phénotypes caractéristiques de la famille ADA. Et si c'était le cas, comment devons nous interpréter les expériences de transcription *in vitro* qui démontreraient la perte de



l'activation par GAL4-VP16 sur une matrice d'ADN nu en utilisant un extrait protéique provenant d'une souche mutante *ada2* (Berger et al, 1992). Certes, les interactions directes observées *in vitro* entre plusieurs protéines Ada et le domaine d'activation de VP16 ou entre Ada2, VP16 et TBP ne sont peut-être pas réelles *in vivo* (Barlev et al, 1995; Marcus et al, 1996). Par contre, les tests de transcription *in vitro*, qui étaient le point de départ du crible ADA (cf. B. I. 2.) me semblent quant à eux plutôt convaincants (Berger et al, 1990; Berger et al, 1992). Enfin, on peut noter que deux mutants du gène *SPT15* (le gène qui code TBP) présentent également une résistance partielle à GAL4-VP16, sans modifier la quantité de l'activateur chimérique. Ceci semble donc suggérer un lien entre les produits des gènes *ADA* et la machinerie transcriptionnelle avec TBP (Marcus et al, 1996).

De toute évidence, l'activité HAT de Gcn5 est cruciale pour l'activation de la transcription. Cependant, je crois qu'il ne faut pas oublier que les protéines identifiées par le crible ADA pourraient bien avoir une fonction additionnelle et complémentaire en pouvant faire le lien entre un certain nombre d'activateurs et la machinerie transcriptionnelle de base.

## **II. Caractérisation biochimique chez la levure**

### **1. Analyse protéomique**

La caractérisation du complexe SAGA a bénéficié de l'utilisation de la technique de spectrométrie de masse. Couplée à de très bonnes purifications biochimiques, cette technique a permis de déterminer très précisément la composition en sous-unité du complexe SAGA. Nous allons voir que, grâce à ces méthodes, des protéines essentielles à la vie de la cellule ont été identifiées dans le complexe SAGA. Dans un second temps, ces techniques ont permis d'identifier de nouvelles sous-unités de SAGA qui lors de ces études n'avait pas de lien évident avec la transcription. J'ai choisi de décrire rapidement ces expériences pour davantage me concentrer sur les liens structuraux et fonctionnels qui existent entre ces protéines.

#### Identification de sous-unités essentielles

Nous avons vu que plusieurs étapes de chromatographie successives ont permis d'isoler un complexe appelé SAGA (Grant et al, 1997). L'analyse par spectrométrie de masse du complexe purifié retrouve des sous-unités déjà connues qui sont Spt3, Spt7, Spt8, Spt20,

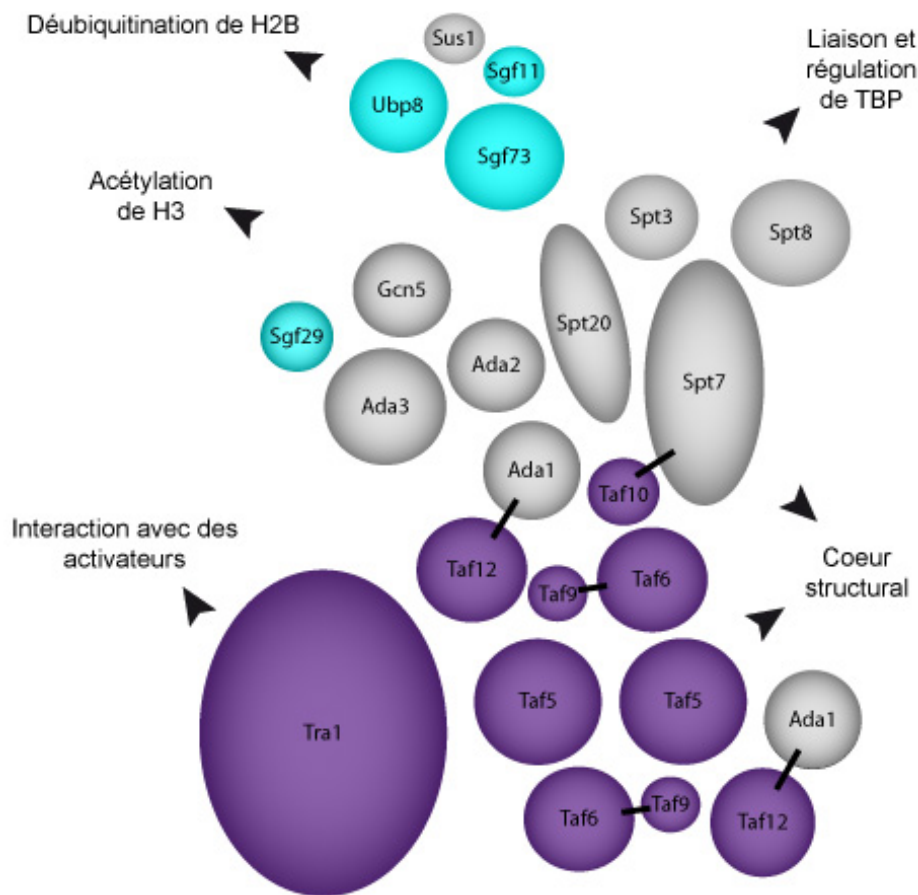
Ada1, Ada2, Ada3 et Gcn5, mais aussi un sous-ensemble de protéines TAF (TBP associated factors) à savoir TAF5, TAF6, TAF9, TAF10 et TAF12 (Figure 8). De façon intéressante, ces protéines sont également des constituants du facteur général de la transcription TFIID. Les TAF présentes dans SAGA semblent jouer un rôle important pour la fonction du complexe. Grant et coll. montrent par exemple qu'en travaillant avec un mutant thermo-sensible de TAF12, l'intégrité du complexe SAGA ainsi que sa capacité à acétyler des nucléosomes sont perdues à une température non permissive (Grant et al, 1998b).

En analysant par un gel SDS PAGE coloré à l'argent le complexe SAGA purifié, J. Workman et coll. ont observé une bande correspondant à une protéine d'environ 400 kDa. Celle-ci a été nommée p400 puis identifiée par spectrométrie de masse comme étant la protéine Tra1 (Figure 8). Comme nous l'avons déjà dit, cette dernière est l'homologue de la protéine humaine TRRAP qui était connue à l'époque pour son implication dans la transformation oncogénique induite par c-Myc et E2F (McMahon et al, 1998). Nous avons également vu que Tra1 possède plusieurs sites d'interaction avec des activateurs qui sont utilisés pour recruter le complexe SAGA au niveau de certains promoteurs (cf. A. III. 2. et C. IV. 3.) (Grant et al, 1998a). La même année, une autre équipe a aussi démontré que Tra1 est une sous-unité de SAGA (Saleh et al, 1998).

#### Nouvelles sous-unités non reliées avec la transcription

Le travail de P. Weil et coll. a permis une avancée majeure pour la caractérisation de la composition de SAGA chez la levure *S. cerevisiae* en identifiant des sous-unités dont la fonction n'avait jamais été associée avec la transcription (Figure 8). A la base, le but de ce travail était d'identifier une liste exhaustive des partenaires spécifiques du complexe TFIID. Pour cela, P. Weil et ses collègues ont purifié de façon systématique le complexe TFIID par immunoprécipitation (IP) en utilisant des anticorps polyclonaux dirigés contre chacune des sous-unités du complexe. Ces purifications ont ensuite été analysées par une technique de spectrométrie de masse de type MudPit (multidimensional protein identification technology). De façon intéressante, certaines protéines ont été retrouvées en grande quantité dans les purifications provenant des TAF présents à la fois dans TFIID et SAGA (TAF5, -6, -9, -10 et -12), alors qu'elles n'ont pas été détectées dans les purifications réalisées à partir des autres TAF présentes uniquement dans TFIID. Comme attendu, on retrouve toutes les sous-unités déjà connues de SAGA comme les protéines Spt ou Ada ainsi que Tra1. Par contre, trois nouvelles protéines sont

## Partie 2 : SAGA, un complexe co-activateur de la transcription



**Figure 8. Identification de sous-unités du complexe SAGA par des approches de protéomique.**

Chaque sous-unité du complexe SAGA de la levure *S. cerevisiae* est représentée par un rond ou un oval dont la taille est à peu près proportionnelle à la masse de la protéine. Les sous-unités essentielles qui ont été identifiées par les premières approches de protéomique sont représentées en violet. Les sous-unités qui ont été identifiées au sein de SAGA par des études de protéomique complémentaires et dont la fonction en lien avec la transcription n'était pas connue sont indiquées en bleu. Les sous-unités qui n'ont pas été identifiées par les approches de protéomique (Sus1) ou celles qui avaient déjà été caractérisées au sein du complexe SAGA sont colorées en gris. La fonction des modules de SAGA définis dans la Figure 4 est indiquée derrière une tête de flèche.

également retrouvées : Sgf29 (SAGA associated factor 29 kDa), Sgf73 (SAGA associated factor 73 kDa), deux protéines de fonction inconnue et l'ubiquitine protéase Ubp8. Des expériences complémentaires d'IP confirment que ces trois protéines sont effectivement des sous-unités de SAGA (Figure 8) (Sanders et al, 2002). SAGA possède donc une ubiquitine protéase et donc probablement une activité enzymatique de déubiquitination comme nous allons le voir en détail plus loin (cf. B. II. 3.). Bien que les auteurs pensaient avoir réussi à identifier toutes les sous-unités de SAGA (Sanders et al, 2002), une réanalyse de leurs données a permis d'identifier une nouvelle protéine de 11kDa faisant également partie du complexe qu'ils ont nommé Sgf11. Pour confirmer que cette

protéine de 11 kDa est bien une sous-unité de SAGA, une souche dans laquelle la protéine est fusionnée à l'étiquette TAP (Tandem Affinity Purification) a été générée. La double purification d'affinité a confirmé l'association de Sgf11 avec l'ensemble des sous-unités de SAGA (Figure 8 et Tableau 1) (Powell et al, 2004).

## **2. Sus1 identifiée par un crible génétique**

La dernière sous-unité de SAGA à avoir été caractérisée est Sus1 (*SI gene Upstream of YSa1*) mais cette protéine n'a pas pu être identifiée par les approches de protéomiques décrites précédemment. En effet, Sus1 est une protéine d'une centaine d'acides aminés, codée par un gène possède deux introns ce qui est très rare chez *S. cerevisiae* et ce qui explique pourquoi ce gène n'avait pas été annoté dans le génome de cet organisme. C'est la recherche de gènes pouvant provoquer une létalité synthétique avec Yra1, un facteur essentiel de la machinerie d'export des ARNm (Kohler & Hurt, 2007; Stewart, 2010) qui a mené à l'identification du gène *SUS1* (Rodriguez-Navarro et al, 2004). L'analyse par spectrométrie de masse des protéines associées à Sus1 a ensuite permis de démontrer que Sus1 est à la fois une sous-unité du complexe SAGA et une sous-unité de TREX2, un complexe impliqué dans l'export des ARNm (Rodriguez-Navarro et al, 2004).

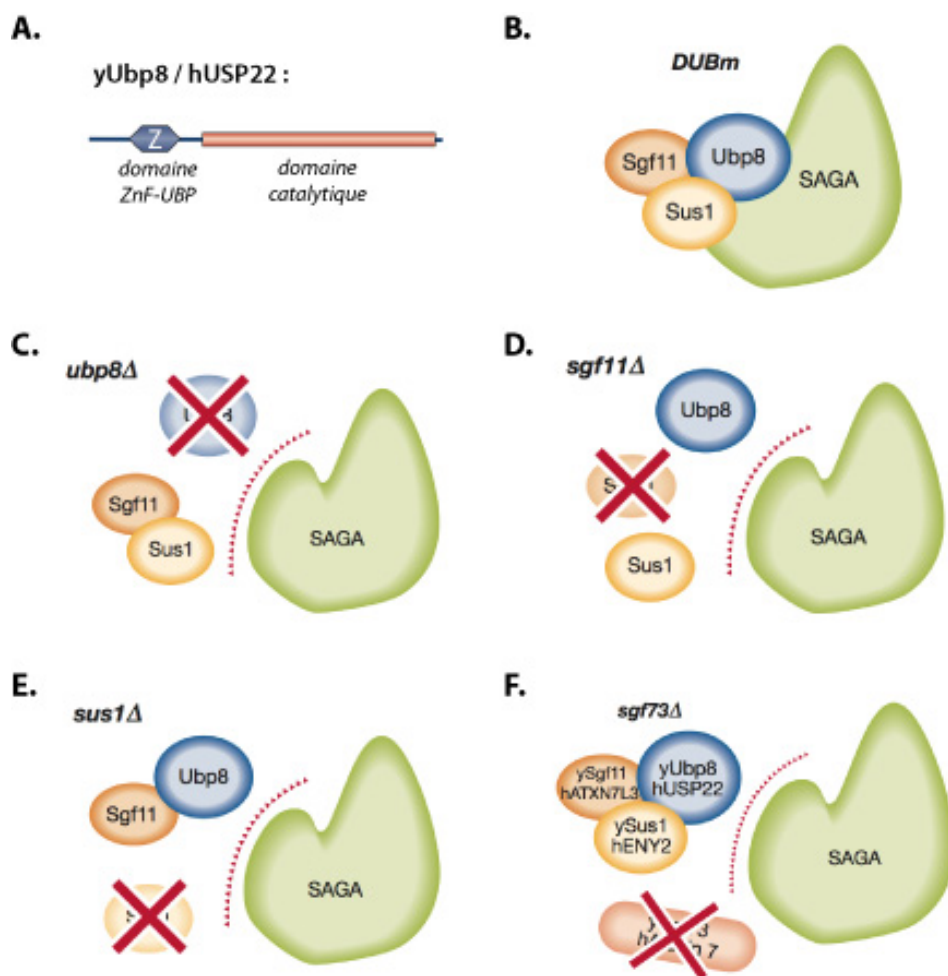
Je voudrais juste souligner que Sus1 est la seule sous-unité de SAGA à avoir d'abord été identifiée chez un métazoaire avant de l'être chez la levure. L'homologue drosophile de Sus1 s'appelle E(y)2 (enhancer of yellow phenotype) et a été identifié par un crible génétique (Georgiev & Gerasimova, 1989; Georgieva et al, 2001). De façon intéressante, le même crible avait permis l'identification du gène e(y)1 qui code la protéine TAF9, une protéine présente à la fois dans TFIID et dans SAGA (Soldatov et al, 1999). Une première caractérisation de la protéine E(y)2 a été réalisée avant que Sus1 ne soit identifiée chez la levure et a indiqué que E(y)2 et TAF9 étaient retrouvées au sein d'un même complexe (Georgieva et al, 2001).

## **3. Identification du module de déubiquitination**

L'analyse par spectrométrie de masse du complexe SAGA a permis de mettre en évidence de nouvelles sous-unités de SAGA (les TAF, Tra1 ainsi que Sgf11, Sgf29, Sgf73 et Ubp8) et de caractériser le dernier module structural et fonctionnel de SAGA. Il s'agit du module de déubiquitination (DUBm) sur lequel mon travail de thèse a porté. Nous allons décrire les travaux ayant montré que, chez la levure, l'activité de déubiquitination est localisée au sein d'un module du complexe SAGA puis présenter les propriétés

structurales et fonctionnelles de ce module (Figure 9 B.) (cf. Parties sur les USP et sur l'ubiquitination de H2B).

Plusieurs études sont arrivées à la conclusion que Sgf11 ne jouait pas de rôle dans l'assemblage du complexe mais qu'elle était nécessaire à l'incorporation de Ubp8 dans SAGA. En effet, à partir d'une souche *sgf11Δ*, seule Ubp8 est absente dans une purification du complexe SAGA alors que toutes les autres sous-unités sont retrouvées (Figure 9 D.) (Ingvarsdottir et al, 2005; Lee et al, 2005; Powell et al, 2004). L'affirmation réciproque est également vraie car dans une souche *ubp8Δ*, Sgf11 n'est pas associée à SAGA (Figure 9 C.) (Ingvarsdottir et al, 2005; Lee et al, 2005). La zone d'interaction entre



**Figure 9. Formation du module de déubiquitination de SAGA chez la levure *S. cerevisiae*.**

**A.** La déubiquitinase Ubp8 (ainsi que son homologue chez l'homme USP22) possède en plus de son domaine catalytique un domaine à doigts de zinc N-terminal appelé domaine ZnF-UBP (zinc finger ubiquitin-specific protease domain). **B.** Représentation schématique du module de déubiquitination (DUBm) de SAGA chez *S. cerevisiae*. **C.** Dans une souche *ubp8Δ*, Sgf11 et Sus1 ne sont pas intégrées à SAGA. **D.** L'absence de Sgf11 provoque la perte de Ubp8 et de Sus1 du complexe SAGA. **E.** La délétion de *SUS1* empêche Sgf11 et Ubp8 d'être associées à SAGA. **F.** Dans une souche dépourvue de Sgf73, l'ensemble des autres sous-unités du DUBm ne sont pas recrutées au sein de SAGA. La figure est adaptée de : (Rodriguez-Navarro, 2009)

Sgf11 et Ubp8 doit être localisée au niveau du domaine à doigt de zinc de Ubp8 appelé ZnF-UBP qui est situé dans la région N-terminale de la protéine en amont du domaine catalytique (Figure 9 A.). En effet, dans des souches de levure présentant des mutations qui affectent le repliement du ZnF-UBP, les deux protéines Ubp8 et Sgf11 ne sont pas intégrées au complexe SAGA (Ingvarsdottir et al, 2005). Ces observations montrent que les associations de Sgf11 et Ubp8 avec SAGA sont interdépendantes.

De façon très intéressante, l'absence de Ubp8 conduit à une augmentation globale du niveau de H2Bub chez la levure. Ceci est démontré en analysant par western blot avec un anticorps anti-Flag la quantité relative de H2B et H2Bub dans une souche exprimant une version fusionnée à l'étiquette Flag de H2B (Daniel et al, 2004; Henry et al, 2003; Ingvarsdottir et al, 2005; Lee et al, 2005). L'observation est d'autant plus marquante que l'absence de Sgf11 ou que des mutations dans le ZnF-UBP de Ubp8 conduisent au même phénotype (Ingvarsdottir et al, 2005; Lee et al, 2005). Cela indique donc que chez la levure, la protéine Ubp8 doit au moins pouvoir interagir avec Sgf11 pour exercer son activité. Dans un prochain chapitre, je décrirai le mécanisme de régulation de l'activité de Ubp8 dans les détails (cf. Partie sur les USP) (Ingvarsdottir et al, 2005).

L'identification de Sus1 a permis de mieux comprendre l'organisation structurale du DUBm. Le même type d'expérience que celles faites pour caractériser le lien entre Sgf11 et Ubp8 ont été réalisées avec Sus1. En l'absence de Sus1, Ubp8 n'est pas incorporée dans SAGA et là encore, l'affirmation réciproque est vérifiée (Figure 9 C. et E.) (Kohler et al, 2006). Pour mieux comprendre comment Sus1 interagit avec ses partenaires, Kohler et coll. ont utilisé une souche dans laquelle Sus1 est fusionnée à l'étiquette TAP. En utilisant cette souche, une purification par affinité permet d'isoler les complexes SAGA et TREX2 (Sus1 est une sous-unité de ces deux complexes : cf. A. I. et B. IV. 4.). Si la souche étiquetée présente une délétion du gène *UBP8*, seul le complexe TREX2 et Sgf11 restent associés à Sus1. Ceci démontre une interaction directe entre Sus1 et Sgf11 qui est indépendante de Ubp8 (Figure 9 C.) (Kohler et al, 2006). Enfin, une purification dans des conditions stringentes montre que Sus1 reste associée à TREX2 et à un sous-complexe de SAGA composé de Sus1, Sgf11, Ubp8 et Sgf73 (Kohler et al, 2006). Cette observation a permis de définir le module de déubiquitination (DUBm) de SAGA.

Différents arguments suggèrent que Sgf73 pourrait être la sous-unité qui fait le lien entre le DUBm et le reste de SAGA. La purification de SAGA dans une souche *sgf73Δ*

indique la perte spécifique de Ubp8, Sgf11 et Sus1 alors que la composition du reste du complexe n'est pas altérée (Figure 9 F.) (Köhler et al, 2008). D'autres expériences du même type ainsi que des analyses par spectrométrie de masse montrent que la région N-terminale de Sgf73 est nécessaire pour permettre l'intégration des autres sous-unités du DUBm au sein de SAGA (Köhler et al, 2008; Lee et al, 2009). Un deuxième type d'expérience démontre que l'absence de Sgf73 conduit à une augmentation globale du niveau de H2Bub similaire à celle observée en l'absence de Ubp8 ou Sgf11.

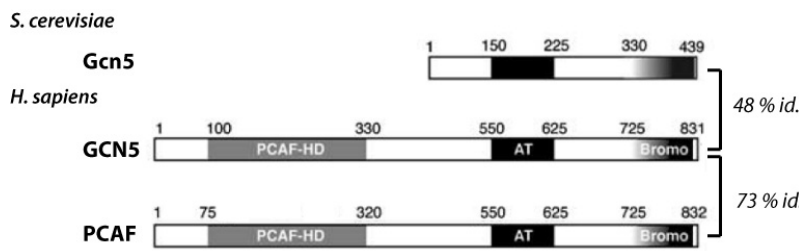
### **III. Identification et caractérisation de SAGA chez l'Homme**

#### **1. Purification de complexes homologues**

Chez l'homme, plusieurs approches indépendantes ont permis de purifier l'homologue du complexe SAGA de levure. Je vais présenter brièvement ces diverses stratégies qui ont conduit à la définition de trois complexes. Dans un premier temps, ceux-ci ont été appelés de façons différentes mais nous verrons que ces trois complexes sont en fait identiques. Pour plus de simplicité, j'utiliserai la nomenclature actuelle pour décrire les sous-unités de ces complexes.

#### Le complexe PCAF

Comme son nom l'indique, la protéine PCAF (p300/CBP Associated Factor) a été caractérisée comme interagissant avec le co-activateur de la transcription p300/CBP. PCAF est constituée par deux régions fonctionnelles. La région C-terminale présente une très forte similarité de séquence avec la protéine de levure Gcn5. PCAF possède d'ailleurs une activité intrinsèque d'acétylation des histones. La région N-terminale, appelée « PCAF homology domain » est, quant à elle, spécifique de PCAF. Il a été suggéré que cette région permette l'interaction avec p300 et CBP (Yang et al, 1996). Même si une telle interaction physique entre PCAF et p300/CBP n'a jamais été étudiée plus en avant, le nom de la protéine est restée. Chez l'homme, une autre protéine présente une très forte similarité de séquence avec PCAF et yGcn5 et a été appelée GCN5 (Figure 10). Lors du clonage du gène codant GCN5, il a été prédit que la taille de cette protéine était similaire à celle de la protéine de levure Gcn5. On sait aujourd'hui que ce n'est pas le cas et que GCN5 et PCAF ont la même structure primaire avec une extension N-terminale puis la région homologue à Gcn5 de levure qui contient un domaine acétyl-transférase et un Bromo-domaine (Figure 10).



**Figure 10. GCN5 et PCAF sont des protéines homologues.**

Chez les vertébrés, GCN5 possède une protéine paralogue appelée PCAF (p300/CBP Associated Factor). Les deux protéines humaines GCN5 et PCAF présentent une forte identité de séquence (73%) ainsi qu'une même organisation en domaines avec une longue région N-terminale (PCAF homology domain) suivie par un domaine acétyl-transférase et un Bromo-domaine. Seuls le domaine acétyl-transférase et le Bromo-domaine sont présents chez Gcn5, l'orthologue de GCN5 et PCAF chez la levure *S. cerevisiae*.

Pour mieux comprendre la fonction de PCAF et de GCN5, l'équipe de Y. Nakatani a déterminé quelles protéines interagissent avec ces deux enzymes d'acétylation des histones. Pour cela, deux lignées cellulaires exprimant une version étiquetée de l'une ou

de l'autre des protéines ont été générées. Après une double purification d'affinité, le contenu en protéines de ces purifications a été analysé par un gel SDS-PAGE coloré à l'argent. Plus d'une vingtaine de bandes y ont été observées et analysées par spectrométrie de masse. Les résultats obtenus avec PCAF ou avec la forme courte de GCN5 sont très similaires. Plusieurs homologues de protéines connues pour interagir avec Gcn5 chez la levure ont été mis en évidence. C'est le cas notamment pour Spt3, Ada2 et Ada3. On peut noter que cette étude a été publiée en même temps que celle de J. Workman qui montre qu'un sous-ensemble de TAF sont des sous-unités de SAGA de levure. Chez l'homme, les protéines TAF9, TAF10 et TAF12, en plus d'être des sous-unités du complexe TFIID, sont également des partenaires de GCN5 ou PCAF. De façon intéressante, chez l'homme, TAF5 et TAF6 sont spécifiques de TFIID alors que deux protéines très similaires à TAF5 et TAF6, appelées TAF5L et TAF6L sont spécifiquement associées avec GCN5 ou PCAF. Les auteurs de l'étude démontrent que comme TAF6, TAF6L peut interagir directement avec TAF9 *in vitro* par son domaine Histone Fold. Enfin, il est montré que les complexes GCN5 ou PCAF peuvent acétyler l'histone H3 dans le contexte du nucléosome ce que ne peut pas faire l'enzyme PCAF seule (Ogryzko et al, 1998).

Nous verrons dans un prochain paragraphe (cf. B. IV. 4.) que chez l'homme, on retrouve GCN5 et PCAF dans le complexes SAGA mais également dans un autre complexe HAT, le complexe ATAC (Ada Two A Containing complex). On sait donc aujourd'hui que la purification des protéines associées à PCAF ou à GCN5 devrait donc permettre d'isoler



ces deux complexes. En analysant cet article (Ogryzko et al, 1998), on pourrait penser que seules des sous-unités de SAGA ont été identifiées, ce qui n'est pas le cas. En effet, l'homologue humain de Ada2 qui a été identifié dans cette étude correspond à ADA2a, une sous-unité spécifique de ATAC, qui est absente de SAGA (Ogryzko et al, 1998). Dans cette étude, la caractérisation des partenaires de la protéine PCAF entière ou de la région C-terminale de GCN5 a conduit à des résultats très similaires. Dans les deux cas, les mêmes sous-unités de SAGA ont été observées. Ceci indique que la région N-terminale de GCN5 n'est pas indispensable pour l'interaction de l'enzyme avec le complexe SAGA. Il serait intéressant de déterminer si cette région joue par contre, un rôle structural dans l'assemblage du complexe ATAC.

D'autre part, il peut sembler surprenant, vu la qualité des expériences réalisées, que l'ensemble des sous-unités de ces deux complexes n'ait pas été caractérisé de suite. Cela est sans doute dû au fait qu'il n'était pas trivial de définir si une protéine identifiée par spectrométrie de masse est une sous-unité spécifique d'un complexe. Cela explique également pourquoi dans un premier temps, on n'a retrouvé au sein des complexes GCN5 et PCAF que des protéines déjà caractérisées comme jouant un rôle dans la transcription. L'exemple des protéines TAF permet bien d'illustrer cette situation. Tous les résultats présentés par les auteurs concernant ce groupe de protéines sont justes et complets (Ogryzko et al, 1998).

#### Le complexe STAGA

L'équipe de R. Roeder s'est intéressée à l'orthologue humain de Spt3 qui, chez la levure, interagit avec TBP (Eisenmann et al, 1992). Cependant, chez l'homme SPT3 n'était pas immunoprécipité par des anticorps anti-TAF4, TAF5 ou TBP. Par contre, la protéine SPT3 est co-purifiée par un anticorps dirigé contre TAF9 indiquant que SPT3 et TAF9 font partie d'un complexe différent de TFIID. Il est aussi montré que ce complexe contient une activité d'acétylation sur les histones H3 libres. Celle-ci est due à la protéine GCN5 qui est immunoprécipitée dans sa forme longue par des anticorps anti-SPT3 ou TAF9. Ce complexe a donc été appelé STAGA (SPT3 TAF9 GCN5 Acetyltransferase) (Martinez et al, 1998). La même équipe a affiné la caractérisation du complexe en générant deux lignées de cellules HeLa qui expriment des versions étiquetées (épitopes HA et Flag) de TAF9 et de SPT3. Après une double purification d'affinité, la composition des complexes a été analysée par spectrométrie de masse. Dans la purification réalisée à partir de TAF9, deux complexes sont purifiés (TFIID et

STAGA) alors que la purification à partir de SPT3, permet d'isoler spécifiquement le complexe STAGA. Des protéines présentes dans les deux types de purification et absentes dans le contrôle ont pu donc être considérées comme des sous-unités de STAGA. Grâce à cette méthode, il a été démontré que TRRAP, ADA3, SGF29, TAF5L, TAF6L, ADA1, et SPT7L sont de nouveaux constituants du complexe (Martinez et al, 2001). Les auteurs suggèrent enfin que la protéine SAP130, un constituant du facteur d'épissage SF3b, pourrait interagir de manière transitoire avec le complexe STAGA (Martinez et al, 2001).

#### Le complexe TFTC

L'équipe de L. Tora travaillant sur les mécanismes d'initiation de la transcription et en particulier sur les protéines TAF a entrepris d'identifier un complexe contenant des TAF mais pas la protéine TBP. Pour démontrer cela, deux IP successives ont été réalisées, à partir d'extraits nucléaires de cellules HeLa. Tout d'abord un anticorps dirigé contre TAF10 permet de retenir les complexes contenant cette protéine. Après une élution peptidique de cette première IP, l'utilisation d'un anticorps anti-TBP permet de séparer TFIID d'autres complexes qui ne possèderaient pas TBP. L'analyse de la fraction dépourvue du complexe TFIID par western blot montre que celle-ci contenait des protéines TAF mais pas de TBP. Ceci démontre donc bien l'existence d'un (ou de plusieurs) complexe(s) contenant des TAF en l'absence de TBP. Ce complexe a donc été appelé TFTC (TBP Free TAF Containing complex) et il contient notamment TAF10 et TAF12 (Wieczorek et al, 1998). Dans un second temps, il a été montré que TFTC contient la protéine GCN5 dans sa forme longue. Ceci confère au complexe au complexe TFTC une activité d'acétylation des histones. L'histone H3 incorporé dans un nucléosome peut être acétylé par TFTC de même que l'histone H1 alors que GCN5 seule ne possède pas d'activité sur un nucléosome. D'autres sous-unités ont également été identifiées : SPT3, ADA3, TAF5L et TRRAP (Brand et al, 1999b).

On sait aujourd'hui qu'au sein de cette purification TFTC, on retrouve SAGA mais également un ensemble d'autres complexes contenant des TAF qui pourraient être des modules d'assemblage du complexe TFIID ou des variantes de TFIID pouvant participer à l'initiation de la transcription *in vivo*. On peut d'ailleurs noter qu'il a été montré que la fraction TFTC peut initier la transcription *in vitro* à partir d'une matrice d'ADN nu contenant ou non une boîte TATA. Ce sont sans doute ces sous-complexes de TAF et non SAGA qui sont responsables de cette activité (Wieczorek et al, 1998).

### Bilan et commentaire

Ces trois stratégies expérimentales ont permis de purifier les complexes PCAF/GCN5, STAGA et TFTC. On sait aujourd'hui avec du recul que les différences observées dans la composition en sous-unités de ces complexes sont dues à des limitations techniques. Nous verrons dans les prochains paragraphes que la composition du complexe SAGA de levure et celle de son homologue humain sont presque identiques. Pour plus de simplicité, nous parlerons donc du complexe SAGA quelque soit l'espèce dans lequel il a été étudié (Pijnappel & Timmers, 2008).

### **2. Identification de SGF29 et de SPT20**

Chez le rat, SGF29 a été identifiée comme étant une sous-unité du complexe SAGA et un partenaire direct de ADA3. Dans un test de double hybride, les auteurs démontrent que le domaine coiled-coil de la région N-terminale de SGF29 interagit avec la région N-terminale de ADA3. Dans la même étude, il est montré que SGF29 est requis pour permettre l'expression optimale de gènes régulés par c-Myc. Enfin, les auteurs suggèrent une implication de la protéine SGF29 dans la transformation oncogénique réalisée par c-Myc (Kurabe et al, 2007).

A partir d'une purification de type TFTC, L. Tora et coll. ont tenté d'identifier de nouvelles sous-unités de SAGA. Une analyse par spectrométrie de masse a permis de mettre en évidence des peptides correspondant à une nouvelle sous-unité potentielle, la protéine p38IP, identifiée initialement sur la base d'une interaction avec la kinase p38 MAPK (mitogen-activated protein kinase). Des analyses bioinformatiques ont permis de mettre en évidence une similarité de séquence entre p38IP et la protéine de levure Spt20. Ceci suggère que p38IP puisse être l'homologue humain de Spt20. Plusieurs évidences biochimiques dont des expériences de co-IP ont montré que la protéine p38IP est bien une sous-unité de SAGA. En effet, un anticorps dirigé contre p38IP permet de purifier les autres sous-unités de SAGA et dans l'expérience réciproque, si on analyse une purification de SAGA par western blot, une bande est révélée par un anticorps anti-p38IP. Cette protéine a donc été renommée SPT20. Les résultats de l'étude suggèrent également que SPT20 joue un rôle important dans l'architecture du complexe. Enfin, l'équipe de L. Tora montre que SAGA régule une classe de gènes induits par un stress du réticulum endoplasmique et qu'en l'absence de SPT20, l'induction optimale de ces gènes est perdue (Nagy et al, 2009). Une autre étude a également démontré que SPT20 était une sous-unité de SAGA (Wang et al, 2008).

### 3. Identification de ATXN7

Avant que Sgf73 soit identifiée comme une sous-unité de SAGA chez la levure, des recherches étaient déjà menées sur son possible homologue humain, la protéine Ataxin-7 (ATXN7). En effet, une expansion de triplets CAG dans la phase codante du gène *ATXN7* est la cause d'une maladie neurodégénérative autosomale dominante appelée l'ataxie spinocérébelleuse de type 7 (SCA7). Avant de décrire les expériences ayant démontrées la présence de ATXN7 au sein du complexe SAGA humain, je voudrais décrire brièvement ce qui est connu de la maladie SCA7.

#### SCA7 est une maladie à expansion polyglutamine

Chez les patients SCA7, on observe une expansion d'un motif polyglutamine à l'extrémité N-terminale de ATXN7. On compte 8 autres pathologies, appelées maladies par expansion polyglutamine, dues au même type de mutation. Parmi elles, on retrouve la maladie de Huntington ainsi que d'autres formes d'ataxies spinocérébelleuses. Ces maladies à expansion polyglutamine présentent des caractéristiques communes : il s'agit de maladies neurodégénératives progressives débutant chez l'adulte. Le seuil pathologique apparaît similaire, avec des allèles mutés portant plus de 40 triplets CAG environ, et la sévérité de la maladie est corrélée à la taille de l'expansion. Il a été clairement démontré grâce à des analyses moléculaires et génétiques que la mutation par expansion polyglutamine confère un gain de fonction à la protéine mutante. Malgré une expression ubiquitaire dans tout l'organisme des protéines mutées, seule une population particulière de neurones est sensible à chaque protéine mutée. Par exemple, SCA7 est la seule maladie par expansion polyglutamine dans laquelle les patients présentent une atteinte de la rétine. Les bases moléculaires de cette spécificité de neurodégénérescence sont encore mal comprises. D'une manière plus générale, les mécanismes physiopathologiques de SCA7 ne sont pas clairement établis. Il apparaît cependant qu'en plus de la toxicité propre des expansions polyglutamine, la fonction normale de la protéine mutée pourrait jouer un rôle dans ces mécanismes physiopathologiques (Gatchel & Zoghbi, 2005).

#### ATXN7 est une sous-unité de SAGA

La fonction normale de ATXN7 a notamment été étudiée par Helmlinger et coll. Différentes approches biochimiques ont permis de démontrer que ATXN7 est une sous-unité de SAGA. En particulier, des expériences de co-IP sur des extraits nucléaires de

cellules HeLa avec un anticorps dirigé contre ATXN7 ont permis de montrer que différentes sous-unités de SAGA sont associées à ATXN7, comme GCN5, TRRAP ou TAF10. De plus, il a été montré que la fraction immunoprécipitée par l'anticorps anti-ATXN7 contenait une activité d'acétylation (Helmlinger et al, 2004b).

#### Altérations transcriptionnelles et chromatinienne

Pour étudier les conséquences de l'expansion Polyglutamine dans ATXN7, deux lignées de souris transgéniques ont été générées, l'une (lignée R7N) exprimant une forme normale de ATXN7 avec 10Q et l'autre (lignée R7E), une forme mutée de la protéine avec 90Q. Dans les deux cas, la phase codante de ATXN7 est placée sous la dépendance du promoteur de la rhodopsine ce qui permet une expression forte et spécifique dans les photorécepteurs de type bâtonnet. Le seul phénotype observé va donc concerner la rétine. Des anomalies de la transcription ont été mises en évidence dans le modèle R7E. ATXN7 mutée induit une diminution rapide et très forte de l'expression de la plupart des gènes spécifiques des bâtonnets qui sont normalement très fortement exprimés. Les gènes les plus touchés sont ceux qui codent pour des protéines impliquées dans le processus de phototransduction. Une analyse temporelle a montré que la baisse de l'expression de la rhodopsine allait de pair avec une diminution des réponses électrophysiologiques des bâtonnets et le raccourcissement de leurs segments. Le premier événement pathologique détecté semblerait donc être une anomalie transcriptionnelle qui expliquerait la perte de fonction des photorécepteurs des souris R7E (Helmlinger et al, 2004a).

D'autre part, l'analyse par microscopie électronique de l'ultrastructure des bâtonnets révèle une différence majeure de l'organisation nucléaire des bâtonnets dans des rétines de souris R7E par rapport aux contrôles. Chez les souris sauvages, les noyaux des bâtonnets ont une organisation particulière et caractéristique de ce type cellulaire, avec un liseré d'euchromatine périphérique entourant une large zone d'hétérochromatine centrale. Dans les rétines des souris R7E, on observe une modification très importante de cette structure chromatinienne avec une décondensation majeure de la chromatine, associée à une forte augmentation du volume du noyau. Ces modifications sont précoces et leur progression est directement corrélée à la répression des gènes spécifiques des photorécepteurs.

Ces altérations transcriptionnelles et chromatinienne suggèrent donc que la fonction de SAGA puisse être en partie altérée et donc impliquée dans les mécanismes

physiopathologiques de SCA7. L'analyse des souris R7E et contrôles montre que l'expansion polyglutamine au sein de ATXN7 n'altère pas le recrutement de GCN5 ou de TAF12 dans le complexe SAGA. De plus, l'activité d'acétylation des histones ou des nucléosomes des complexes purifiés n'est pas affectée par l'expansion polyglutamine indiquant que dans la rétine des souris R7E le complexe SAGA garde son activité d'acétylation. Par contre des expériences de ChIP suggèrent une modification du recrutement de SAGA et une acétylation anormale sur certains gènes (Helmlinger et al, 2006). Cependant, d'autres travaux utilisant la levure ou des lignées cellulaires, suggèrent au contraire que l'expansion polyglutamine dans ATXN7 inhibe l'activité HAT du complexe SAGA (McMahon et al, 2005; Palhan et al, 2005).

Mon travail de thèse a permis de montrer que le domaine N-terminal de ATXN7 participe à la formation du DUBm ainsi qu'à la régulation enzymatique de USP22 (cf. Partie Résultats). La région de ATXN7 qui contient l'expansion polyglutamine étant située à proximité de ce domaine, il serait intéressant de déterminer l'effet de l'expansion polyglutamine sur la formation du DUBm ainsi que ses éventuels effets sur l'activité DUB de SAGA *in vitro* et *in vivo*.

#### **4. Identification d'un module de déubiquitination**

Nous avons vu que chez la levure, une activité enzymatique de déubiquitination a été caractérisée au sein du complexe SAGA. Cette activité est portée par la protéine Ubp8 qui interagit fortement avec Sgf11, Sgf73 et Sus1 pour former un module de déubiquitination des histones. Chez l'homme, deux études ont permis de mettre en évidence une telle activité de déubiquitination dans le complexe SAGA.

L'équipe de S. McMahon s'est intéressée à la protéase USP22 car elle a été identifiée comme faisant partie d'une signature de 11 gènes qui sont surexprimés dans un certain nombre de cancers de mauvais pronostic (Glinsky et al, 2005). Parmi ces gènes, on retrouve BMI-1 et RNF2/Ring1B qui sont des sous-unités du complexe PRC1 (Polycomb Repressive Complex 1) et qui catalysent l'ubiquitination de H2A. Comme une similarité de séquence a été observée avec Ubp8, les auteurs ont émis l'hypothèse que USP22 était une sous-unité du complexe SAGA chez l'homme. Pour le démontrer, différentes approches biochimiques ont été employées. Par exemple, une lignée cellulaire exprimant USP22 fusionnée à l'étiquette Flag a été générée. Le contenu en protéines d'une IP anti-Flag a été analysé par spectrométrie de masse et par western blot. Plusieurs sous-unités spécifiques de SAGA ont été trouvées comme TRRAP et GCN5.

De plus, il a été montré que USP22 est co-purifiée par GCN5. Enfin, une activité d'acétylation des histones est associée au complexe purifié à partir de USP22. Les auteurs montrent également que le complexe SAGA purifié à partir de USP22 ou à partir de GCN5 possède une activité de déubiquitination de l'histone H2B monoubiquitiné *in vitro* (Zhang et al, 2008). Dans la partie résultats, nous parlerons de la caractérisation de l'activité DUB de SAGA *in vivo*.

Une autre approche a été utilisée dans l'équipe de L. Tora qui avait déjà caractérisé ATXN7 comme l'homologue de Sgf73, faisant partie du complexe SAGA humain (cf. B. III. 3.) et (Helmlinger et al, 2004b). Pour identifier de potentiels homologues humains des autres sous-unités du DUBm de levure, des analyses bioinformatiques ont été réalisées. Une similarité de séquence a été observée entre USP22, ATXN7L3 et ENY2 d'une part et Ubp8, Sgf11 et Sus1 d'autre part. Pour démontrer que ces trois protéines sont des sous-unités de SAGA chez l'homme, des anticorps spécifiques de chacune ont été générés. Ceux-ci ont notamment été utilisés pour des expériences d'IP qui ont permis de co-purifier plusieurs sous-unités de SAGA comme TRRAP, GCN5 ou TAF10. Des expériences réciproques ont permis de montrer que USP22 et ATXN7L3 sont présentes dans des purifications de SAGA. A partir de ces observations, les auteurs ont pu conclure que ces trois protéines sont des sous-unités du complexe SAGA. Des expériences visant à déterminer si ces nouvelles sous-unités de SAGA forment un module structural ont aussi été menées. Pour cela, différentes combinaisons de cDNA de sous-unités du DUBm ont été transfectées dans des cellules HEK293T. Des IP sur ces protéines recombinantes ont ensuite permis de démontrer que ATXN7L3, ENY2 et USP22 peuvent s'assembler dans la cellule. De plus, une interaction directe entre ATXN7L3 et ENY2 d'une part et entre ATXN7L3 et USP22 d'autre part a été observée. Dans une autre approche, des cellules HeLa ont été traitées avec un ARN interférant pour réduire drastiquement la quantité de ATXN7L3 dans la cellule. Une IP anti-GCN5 réalisée sur un extrait protéique provenant de ces cellules a ensuite permis de montrer que USP22 est perdue de SAGA en l'absence de ATXN7L3. Enfin, des expériences *in vitro*, à partir de protéines recombinantes suggèrent une interaction entre ATXN7 et les autres sous-unités du DUBm. Ceci montre que chez l'homme, ATXN7L3, ENY2 et USP22 font partie d'un module structural et que ATXN7L3 joue un rôle crucial dans la formation de celui-ci. De plus, ces résultats indiquent que ATXN7

pourrait être impliqué dans le recrutement du module au sein de SAGA (Zhao et al, 2008).

Une partie de mon travail de thèse a consisté à déchiffrer le réseau d'interactions entre les sous-unités du DUBm pour comprendre l'importance fonctionnelle de ces interactions et nous reviendrons sur ce point dans la partie Résultats.

## **IV. Evolution du complexe SAGA**

### **1. Conservation au cours de l'évolution**

Chez *S. pombe*, l'analyse par spectrométrie de masse d'un complexe SAGA purifié par une double purification d'affinité de type TAP a montré que la composition du complexe SAGA est identique à celle de *S. cerevisiae* (Helmlinger et al, 2008).

Plusieurs études ont été réalisées afin de déterminer la composition en sous-unité du complexe SAGA chez la mouche *Drosophila melanogaster* (Guelman et al, 2006; Kurshakova et al, 2007; Kusch et al, 2003; Muratoglu et al, 2003; Weake et al, 2008; Weake et al, 2009). Une analyse par spectrométrie de masse en utilisant la technique de MudPit a permis une caractérisation complète de SAGA dans cet organisme (Weake et al, 2009). Une protéine homologue de chacune des sous-unités du complexe SAGA de levure a pu être identifiée, à l'exception de Spt8. On peut noter que l'identification de l'orthologue de Sgf73/ATXN7 n'a pas été aisée dans cet organisme. En effet, au sein de cette famille de protéines, seul le domaine central appelé domaine SCA7 est suffisamment conservé pour être utilisé pour l'identification d'homologues par des analyses bioinformatiques de type BLAST (cf. Partie Résultats) et (Helmlinger et al, 2004b). Cependant, bien que ce domaine SCA7 soit très conservé chez tous les autres homologues de Sgf73/ATXN7 connus, il n'est pas retrouvé au sein de l'orthologue de drosophile comme l'a indiqué l'identification par MudPit de cet orthologue (cf. Partie Résultats) et (Weake et al, 2009).

Enfin, une étude récente a permis de montrer que la composition du complexe SAGA dans les cellules ES de souris (Embryonic Stem cells) était identique à celle du complexe SAGA chez l'homme (Krebs et al, 2010).

### **2. Lien avec TBP chez les métazoaires ?**

En terme de composition en sous-unité, la seule différence majeure entre le complexe SAGA chez les levures et les métazoaires concerne l'absence d'un homologue



de Spt8 chez ces derniers. Cette différence pourrait paraître minime et sans importance, mais je ne pense pas que ce soit le cas. En effet, elle pose la question de savoir si la protéine TBP est également régulée par SAGA chez les eucaryotes supérieurs. Chez la levure, nous avons vu qu'il existe une interaction génétique entre *SPT3* et *SPT15*, le gène codant TBP (Eisenmann et al, 1992; Eisenmann et al, 1994; Laprade et al, 2007). Une autre étude a démontré une interaction directe entre Spt3 et TBP *in vitro* et *in vivo* et suggère que TBP est principalement régulée par Spt3 mais que Spt8 participe aussi à ce processus (Mohibullah & Hahn, 2008). D'ailleurs, d'autres études ont montré une interaction directe entre Spt8 et TBP *in vitro* (Sermwittayawong & Tan, 2006; Sterner et al, 1999; Warfield et al, 2004) et il semble donc que cette protéine joue aussi un rôle dans la déposition de TBP sur certains promoteurs. Dans ce cas, l'absence d'homologue de Spt8 dans un complexe SAGA chez les eucaryotes supérieurs empêche-t-elle SAGA de réguler TBP ? On sait par exemple que TBP n'est pas détectée dans des purifications de SAGA chez les métazoaires (Martinez et al, 1998; Weake et al, 2009; Wieczorek et al, 1998) alors TBP peut être observée dans des purifications du complexe SAGA chez la levure (Mohibullah & Hahn, 2008). De plus, TBP semble être régulée de façon différente chez les mammifères, où la protéine est uniquement retrouvée au sein des complexes TFIID et B-TFIID (Choukrallah et al, 2011) et la levure, où en plus d'être une sous-unité de TFIID, elle peut exister sous une forme libre et interagir de façon transitoire avec des complexes comme SAGA, NC2 ou la protéine Mot1 (Spedale et al, 2012; Venters & Pugh, 2009). Cependant, chez les eucaryotes supérieurs, SAGA possède un homologue de Spt3 et on ne peut donc pas exclure une implication de cette protéine dans la régulation de la dynamique de TBP. Pour apporter une réponse à cette question, il sera sans doute possible d'analyser les conséquences d'une délétion de *SPT3* chez la drosophile ou chez la souris pour déterminer si SAGA a perdu ou non sa fonction de régulation de TBP au cours de l'évolution et s'il ne conserve que ses fonctions de modification de la chromatine chez les métazoaires.

### **3. Duplication de gènes codant des sous-unités de SAGA**

Au cours de l'évolution, des duplications de gènes codants des sous-unités de SAGA ont été observées. Une partie de ces duplications concerne des protéines qui sont partagées avec d'autres complexes chez *S. cerevisiae*. Par exemple, deux gènes homologues au gène *TRA1* de *S. cerevisiae*, *TRA1* et *TRA2* sont retrouvés chez *S. pombe*, le premier codant une sous-unité spécifique de SAGA et le second une sous-unité de NuA4 (Helmlinger et al,

2008; Helmlinger et al, 2011; Shevchenko et al, 2008). Cette duplication n'est pas retrouvée chez d'autres organismes, mais d'autres duplications de gènes sont plus largement conservées. Chez les métazoaires, il existe en effet des paralogues de TAF5 et de TAF6 qui sont spécifiques de SAGA et appelés respectivement WDA (Will Decrease Acetylation) et SAF6 (SAGA associated factor 6) chez la drosophile (Guelman et al, 2006; Weake et al, 2009) ou TAF5L et TAF6L chez les mammifères (Krebs et al, 2010; Nagy et al, 2009; Ogryzko et al, 1998).

D'autre part, chez les vertébrés, certaines sous-unités de SAGA sont codées par plusieurs gènes paralogues. C'est le cas notamment des enzymes HAT GCN5 et PCAF, de ATXN7, ATXN7L1 et ATXN7L2 et enfin de TAF9 et de TAF9b. Bien qu'une fonction spécifique ne soit pas associée à l'ensemble de ces paralogues, il semblerait que au moins pour certains d'entre eux, la fonction de ces différents paralogues ne soit pas totalement redondante. Par exemple, des souris *Gcn5* *-/-* ne sont pas viables alors que des souris *Pcaf* *-/-* sont viables et ne présentent que des problèmes neurologiques comme des troubles de la mémoire (Maurice et al, 2008). Nous avons également montré que la proportion d'ARNm codant les différents paralogues ATXN7, -L1 et -L2 variait d'un tissu à l'autre, chez la souris. Il est donc possible que plusieurs paralogues existent pour permettre à une cellule de réguler la quantité de certaines sous-unités de SAGA de façon spécifique. Enfin, il a été montré que TAF9 et TAF9b étaient à la fois des sous-unités du complexe SAGA et du complexe TFIID et des analyses transcriptionnelles réalisées dans des cellules traitées par un ARN interférent dirigé contre TAF9 ou TAF9b ont indiqué que ces deux protéines ne régulaient pas le même groupe de gènes (Frontini et al, 2005).

#### **4. Autres complexes contenant des sous-unités de SAGA**

De nombreuses sous-unités de SAGA sont également retrouvées dans d'autres complexes (Tableau 1), et la distribution de certaines de ces protéines entre plusieurs complexes est largement conservée au cours de l'évolution. C'est notamment le cas de la sous-unité Sus1/ENY2 du DUBm de SAGA qui a également été identifiée au sein du complexe TREX2 chez la levure, la drosophile et l'homme (Jani et al, 2012; Kurshakova et al, 2007; Rodriguez-Navarro et al, 2004). Ce complexe est impliqué dans l'export des ARNm et il est localisé au niveau des pores nucléaires. Pour des revues sur TREX2 voir : (Garcia-Oliver et al, 2011; Rondon et al, 2010). D'autre part, un sous-ensemble de protéines TAF (TAF5, -6, -9, -10, -12) est partagé entre SAGA et le facteur général de la

transcription TFIID (cf. Partie sur les mécanismes d'initiation de la transcription). TAF5 et TAF6 sont seulement trouvées à la fois dans SAGA et TFIID chez les levures (cf. B. IV. 3.), mais par contre, TAF9, TAF10 et TAF12 sont présentes dans les deux complexes chez l'ensemble des eucaryotes. Enfin, à l'exception de *S. pombe* qui possède deux versions de Tra1 (cf B. IV. 3.), Tra1/TRRAP est présent dans deux complexes d'acétylation des histones ; le complexe SAGA et le complexe NuA4. Il a été montré que NuA4 acétylait la queue N-terminale de l'histone H4 et qu'il participait à la régulation de la transcription, de la progression du cycle cellulaire et de la réparation de l'ADN (Doyon & Cote, 2004).

Par ailleurs, certaines sous-unités de SAGA sont retrouvées dans des complexes qui sont spécifiques à des espèces de levures ou aux métazoaires. Chez la levure *S. cerevisiae*, il existe notamment un complexe appelé SALSA (SAGA altered Spt8 absent) ou SLIK (SAGA-like) qui ressemble énormément à SAGA mais qui est caractérisé par l'absence de Spt8, consécutive à un clivage protéolytique de l'extrémité C-terminale de Spt7 (Pray-Grant et al, 2002; Spedale et al, 2010; Sterner et al, 2002a). Bien que des différences fonctionnelles aient pu être observées entre SAGA et SLIK, pour le moment, aucune fonction générale et spécifique n'a été clairement associée au complexe SLIK (Spedale et al, 2010). D'autre part, chez *S. cerevisiae*, il a été montré que les sous-unités du HATm (Ada2, Ada3, Sgf29 et Gcn5) faisaient également partie d'un complexe appelé Ada qui contient deux autres sous-unités spécifiques, et qui peut acétyler l'histone H3 dans le contexte d'un nucléosome (Grant et al, 1997; Lee et al, 2011a).

Enfin, il faut noter que chez les eucaryotes supérieurs, les sous-unités du HATm de SAGA (ADA3, SGF29 et GCN5) sont également retrouvées au sein d'un complexe co-activateur de la transcription appelé ATAC (Ada Two A Containing). Ce complexe tire son nom du fait que chez les eucaryotes supérieurs, il existe deux orthologues de la protéine de levure Ada2, ADA2a et ADA2b qui sont respectivement spécifiques de ATAC et de SAGA (Nagy & Tora, 2007). Plusieurs études suggèrent que ATAC contient une seconde activité d'acétylation grâce à sa sous-unité ATAC2 (Suganuma et al, 2008). D'autre part, de nombreux travaux récents ont montré que ATAC était requis pour le développement des insectes et des mammifères et que le complexe était impliqué dans la régulation de la transcription de gènes qui ne sont souvent pas contrôlés par SAGA (Guelman et al, 2009; Krebs et al, 2011; Nagy et al, 2010; Pankotai et al, 2010).

## **C. Modularité structurale et fonctions de SAGA**

Suite aux premières caractérisations biochimiques du complexe SAGA et à l'identification de nombreuses sous-unités par cribles génétiques, plusieurs études ont voulu préciser l'organisation fonctionnelle du complexe SAGA. Le premier modèle ainsi obtenu propose que SAGA soit divisé en trois principaux modules fonctionnels sans exclure que d'autres modules puissent exister. Les auteurs de ces études suggèrent ainsi que Gcn5, Ada2 et Ada3 forment un module dont la fonction principale est d'acétyler les histones, que Spt3 et Spt8 régulent la liaison de TBP avec un certain nombre de promoteurs et que Ada1, Spt20 et Spt7 soient définies comme sous-unités structurales de SAGA sur la base d'expériences montrant que la délétion de ces protéines cause une perte de l'intégrité du complexe SAGA (Figure 7).

### **I. Structure tridimensionnelle de SAGA**

#### **1. Structure modulaire**

La structure tridimensionnelle du complexe SAGA de levure à basse résolution (3,1 nm) obtenue par microscopie électronique nous renseigne également quant à l'organisation modulaire du complexe SAGA. Une purification d'affinité TAP (Tandem Affinity Purification) de SAGA à partir de la sous-unité structurale Spt7 a été traitée par des sels de métaux lourds avant d'être observée par microscopie électronique. Des particules allongées de taille 27 x 18 nm ont été observées. A partir des images enregistrées, un modèle du complexe SAGA a pu être généré. Il en ressort que SAGA est composé principalement par quatre domaines structuraux agencés les uns à la suite des autres. Un cinquième domaine apparaît dans une sous-classe d'images et semble être beaucoup plus flexible (Figure 11) (Wu et al, 2004). Précédemment, la structure du complexe SAGA humain avait été analysée par microscopie électronique et il apparaît que l'organisation allongée avec 5 domaines principaux est conservée entre le complexe humain et celui de levure (Figure 11 A.) (Brand et al, 1999a; Wu et al, 2004).

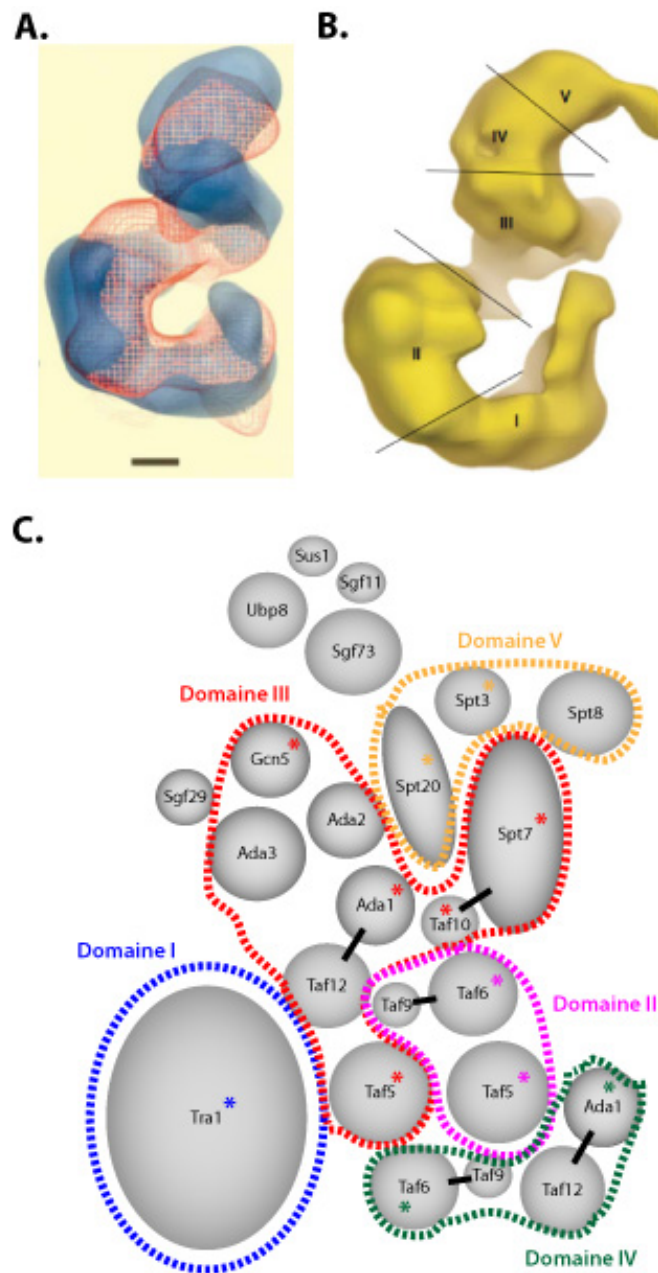
#### **2. Localisation des sous-unités**

Pour localiser la position de plusieurs sous-unités de SAGA, des expériences d'immunomarquages ont été réalisées. Pour cela, des purifications de SAGA ont été incubées avec des anticorps spécifiques de diverses sous-unités avant d'être traitées pour être observées en microscopie électronique. A partir du modèle structural du

complexe SAGA, 8 zones équivalentes ont été définies tout autour du complexe. Pour chaque anticorps, la densité additionnelle a été localisée par rapport au modèle de base pour déterminer dans quelle région chacune des sous-unités étudiées se situe et pour avoir une idée du nombre de copies de la sous-unité au sein du complexe.

### **3. Cœur structural de SAGA**

Ces expériences de microscopie électronique ont permis de commencer à déchiffrer l'organisation structurale du complexe SAGA. Nous avons vu que celui-ci partage avec TFIID un sous-ensemble de protéines TAF. On retrouve notamment TAF5, une protéine importante pour l'architecture du complexe TFIID (Leurent et al, 2004) ainsi que plusieurs TAF possédant un domaine Histone Fold. TAF6 et TAF9 interagissent ensemble via ce domaine alors que TAF10 et TAF12 forment un hétérodimère avec deux protéines spécifiques de SAGA, respectivement Spt7 et Ada1 (Gangloff et al, 2001; Gangloff et al, 2000). L'immunolocalisation des protéines TAF5, TAF6 et TAF10 au sein de SAGA a permis de montrer que ces protéines sont présentes dans les domaines II, III et IV au cœur de la structure du complexe (Figure 11 B. et C.). Grâce à des expériences similaires, la position de ces TAF a aussi été déterminée au sein du complexe TFIID (Leurent et al, 2002; Leurent et al, 2004). Ces informations ont servi de référence pour superposer les structures des deux complexes et le résultat suggère que TFIID et SAGA possèdent un cœur structural très similaire. D'autre part, il semblerait que les protéines TAF5, TAF6 et TAF9 soient présentes en deux copies et qu'elles puissent servir de plateforme d'assemblage pour chacun des deux complexes (Wu et al, 2004).



**Figure 11. Structure tridimensionnelle de SAGA obtenue par microscopie électronique.**

**A.** Superposition de la structure de complexe SAGA de la levure *S. cerevisiae* (en bleu) (Wu et al, 2004) avec celle du complexe SAGA humain (en rouge) (Brand et al, 1999a). La barre représente 3,7 nm. **B.** Structure du complexe SAGA de la levure *S. cerevisiae* contenant le cinquième domaine plus flexible. Les 5 domaines structuraux du complexe SAGA sont mis en évidence. **C.** Représentation schématique du contenu en sous-unités des 5 domaines structuraux du complexe SAGA. La localisation des sous-unités marquées par une étoile a été déterminée expérimentalement. Les sous-unités n'étant pas incluses dans un des cinq domaines n'étaient pas encore identifiées au sein de SAGA au moment où cette étude structurale a été réalisée. La figure est adaptée de : (Timmers & Tora, 2005; Wu et al, 2004).

TAF10 et TAF12 sont également trouvées à deux endroits dans la structure du complexe TFIID (Leurent et al, 2002). La localisation de TAF12 n'a pas été analysée directement dans SAGA, mais son partenaire Ada1 a été observé à la fois dans les

domaines III et IV. La situation concernant TAF10 est un petit peu plus complexe. Au sein de SAGA, deux sites de fixation des anticorps anti-TAF10 ont été localisés dans les domaines II et III. Cependant, les auteurs n'excluent pas qu'il s'agisse en fait d'un site unique étant donné que les images obtenues présentent un fort degré de symétrie. D'autres observations suggèrent d'ailleurs qu'une seule molécule de TAF10 est présente dans SAGA. Premièrement, au sein de TFIID, TAF10 forme des hétérodimères avec deux autres TAF, TAF3 et TAF8 (Gangloff et al, 2001) alors qu'il n'a qu'un seul partenaire au sein de SAGA, Spt7. De plus, une seule copie de la protéine Spt7 a été localisée dans le domaine III de SAGA, cela suggère donc qu'une seule molécule de TAF10 est présente dans le complexe SAGA. L'ensemble de ces résultats indique donc que la protéine TAF5 ainsi que les paires Histone Fold TAF6-9, TAF10-Spt7 et TAF12-Ada1 se situent au cœur de la structure du complexe SAGA ce qui est en accord avec leur fonction structurale supposée dans la formation de l'ossature du complexe. En accord avec cette idée, le complexe SAGA purifié à partir de Spt7 dans une souche *ada1Δ* perd au moins TAF12, Tra1, Spt3, Ada3, Ada2, Sgf29 et Gcn5. De plus la structure de ce complexe minimal correspond au domaine III avec une partie des domaines II et IV de SAGA dans lesquels se situent les protéines structurales dont nous venons de parler (Wu et al, 2004).

#### **4. Localisation des fonctions co-activatrices de SAGA**

Enfin, des sous-unités liées à une fonction spécifique de SAGA sont retrouvées dans différents modules. Tra1, dont la fonction principale est d'interagir avec des activateurs (cf. A. III. 2.) se situe dans le domaine I. Gcn5, l'enzyme qui catalyse l'acétylation des histones H3 est présente dans le domaine III. Enfin, Spt3 dont la fonction est associée à la régulation de TBP au niveau de certains promoteurs chez la levure (cf. A. I.) est présent dans le domaine V. Il semble donc que chaque module structural de SAGA puisse être associée à une fonction particulière. La modularité du complexe est donc retrouvée au niveau structural (Figure 11 C.) (Wu et al, 2004).

L'utilisation de nouvelles techniques de microscopie électronique (Cryo Electron Microscopy) a permis à l'équipe de Patrick Schultz d'obtenir une structure beaucoup plus précise du complexe TFIID (Papai et al, 2010; Papai et al, 2009). Une étude est actuellement en cours pour obtenir la structure de SAGA en utilisant cette nouvelle méthode et il sera bien sûr intéressant de localiser le DUBm au sein de cette future structure afin de connaître la position relative de l'ensemble des modules fonctionnels de SAGA.

## **II. Analyse protéomique et architecture de SAGA**

### **1. Approche expérimentale**

Une étude récente a eu pour but de mieux définir l'organisation structurale du complexe SAGA par spectrométrie de masse chez la levure *S. cerevisiae* (Lee et al, 2011a). Pour cela, les auteurs ont analysé de nombreuses purifications d'affinité du complexe SAGA par la technique de MudPit, en évaluant l'abondance de chaque sous-unité du complexe d'après sa valeur dNSAF (distributed Normalized Spectral Abundance Factors).

L'utilisation de mutants permet de déterminer les sous-unités qui restent associées avec la protéine utilisée pour la purification en l'absence d'une autre sous-unité du complexe. 34 purifications ont ainsi été réalisées à partir d'une combinaison de 10 souches étiquetées et de 10 souches délétées donnant pour chaque purification une valeur de dNSAF pour l'ensemble des sous-unités du complexe. Un profil de 34 valeurs dNSAF a ainsi été obtenu pour chaque protéine de SAGA. Ces profils dNSAF ont enfin été catégorisés pour essayer de définir des modules structuraux (Lee et al, 2011a).

Tout d'abord, le complexe SAGA a été purifié à partir de souches sauvages exprimant chacune l'étiquette TAP fusionnée à l'extrémité C-terminale d'une sous-unité du complexe. Quelque soit la protéine étiquetée utilisée pour la purification du complexe, les valeurs dNSAF calculées pour les sous-unités de SAGA sont toujours très similaires. Ceci semble indiquer que le complexe SAGA est très stable chez la levure (Lee et al, 2011a).

### **2. Les 4 modules structuraux**

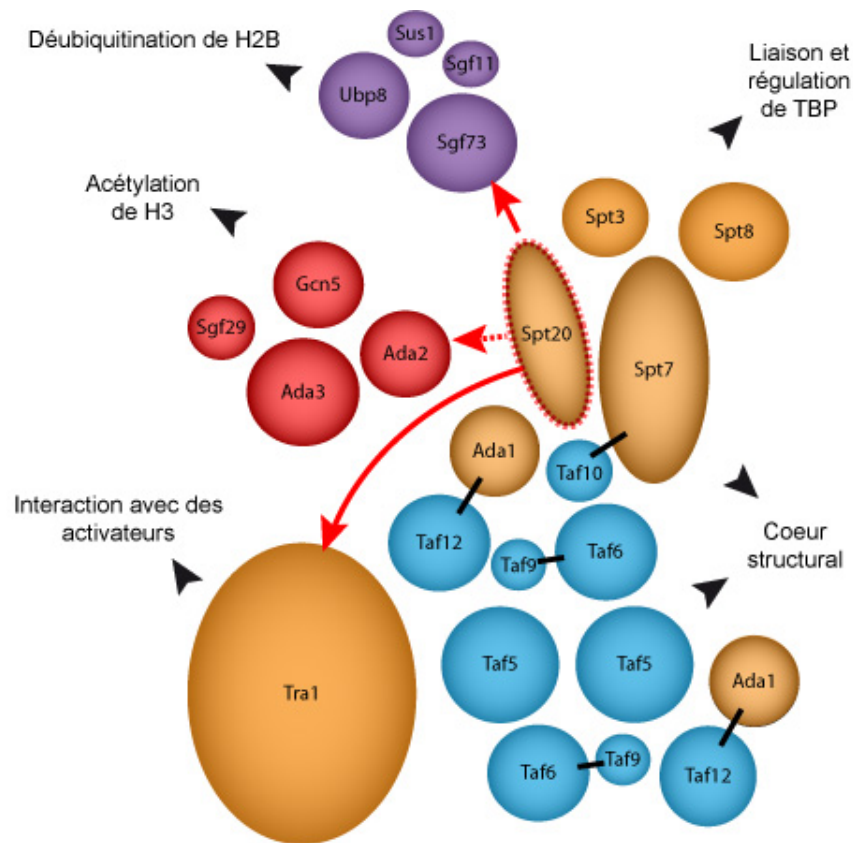
L'utilisation de purifications à partir de souches mutantes a ensuite permis de définir quatre modules. D'après cette étude, le DUBm contient : Sgf11, Sus1, Ubp8 et Sgf73, et cette dernière permet de recruter le DUBm dans SAGA. Le HATm est composé de Ada2, Ada3, Sgf29 et Gcn5 et semble être intégré à SAGA par la sous-unité Ada2. Enfin, deux modules formant l'ossature structurale de SAGA ont été caractérisés. Un premier, appelé module SPT, contient Spt3, Spt7, Spt8, Spt20, Ada1 et Tra1. Enfin, un second module structural, le module TAF, possède TAF5, TAF6, TAF9, TAF10 et TAF12 (Figure 12) (Lee et al, 2011a).

### **3. Rôle structural et fonctionnel de Spt20**

De façon intéressante, cette étude permet de mieux comprendre le rôle de la protéine Spt20 dans l'organisation structurale et fonctionnelle de SAGA. Précédemment,



## Partie 2 : SAGA, un complexe co-activateur de la transcription



**Figure 12. Modularité du complexe SAGA définie par une approche de protéomique.**

L'analyse par spectrométrie de masse de purifications de SAGA réalisées à partir de souches délétées de différentes sous-unités du complexe SAGA a permis de définir 4 modules structuraux principaux (Lee et al, 2011a). Le DUBm, le HATm, le module structural SPT et le module structural TAF sont représentés respectivement en violet, en rouge, en orange et en bleu. La sous-unité Spt20 (entourée par des pointillés rouges) est nécessaire au recrutement du DUBm et de Tra1 dans SAGA (représenté par une flèche rouge). De plus, Spt20 stabilise le recrutement du HATm au sein du complexe (représenté par une flèche en pointillés rouges). La fonction des modules de SAGA définis dans la Figure 4 est indiquée derrière une tête de flèche.

il a été montré que Spt20 jouait un rôle important dans l'intégrité du complexe (Wu & Winston, 2002). Cependant, cette nouvelle étude montre qu'en purifiant SAGA dans une souche *spt20Δ*, l'ensemble des protéines TAF de SAGA, ainsi que Ada1, Spt3, Spt7, Spt8 et le HATm restent associées entre elles. Par contre, les sous-unités du DUBm ainsi que Tra1 ne sont pas co-purifiées avec le reste de SAGA. Ceci indique donc que Spt20 doit participer au recrutement du DUBm et la protéine Tra1 au sein du complexe SAGA (Figure 12). De plus, dans une souche *spt20Δ*, les auteurs ont également purifié les protéines associées à Ubp8 et ont observé que les quatre sous-unités du DUBm sont liées normalement entre elles et qu'elles forment donc un module libre indépendant de SAGA. La question est de savoir si ce DUBm est capable déubiquitiner H2B *in vivo*. En

effet, dans une souche sauvage de levure, le DUBm est intégré à SAGA et son activité permet de maintenir un niveau global faible de H2Bub. Par contre dans une souche *ubp8Δ*, l'activité DUB de SAGA est perdue ce qui a pour conséquence d'augmenter très fortement le niveau global de H2Bub. De façon intéressante, dans une souche *spt20Δ* ou dans un autre mutant dans lequel on a un DUBm intègre mais indépendant de SAGA, le niveau global de H2Bub est intermédiaire entre la situation sauvage et la situation *ubp8Δ* (Henry et al, 2003; Köhler et al, 2008). Cette observation indique donc que la déubiquitination optimale de H2B requiert non seulement la formation d'un DUB stable et intègre mais en plus l'intégration de celui-ci au sein de SAGA. A ce propos, on peut donc imaginer que le DUBm est incorporé au sein du complexe SAGA pour bénéficier d'un environnement structural favorable lui permettant de faciliter ses déplacements sur la chromatine.

L'association du HATm avec le cœur structural de SAGA semble également fragilisée en l'absence de Spt20 (Figure 12). Certes, le HATm est présent avec les sous-unités structurales dans des purifications faites à partir de TAF5, de TAF9 ou de Ada1, mais par contre, en purifiant SAGA depuis la sous-unité Ada2 du HATm, seules les sous-unités de ce module sont co-purifiées. Ceci montre également, qu'un HATm indépendant de SAGA peut exister *in vivo*.

En conclusion, on peut donc supposer que les différents modules de SAGA sont pré-assemblés *in vivo* avant d'être intégrés au sein du complexe SAGA. Cependant, on a vu (cf. B. I.) qu'une souche de *S. cerevisiae spt20Δ* présente divers phénotypes caractéristiques d'une perte des différentes fonctions de SAGA. Or, dans ces levures, il apparaît que les différents modules fonctionnels sont assemblés dans la cellule mais pas associés les uns aux autres au sein de SAGA. Ces expériences semblent donc montrer que chacun des modules de SAGA doit être incorporé au sein du complexe pour exercer pleinement sa fonction.

#### **4. Organisation du cœur structural**

Le deuxième point que je voudrais discuter brièvement concerne l'ossature structurale du complexe. Parmi les sous-unités impliquées, on retrouve un sous-ensemble de protéines TAF qui sont partagées avec le facteur général de la transcription TFIID. De ce fait, les gènes codant ces protéines sont essentiels et ne peuvent pas être délétés facilement. La stratégie expérimentale mise en œuvre par les auteurs de cette étude permet donc de mieux caractériser la fonction structurale de ces sous-unités. Pour

cela, des purifications à partir de plusieurs TAF ont été réalisées dans des souches présentant la délétion d'autres sous-unités dites structurales. Il est également possible de mettre en relief ces nouveaux résultats avec le modèle généré par microscopie électronique. Les expériences de spectrométrie de masse suggèrent que le cœur structural de SAGA se divise en deux sous-parties avec d'un côté les protéines TAF (-5-6-9-10-12) qui semblent interagir fortement entre elles, et d'un autre côté, des sous-unités du module SPT, avec notamment Spt7, Spt20 et Ada1. Une telle séparation en deux cœurs structuraux indépendants semble surprenante étant donné que des interactions fortes ont été observées entre des sous-unités de ces deux modules. En effet, Ada1 et TAF12 ainsi que Spt7 et TAF10 sont supposées former des dimères par leur domaine de Histone Fold (Gangloff et al, 2001; Gangloff et al, 2000). Cette étude pose donc la question du rôle exact joué par ces interactions et met en doute une partie des observations réalisées en microscopie électronique. En effet, on peut se demander si les dimères de type histone jouent un rôle crucial dans la formation du cœur structural de SAGA ou s'ils sont simplement impliqués dans la fixation des deux modules TAF et SPT. Des études analysant l'effet de la délétion des domaines Histone Fold de Ada1 et de Spt7 sur l'intégrité structurale de SAGA semblent nécessaires. On peut enfin se demander si la nette séparation observée pour les profils des sous-unités TAF par rapport aux sous-unités Spt n'est pas due au fait que les TAF sont également présentes au sein du complexe TFIID et il me semble que cela pourrait biaiser les résultats obtenus par cette approche.

## **5. Vers une meilleure définition architecturale de SAGA**

Jusqu'à présent, seule la délétion de sous-unités entières de SAGA a été analysée par ce type d'analyse en spectrométrie de masse. Il sera sans doute également intéressant d'étudier les effets des délétions de domaines de protéines du complexe. Cela permettra de déterminer si tel ou tel domaine joue un rôle dans l'assemblage de modules de SAGA ou dans l'ancrage d'un module avec un ou plusieurs autres. Comme nous l'avons vu dans le paragraphe précédent, l'incorporation des modules au sein de SAGA est nécessaire pour leur fonction *in vivo*. Mieux comprendre la relation structurale et fonctionnelle qui existe entre ces modules permettra sans doute de mieux comprendre le mode de fonctionnement du complexe SAGA.

### III. Analyse des transcriptomes et modules fonctionnels :

#### 1. Chez *S. cerevisiae*

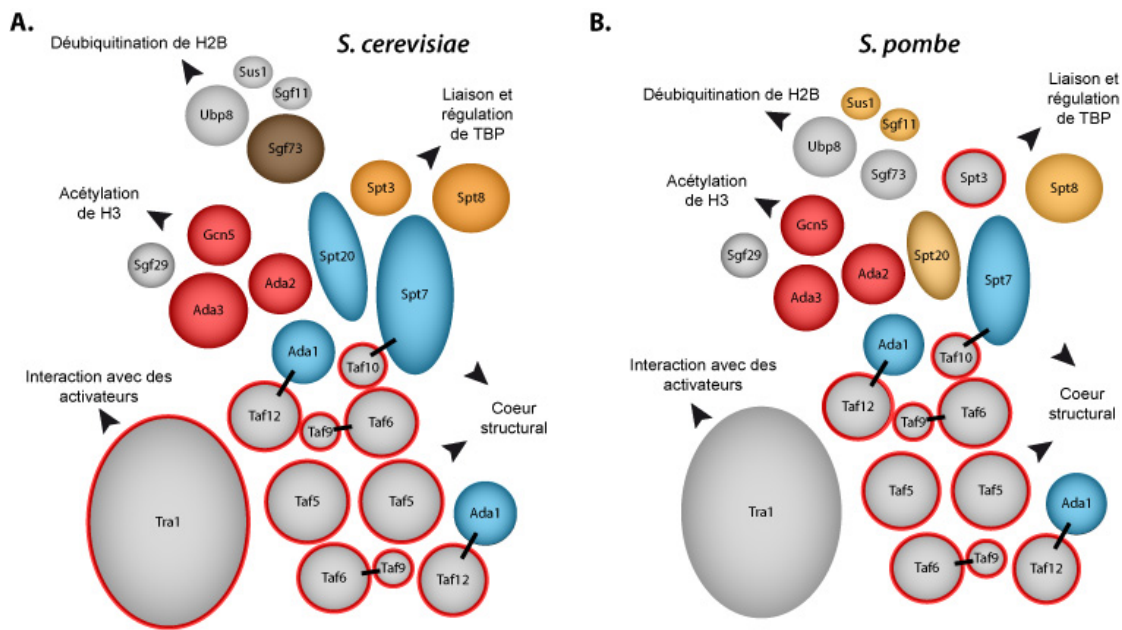
##### Approche expérimentale

Chez la levure *S. cerevisiae*, le profil d'expression des gènes d'un mutant de délétion de chacune des sous-unités non essentielles de SAGA a été analysé par la technique de puces transcriptionnelles ( $\mu$ -array) et comparé au profil d'une souche sauvage. L'ensemble des mutants de SAGA ne donne pas un profil similaire de dérégulation de l'expression des gènes. Au contraire, la classification des profils observés, permet de retrouver les modules structuraux et fonctionnels définis par microscopie électronique, par spectrométrie de masse ou par les screens génétiques (Lenstra et al, 2011).

##### Définition de 3 modules fonctionnels

La délétion des sous-unités structurales Spt20, Ada1 ou Spt7 cause la dérégulation du même groupe de gène et est responsable du phénotype transcriptionnel le plus fort, avec le plus grand nombre de gènes dérégulés (Figure 13 A.). Tous les autres mutants de SAGA analysés causent une modification de l'expression d'un nombre plus limité de gènes au sein de ce même groupe. De façon intéressante, *spt3 $\Delta$*  et *spt8 $\Delta$*  ont un profil similaire et régulent un sous-groupe de gènes différent de celui régulé par *gcn5 $\Delta$* , *ada2 $\Delta$*  et *ada3 $\Delta$* . Ceci illustre donc bien la présence d'un module fonctionnel autour de Gcn5 (permettant l'acétylation des histones), et d'un autre contenant les protéines Spt3 et Spt8 (pouvant réguler l'interaction de TBP avec certains promoteurs) (Figure 13 A.). On peut également noter que ces deux modules ne régulent que très peu de gènes en commun. De plus, des effets opposés sur l'expression de certains gènes sont observés avec des mutants de l'un ou l'autre de ces deux modules fonctionnels.

Enfin, on peut noter que certains mutants n'affectent presque pas l'expression de gènes dans les conditions utilisées dans cette étude. C'est le cas de Sgf29 (une sous-unité du HATm) ainsi que des sous-unités du DUBm. Seule la souche *sgf73 $\Delta$*  provoque la répression d'un petit nombre de gènes (~30) ce qui s'explique sans doute par le fait que Sgf73 fait le lien entre SAGA et le module DUB (Figure 13 A.). Le phénotype observé est d'ailleurs probablement causé par l'absence du domaine de Sgf73 qui interagit avec SAGA (Lenstra et al, 2011).



**Figure 13. Analyse des transcriptomes et définition de modules fonctionnels.**

Le profil d'expression des gènes de tous les mutants de délétion viables des sous-unités de SAGA chez les levures *S. cerevisiae* (A.) et *S. pombe* (B.) a été analysé. Ceci a permis de définir différents modules fonctionnels (Helmlinger et al, 2011; Lenstra et al, 2011). Les sous-unités entourées en rouge sont des sous-unités essentielles alors que les sous-unités dont la délétion ne provoque pas de changements transcriptionnels importants sont uniquement en gris. A. Chez *S. cerevisiae*, trois groupes de sous-unités ont été définis : un premier contenant Ada1, Spt7 et Spt20 (en bleu), un second avec Gcn5, Ada2 et Ada3 (en rouge) et un troisième avec Spt3 et Spt8 (en orange). La délétion de Sgf73 provoque la dérégulation d'un petit nombre de gènes mais le profil de dérégulation peut être difficilement regroupé avec celui d'autres sous-unités. B. Chez *S. pombe*, trois groupes de sous-unités ont pu être définis : un premier contient les sous-unités Ada1 et Spt20 (en bleu), un second, les sous-unités Gcn5, Ada2 et Ada3 (en rouge) et un troisième, les sous-unités Spt8, Spt20, Sgf11 et Sus1 (en orange). Notons que chez *S. pombe*, Spt3 est une sous-unité essentielle mais que Tra1 ne l'est pas. A. et B. La fonction des modules de SAGA définis dans la Figure 4 est indiquée derrière une tête de flèche.

## 2. Chez *S. pombe*

### Approche expérimentale

Chez la levure *S. pombe*, la composition en sous-unités du complexe SAGA est identique à celle du complexe de la levure *S. cerevisiae* (Helmlinger et al, 2008). A partir des mutants de délétion des sous-unités non essentielles de SAGA, le même type d'analyse des profils d'expression des gènes a été réalisé. Alors qu'une partie des conclusions est similaire à celles faites chez *S. cerevisiae*, sur de nombreux points, des différences importantes sont observées. Au niveau de la viabilité des souches, tout d'abord, *tra1* est essentiel chez *S. cerevisiae*, mais sa délétion n'est pas létale chez *S. pombe*. A l'inverse, *spt3Δ* chez *S. pombe* n'a pas pu être analysée car la souche présente de très gros problèmes de croissance (Helmlinger et al, 2011).

Le complexe SAGA chez *S. pombe* est modulaire

Comme chez *S. cerevisiae*, divers profils de modifications de l'expression des gènes sont observés soulignant que l'aspect modulaire du complexe SAGA est conservé chez *S. pombe*. Ces profils ont pu être catégorisés en trois classes. La première comprend les souches *spt7Δ* et *ada1Δ*, la deuxième, les souches *ada2Δ*, *ada3Δ*, et *gcn5Δ*, et la classe 3, les souches *spt20Δ*, *spt8Δ*, *sgf11Δ* et *sus1Δ*. Les profils des autres mutants analysés (*ubp8Δ*, *sgf73Δ*, *sgf29Δ* et *tra1Δ*) ne peuvent pas être regroupés et le nombre de gènes dont l'expression est modifiée de façon significative est limité (Figure 13 B.).

Concernant les trois groupes décrits ci-dessus, les classes 1 et 3 régulent le plus grand nombre de gènes. Par ailleurs, seule une petite fraction des gènes est régulée par au moins deux des trois catégories de mutants. En effet, l'expression de la plus grande partie des gènes est spécifiquement altérée dans les mutants de l'une des classes. Ceci montre clairement que SAGA régule la plupart de ses gènes cibles avec un module fonctionnel bien particulier (Helmlinger et al, 2011).

Comparaison avec les résultats obtenus chez *S. cerevisiae*

Si l'on compare les trois groupes définis chez *S. pombe* avec ceux caractérisés chez *S. cerevisiae*, un seul est identique. Il s'agit de la classe 2 qui correspond au module d'acétylation des histones avec Ada2, Ada3 et Gcn5 (Figure 13 B.). Un autre correspondant aux sous-unités structurales de SAGA est relativement similaire avec Spt7 et Ada1 qui sont retrouvées dans les deux espèces (Figure 13 B.). La dernière classe est par contre très différente. Spt20 et Spt8 sont groupés avec deux des quatre sous-unités du DUBm. Ceci est très surprenant et difficile à expliquer avec ce qui est connu de l'organisation structurale de SAGA. Cela suggère que Sgf11 et Sus1 ne sont pas uniquement impliqués dans la régulation de l'activité de débubiquitination de Ubp8 chez *S. pombe* (Figure 13 B.) (Helmlinger et al, 2011).

Analyse phénotypique des mutants de *S. pombe*

Dans la même étude, les auteurs ont également caractérisés ces mutants de SAGA en mesurant leur capacité à se développer sur une quarantaine de milieux sélectifs. Une classification des souches en fonction de leurs phénotypes de croissance permet de préciser les conclusions faites grâce à l'analyse de l'expression des gènes. Les mutants *spt7Δ*, *ada1Δ* et *spt3Δ* ne sont pas capables de se développer sur la quasi totalité des milieux analysés et les protéines correspondantes semblent donc être les sous-unités

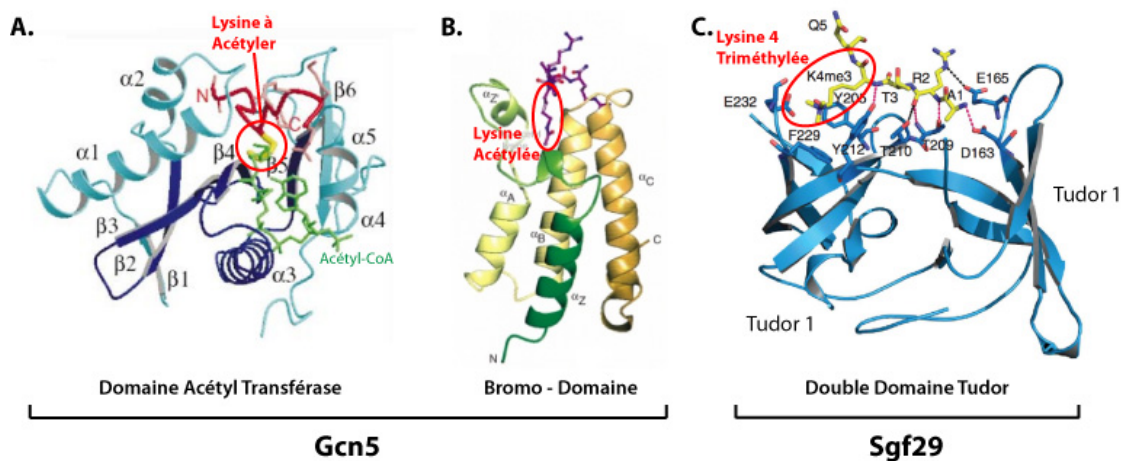
structurales du complexe. Les phénotypes des souches *spt20Δ* et *spt8Δ* sont aussi très similaires et un petit moins marqués. Les sous-unités du HATm se regroupent également très bien même si le mutant *sgf29Δ* présente des phénotypes plus modérés que les trois autres mutants *gcn5Δ*, *ada2Δ* et *ada3Δ*. Enfin, pour ce qui concerne le DUBm, les souches ne présentent un phénotype de retard de croissance que sur un nombre très limité de milieux sélectifs. De plus, la corrélation entre les phénotypes observés pour les quatre sous-unités du module n'est pas aussi importante que pour les autres modules. Une éventuelle fonction des sous-unités de ce module dans un autre processus que la régulation de la déubiquitination des histones reste à explorer. Enfin, le mutant *tra1Δ* présente aussi un nombre faible de phénotypes ce qui semble indiquer que la fonction de cette protéine n'est pas cruciale pour le complexe SAGA chez *S. pombe* (Helmlinger et al, 2011).

#### **IV. Analyses structurales et fonctionnelles de modules de SAGA**

##### **1. Le HAT module**

Il est admis que l'acétylation de lysines de la queue N-terminale de l'histone H3 par le HATm de SAGA facilite l'initiation de la transcription par la PolII en participant au relâchement de la chromatine au niveau de certains promoteurs et en facilitant le recrutement d'autres complexes transcriptionnels qui possèdent un domaine de liaison de lysines acétylées (Nagy & Tora, 2007). Chez *S. cerevisiae*, une étude suggère également que SAGA pourrait acétyler les histones H3 au niveau de la région transcrite du gène *GAL1* ce qui faciliterait l'élongation de la transcription, notamment en stimulant l'éviction des nucléosomes (Govind et al, 2007). Chez la levure *S. pombe*, il a aussi été montré que l'acétylation des histones par Gcn5 participait à l'éviction des nucléosomes de la région transcrite de gènes, régulés par un stress (Sanso et al, 2011).

Dès la purification des complexes SAGA chez la levure ou chez l'homme, il a été démontré que *in vitro* seul le complexe SAGA mais pas l'enzyme GCN5 isolée était capable d'acétyler l'histone H3 dans le contexte du nucléosome (Brand et al, 1999b; Grant et al, 1997; Ogryzko et al, 1998). Depuis, la structure cristallographique du domaine acétyl-transférase de la protéine Gcn5 de *Tetrahymena* a été résolue en présence d'acétyl-CoA (le donneur de groupement acétyl) et d'un peptide correspondant à la queue N-terminale de l'histone H3 (Figure 14 A.) (Rojas et al, 1999). D'autre part, la structure du Bromo-domaine de Gcn5 de *S. cerevisiae* démontre que ce domaine peut



**Figure 14. Structure de domaines du HATm de SAGA.**

**A.** Structure du domaine acétyl-transférase de Gcn5 de Tetrahymena (en bleu) en présence d'acétyl-CoA (donneur de groupement acétyl, en vert) et d'un peptide correspondant à la queue de l'histone H3 (en rouge, la lysine à acétyle est en jaune). La référence PDB est : 1QSN. **B.** Structure du Bromo-domaine de Gcn5 de *S. cerevisiae* en complexe avec un peptide correspondant à la queue N-terminale de l'histone H4 dans lequel la lysine 16 est acétylée. La référence PDB est : 1E6I. **C.** Structure du double domaine Tudor de la protéine Sgf29 de la levure *S. cerevisiae* en présence d'un peptide correspondant à la queue N-terminal de l'histone H3 et qui contient une lysine 4 triméthylée. Cette modification est reconnue par la poche formée par le domaine Tudor 1. La référence PDB est : 3MP1. La figure est adaptée de : (Bian et al, 2011; Owen et al, 2000; Rojas et al, 1999).

interagir avec l'histone H4 acétylé sur sa lysine 16 (Figure 14 B.) (Owen et al, 2000). Une autre étude indique que le Bromo-domaine de Gcn5 participe au recrutement du complexe SAGA sur une chromatine acétylée (Hassan et al, 2002). Par ailleurs, une étude structure fonction récente a démontré que le double domaine Tudor de Sgf29 (une sous-unité du HATm) peut reconnaître de façon spécifique la lysine 4 de l'histone H3 lorsque celle-ci est di ou triméthylée et que cette interaction semble nécessaire pour permettre à Gcn5 d'exercer son activité HAT *in vivo* (Figure 14 C.) (Bian et al, 2011). Enfin, on peut noter que chez la levure *S. cerevisiae* ainsi que chez l'homme, il a été montré que Ada2/ADA2b jouait un rôle important pour stimuler l'activité HAT de Gcn5/GCN5 sur des nucléosomes *in vitro* (Gamper et al, 2009; Sterner et al, 2002b). Cependant, il manque encore une structure de l'ensemble du HATm qui permettrait de comprendre comment les différentes sous-unités du HATm participent précisément à la régulation de l'activité d'acétylation de Gcn5. De plus, de nombreuses questions restent encore sans réponse précise, notamment au sujet de la spécificité d'acétylation de lysines par le HATm de SAGA.

En effet, chez la levure, la délétion de Gcn5 cause une perte presque totale de l'acétylation de deux résidus de l'histone H3, la lysine 9 et la lysine 18 (Bian et al, 2011).



Dans des cellules murines, une analyse par spectrométrie de masse a indiqué que les délétions combinées de GCN5 et de PCAF conduisent à la perte de l'acétylation de la lysine 9 de l'histone H3 mais qu'aucune des autres acétylations analysées dans l'étude n'est affectée (Jin et al, 2011). Cela démontre donc que Gcn5 est la seule enzyme HAT à pouvoir acétyler efficacement la lysine 9 de l'histone H3 *in vivo*. De façon très surprenante, des résultats apparemment contradictoires ont été observés lors de tests d'acétylation *in vitro*. Dans ce cas, il semble en effet que la cible principale de SAGA ou de l'enzyme Gcn5 recombinante seule soit la lysine 14 de l'histone H3 et qu'un peptide muté à cette position en Arginine ne soit presque pas acétylé (Brand et al, 1999b) et (Résultats obtenus dans le laboratoire de Laszlo Tora). L'ensemble de ces observations suggère donc que les différentes sous-unités du HATm jouent un rôle important dans la régulation de l'activité acétyl-transférase de Gcn5. On peut par exemple imaginer que *in vivo*, la reconnaissance de modifications de la chromatine par des domaines du HATm puisse être requise pour positionner de façon optimale le domaine catalytique de Gcn5 par rapport à son substrat.

### 2. Le DUB module

Chez *S. cerevisiae*, le DUBm est constitué par les 4 sous-unités Sgf11, Sgf73, Sus1 et Ubp8 alors que le module humain est formé par les protéines ATXN7L3, ATXN7, ENY2 et USP22. En 2010, deux groupes ont publié la structure tridimensionnelle du DUBm de levure, et je vais présenter ces travaux en détail dans un prochain chapitre (Kohler et al, 2010; Samara et al, 2010) (cf. Partie sur les USP). Notons simplement ici, que ces résultats structuraux mettent en lumière comment le complexe réseau d'interactions existant entre les sous-unités du DUBm permet de stimuler l'activité enzymatique de Ubp8 (Kohler et al, 2010; Samara et al, 2010) (cf. Partie sur les USP). D'autre part, nous verrons dans la partie Résultats, que chez l'homme, les différentes sous-unités du DUBm sont également impliquées dans la régulation allostérique de l'activité de USP22.

Concernant la nature des protéines déubiquitinées par Ubp8/USP22, je voudrais souligner que chez la levure comme chez l'homme, la principale cible *in vivo* de l'activité DUB de SAGA est l'histone H2B (Henry et al, 2003) (cf. Partie Résultats). En effet, dans ces deux organismes, la perte de l'activité DUB de SAGA conduit à une augmentation globale du niveau de H2Bub (environ 5X chez l'homme et 10X chez la levure) (Henry et al, 2003) (cf. Partie Résultats). Chez l'homme, nos résultats indiquent que SAGA pourrait également participer à la déubiquitination de H2A mais de façon beaucoup moins

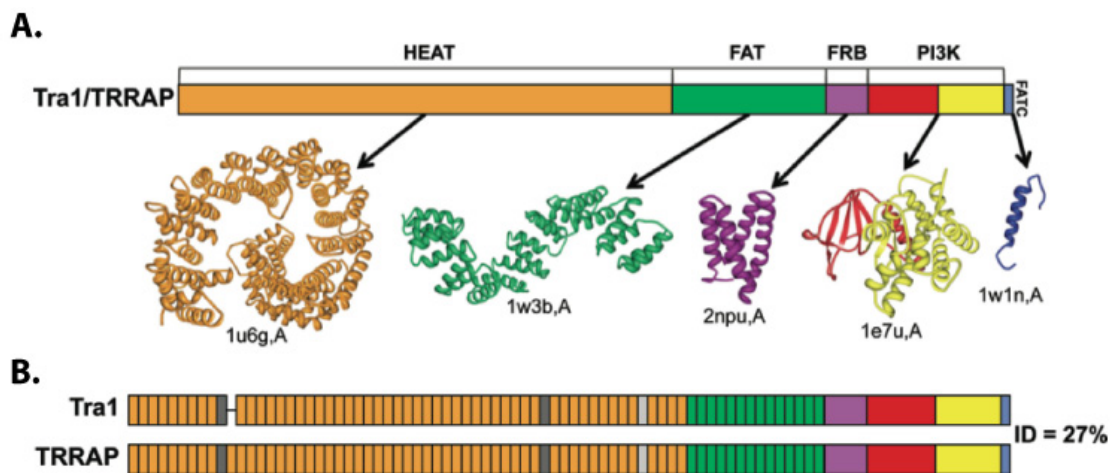
prononcée que pour H2B. Par ailleurs, plusieurs études récentes suggèrent que l'activité DUB de SAGA pourrait aussi agir sur des substrats non histone. Une étude a par exemple proposé que USP22 pouvait déubiquitiner TRF1, une protéine associée avec les télomères, ce qui permettait de contrôler la quantité cellulaire de cette protéine et de maintenir l'intégrité des télomères (Atanassov et al, 2009). Pour des détails sur les substrats non histone de Ubp8/USP22 voir la revue : (Weake & Workman, 2011).

Dans la partie discussion, nous reparlerons abondamment de la régulation allostérique de Ubp8/USP22 par les autres sous-unités du DUBm ainsi que des aspects fonctionnels concernant ce module. Dans une prochaine partie, je présenterai aussi ce qui est connu sur l'ubiquitination de H2B (cf. Discussion et Partie sur H2Bub).

### 3. Tra1 et organisation structurale de SAGA

#### Organisation modulaire de Tra1/TRRAP

Chez la levure, Tra1 a une taille d'environ 400 kDa et est donc une des plus grosses protéines de la cellule. Vu sa taille, Tra1 représente environ 25% de la masse de SAGA et il a été supposé que cette protéine exerce une fonction dans l'architecture du complexe. Nous avons aussi vu que cette sous-unité de SAGA joue un rôle dans le recrutement du complexe par des activateurs (cf. A. III. 2.). Chez *S. cerevisiae*, Tra1 est également



**Figure 15. Organisation en domaines des protéines Tra1 et TRRAP.**

**A.** Les cinq types de domaines retrouvés au sein de Tra1 et de TRRAP sont indiqués sur le schéma et une structure représentative de chaque domaine est présentée avec sa référence PDB. **B.** Comparaison de l'organisation structurale qui a été prédite pour Tra1 et TRRAP. Chaque petit rectangle localisé au sein des domaines HEAT et FAT correspond respectivement à un motif HEAT (structure en orange en A.) ou un motif TPR (structure en vert en A.). Les rectangles gris foncés indiquent des régions qui sont prédites comme dépliées, et les rectangles gris clairs indiquent une région qui ne contient pas de motif HEAT. L'absence d'une répétition d'un motif HEAT dans la séquence de Tra1 est indiquée par un trait noir. Le pourcentage d'identité entre Tra1 et TRRAP est de 27%. La figure a été adaptée de : (Knutson & Hahn, 2011)

présente dans un autre complexe HAT essentiel appelé NuA4. Chez l'homme, on retrouve également TRRAP, l'homologue de Tra1, dans le complexe HAT TIP60 (HIV Tat-interacting protein 60) (McMahon et al, 1998).

Tra1 est un membre de la famille de protéines PIKK (phosphatidylinositol 3-kinase (PI3K)-related protein kinase) mais cette protéine est catalytiquement inactive bien qu'elle conserve un domaine de type PI3K. Comme les autres membres de la famille PIKK, Tra1 est composée par 4 autres types de domaines principalement structurés par des hélices  $\alpha$ . Plus de la moitié de Tra1 est formée par le domaine HEAT (Huntingtin, elongation factor 3, PR65/A, and TOR) qui est constitué par la répétition du motif du même nom. Le domaine FAT (FRAP, ATM, and TRRAP) est formé par la répétition du motif TPR (tetra-tricopeptide). Enfin, le domaine PI3K est entouré par un domaine FRB (FKBP12 rapamycin binding) et un domaine FATC (FAT C-terminal) (Figure 15 A.). Des études structurales ont proposé que ces répétitions de structures hélicoïdales pourraient s'assembler en des super-hélices allongées, dans le sillon desquels des partenaires protéiques pourraient interagir (Knutson & Hahn, 2011).

#### Description de l'analyse structure fonction

Une étude bioinformatique a récemment permis d'analyser la structure secondaire de Tra1 en définissant de façon très précise les limites des différentes unités structurales qui composent les répétitions HEAT et FAT (Knutson & Hahn, 2011). Il en ressort que même si l'identité de séquence entre Tra1 et TRRAP est modeste (27%), l'organisation structurale de ces deux protéines semble très fortement conservée (Figure 15 B.).

Tra1 est une protéine essentielle chez *S. cerevisiae*. Pour analyser sa fonction, les auteurs de l'étude ont utilisé les résultats de leur analyse bioinformatique pour réaliser 44 délétions systématiques d'une centaine de résidus depuis l'extrémité N-terminale de la protéine jusqu'à son extrémité C-terminale (Knutson & Hahn, 2011).

#### Rôle dans la régulation de l'activité HAT

Plus de deux tiers des mutants ne sont pas viables et ce phénotype est retrouvé avec des délétions provenant de chacun des 5 domaines de la protéine. De plus, pour l'ensemble des mutants non viables, une perte de l'interaction de Tra1 avec les autres sous-unités de SAGA ou celles de NuA4 est observée. Pour démontrer cela, les auteurs ont exprimé les différentes versions mutées de Tra1 avec une étiquette Flag dans des souches sauvages et ont observé par western blot si plusieurs sous-unités de SAGA étaient co-

purifiées par une IP anti-Flag. Le fait que l'interaction entre Tra1 et SAGA soit perdue avec des mutants de chacun des domaines de Tra1 suggère que plusieurs points d'ancrages existent entre Tra1 et le reste de SAGA et que de tous ces contacts sont requis pour stabiliser l'intégration de la protéine au sein du complexe (Knutson & Hahn, 2011).

D'autre part, les auteurs indiquent que dans deux mutants, l'interaction entre Tra1 et les sous-unités du HATm est perdue alors que l'interaction avec des sous-unités structurales comme Ada1 ou TAF12 est conservée. Enfin, les auteurs suggèrent qu'un seul mutant de Tra1 altère l'activité d'acétylation de Gcn5 sur H3K9 *in vivo* sans modifier la composition du complexe SAGA. En effet, des expériences de ChIP montrent que Tra1 et Gcn5 sont bien recrutés aux niveaux de certains promoteurs mais que le niveau d'acétylation de H3K9 est plus faible à ces mêmes promoteurs (Knutson & Hahn, 2011). Une telle observation a également été réalisée par un autre groupe pour un mutant ponctuel du domaine PI3K de Tra1 (Mutiu et al, 2007).

Le fait que Tra1 soit une sous-unité essentielle limite donc cette analyse structure fonction et empêche notamment de définir avec précision si Tra1 est requis pour l'assemblage du complexe SAGA chez *S. cerevisiae*. On peut noter que chez *S. pombe*, Tra1 est spécifique de SAGA, que sa délétion n'est pas létale et que dans une telle souche, un complexe SAGA peut se former (Helmlinger et al, 2011). Cette analyse semble par contre indiquer que certains domaines de Tra1 participent à l'assemblage et à la régulation de l'activité d'acétylation des histones de SAGA (Knutson & Hahn, 2011). Il serait intéressant d'étudier si une telle observation est aussi vraie chez *S. pombe* ou si elle est spécifique de *S. cerevisiae*.

## **V. Données additionnelles pour définir la structure du complexe SAGA**

### **1. Modifications post-traductionnelles des sous-unités de SAGA**

En 2009, Mischerikow et al ont établi une liste précise de plusieurs modifications post-traductionnelles des sous-unités de SAGA chez *S. cerevisiae* par une approche de spectrométrie de masse (Mischerikow et al, 2009). De façon intéressante, il a été montré que plusieurs sous-unités du complexe peuvent être hyper acétylées. C'est notamment le cas de Sgf73 et Spt7 (Mischerikow et al, 2009). Une autre étude a montré qu'en plus de ces deux protéines, Ada3 pouvait aussi être acétylée et que l'ensemble des ces acétylations dépendait de Gcn5. De plus, les auteurs de cette étude ont démontré que Gcn5 acétylait ces trois sous-unités au sein du complexe dans lequel elle se situe (Cai et

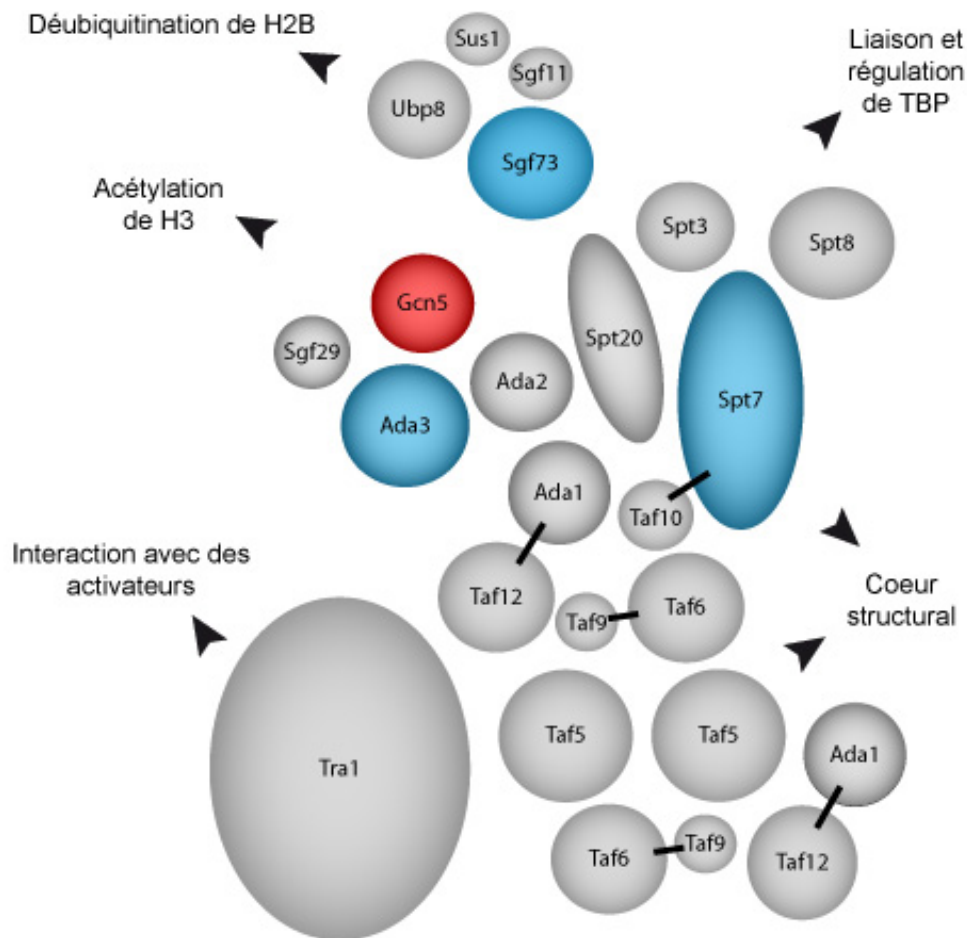
al, 2011). On peut donc imaginer que certaines régions des sous-unités Spt7, Sgf73 et Ada3 sont assez flexible ou qu'elles se situent en permanence à proximité du site catalytique de Gcn5 (Figure 16).

## 2. Stœchiométrie des différentes sous-unités

Les seules informations disponibles sur ce sujet concernent le complexe SAGA chez la levure *S. cerevisiae*.

### Cœur structural de SAGA

La localisation de certaines sous-unités de SAGA par microscopie électronique nous donne une idée de la stœchiométrie au sein du complexe SAGA. Il semble notamment clair que l'on ait un cœur structural composé par un dimère des trois TAF, TAF5, TAF6 et



**Figure 16. Proximité spatiale de certaines sous-unités de SAGA avec Gcn5.**

Chaque sous-unité du complexe SAGA de la levure *S. cerevisiae* est représentée par un rond ou un oval dont la taille est à peu près proportionnelle à la masse de la protéine. Les sous-unités qui peuvent être acétylées par Gcn5 au sein d'un même complexe SAGA sont représentées en bleu. Gcn5 qui possède l'activité acétyltransférase est en rouge. Les autres sous-unités du complexe sont en gris. La fonction des modules de SAGA définis dans la Figure 4 est indiquée derrière une tête de flèche.

TAF9. Deux molécules du dimère Ada1 - TAF12 semblent également faire partie de ce cœur ainsi qu'une copie du dimère Spt7 - TAF10 (Wu et al, 2004).

#### HAT module

Les immunomarquages réalisés par microscopie électronique indiquent également une seule position pour Gcn5 au sein du complexe SAGA (Wu et al, 2004). Des données fonctionnelles plus récentes vont également dans le sens de la présence d'un seul HATm par complexe SAGA et cela semble être aussi le cas chez la souris (Cai et al, 2011; Krebs et al, 2010).

#### Le complexe SAGA contient une copie du DUBm

Une étude récente a montré qu'une seule molécule de Sus1 est présente par complexe SAGA alors que cette protéine est présente en deux copies dans le complexe d'export des ARNm TREX2 (Jani et al, 2009). Pour démontrer cela, une souche diploïde exprimant deux versions de Sus1, Sus1-TAP et Sus1-Flag, a été générée. Après une chromatographie d'affinité sur des billes IgG, les complexes ont été élués par clivage avec la protéase TEV. Une partie des éluats a été analysée sur un gel SDS-PAGE coloré au bleu de coomassie et des bandes caractéristiques des complexes SAGA et TREX2 ont été observées. Le reste des éluats a été utilisé pour réaliser une IP anti-Flag. Le contenu en protéines d'une élution peptidique de cette IP a de nouveau été étudié par un gel SDS-PAGE coloré au bleu de coomassie. Dans ce cas, seules les bandes correspondant au complexe TREX2 sont visibles. Ceci indique que deux molécules de Sus1 sont présentes dans le complexe TREX2 et une seule dans le complexe SAGA (Jani et al, 2009). Vu ce qui est connu des interactions de Sus1 avec ses partenaires du DUBm (cf. Partie sur les USP), on peut conclure que SAGA ne comporte également qu'une seule copie de Sgf73, Ubp8, et Sgf11.

#### Autres sous-unités et perspectives

Vu sa taille, une seule molécule de Tra1 doit être trouvée au sein de SAGA. Enfin, pour ce qui est du module qui interagit avec TBP, une seule localisation pour Spt3 a été observée en microscopie électronique. Etant donné que Spt8 interagit directement avec la région C-terminale de Spt7 et qu'une seule copie de cette protéine semble présente dans SAGA, il est donc probable que Spt8 soit également trouvée une fois dans le complexe (Stern et al, 2002a; Wu et al, 2004; Wu & Winston, 2002).

La stratégie utilisée dans le cas de Sus1 n'est pas très complexe à mettre en œuvre et pourrait notamment être utile pour lever certaines ambiguïtés observées après les immunomarquages de microscopie électronique. Cela pourrait non seulement confirmer le modèle actuel mais aussi être utile si un éventuel futur modèle de SAGA est généré par des techniques de microscopie électronique plus récentes. Plus simplement, il serait même possible de transformer des levures délétées du gène codant la protéine d'intérêt avec deux plasmides qui exprimeraient deux versions étiquetées de cette protéine. Après avoir vérifié l'expression des deux formes de la protéine, il faudrait réaliser une purification de SAGA à partir d'une des étiquettes. Cette purification pourrait alors être analysée par western blot en utilisant des anticorps spécifiques de chacune des étiquettes. Si une seule bande correspondant à l'étiquette utilisée pour la purification est révélée, cela signifierait que la protéine est présente une seule fois dans le complexe. Si par contre, une bande est observée pour chacune des étiquettes, deux copies de la protéine sont présentes dans le complexe. Pour vérifier que la sous-unité n'est pas retrouvée trois fois au sein du complexe, il suffirait d'utiliser la même stratégie en transformant la souche de levure avec une troisième version étiquetée de la protéine.

## VII Conclusions

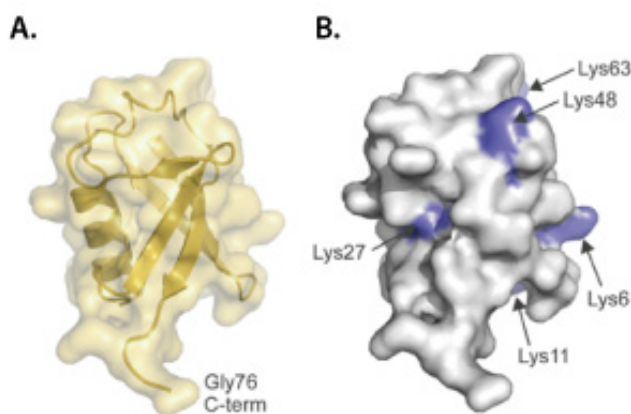
Au total, les différentes approches analysées dans ce chapitre s'accordent pour décrire une structure modulaire du complexe SAGA. En particulier, les différentes études rapportent toutes l'existence d'un HATm composé de Ada2, Ada3, Sgf29 et Gcn5 et d'un DUBm composé de Sgf11, Sgf73, Sus1 et Ubp8. Enfin, bien qu'un cœur structural de SAGA soit proposé dans les diverses études, l'organisation précise de ce cœur en un ou deux modules structuraux et le mode d'assemblage de tous les modules entre eux est encore mal compris. En particulier, le rôle de Tra1/TRRAP, l'organisation structurale du module SPT et une éventuelle dynamique des différents modules les uns par rapport aux autres restent à explorer. Enfin, la conservation de cette organisation structurale au cours de l'évolution reste à démontrer. D'une façon plus générale, de nouvelles études seront nécessaires pour déterminer si toutes les fonctions exercées par le complexe SAGA de levure sont également exercées par le complexe SAGA humain. L'absence d'un homologue de Spt8 chez l'homme pose par exemple la question de la conservation de la fonction de la régulation de TBP par SAGA comme nous l'avons déjà discuté.

## Partie 3 : Les Déubiquitinasés de la famille USP

### A. Le monde de l'ubiquitine

#### I. Machinerie de mise en place

L'ubiquitine est une protéine de 76 acides aminés dont la séquence est strictement conservée chez les eucaryotes (Figure 17) (Hershko et al, 2000). La découverte de l'ubiquitine remonte à une trentaine d'années et l'on sait aujourd'hui que l'ubiquitination, c'est à dire la modification covalente de protéines par l'ubiquitine, est une modification post-traductionnelle aussi largement répandue que la phosphorylation (Wilkinson, 2009).



**Figure 17. La molécule d'ubiquitine.**

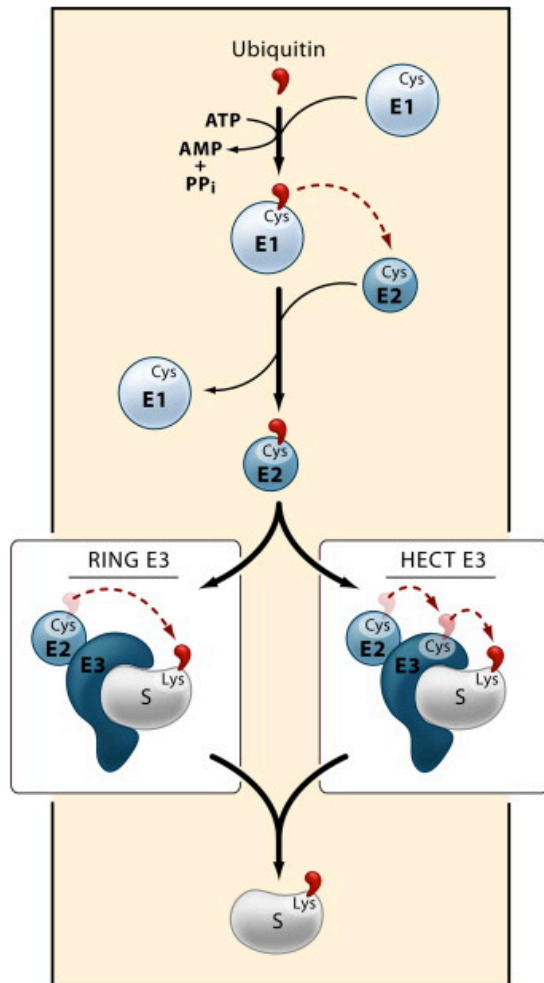
**A.** Représentation de la structure de l'ubiquitine (Code PDB : 1UBQ). La Glycine 76 de l'extrémité C-terminale est mise en évidence. **B.** Localisation des principales lysines qui sont impliquées dans la formation de chaînes de polyubiquitine (en bleu). La figure est adaptée de (Hurley et al, 2006).

Une liaison isopeptidique entre le groupement carboxyl de la glycine C-terminale de l'ubiquitine et le groupement  $\epsilon$ -amine d'une lysine d'une protéine cible peut être synthétisée par l'action successive de trois enzymes : une enzyme d'activation de l'ubiquitine (E1), une enzyme de conjugaison de l'ubiquitine (E2) et une ubiquitine ligase (E3) (Figure 18). Brièvement,

l'ubiquitine est tout d'abord thiol-estérifiée au niveau de son extrémité C-terminale par l'enzyme E1 ce qui consomme une molécule d'ATP, puis l'ubiquitine est conjuguée sur une E2. Il existe enfin deux familles d'E3 ligases qui sont caractérisées par la présence d'un domaine RING (Really Interesting New Gene) ou d'un domaine HECT (Homologous to the E6-AP Carboxyl Terminus) et dont le mode de fonctionnement diffère quelque peu. La première classe d'E3 ligases va interagir avec l'E2 pour l'assister dans le transfert de l'ubiquitine sur un substrat. Par contre, dans le cas des E3 ligases de type HECT, l'ubiquitine est d'abord transférée sur l'E3 ligase qui va ensuite reconnaître son



substrat et l'ubiquitine (Figure 18) (Hurley et al, 2006; Wilkinson, 2009), ainsi que des revues décrivant dans les détails les mécanismes d'ubiquitination : (Belgareh-Touze et

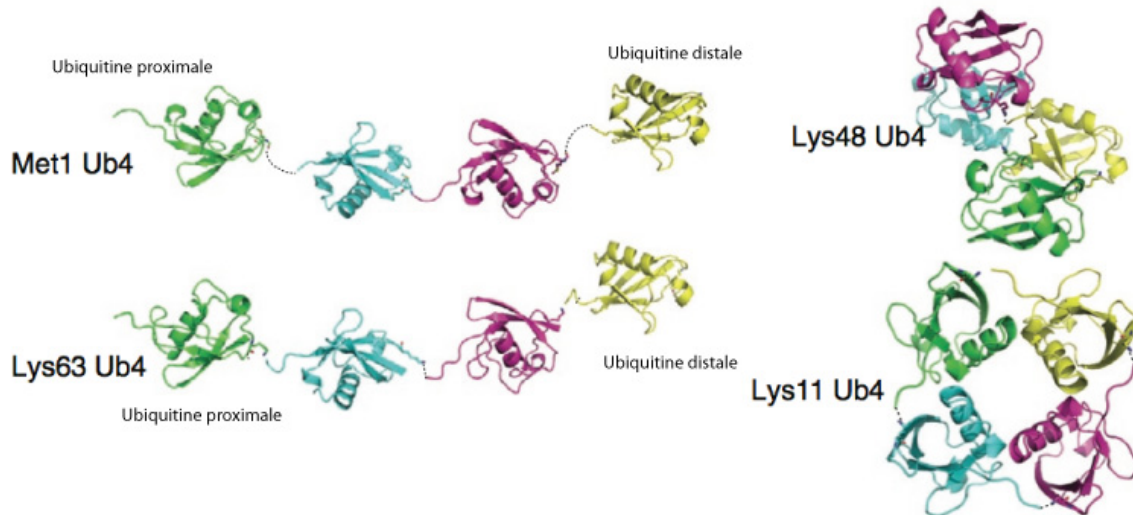


**Figure 18. La machinerie d'ubiquitination.**

Représentation schématique de la cascade enzymatique menant à l'ubiquitination d'un substrat (S). Dans un premier temps, l'hydrolyse d'une molécule d'ATP permet la formation d'une liaison thio-ester entre une cystéine de l'enzyme d'activation E1 et l'extrémité C-terminale de l'ubiquitine. Celle-ci est ensuite transférée sur une cystéine d'une enzyme de conjugaison E2. Enfin, l'ubiquitine est fixée sur une lysine d'un substrat par une enzyme E3 ligase. Les E3 ligases possédant un domaine RING facilitent le transfert de l'ubiquitine de l'E2 vers le substrat alors que pour les E3 ligases présentant un domaine HECT, l'ubiquitine lie transitoirement l'E3 ligase avant d'être transférée sur le substrat. Dans les deux cas, les E3 ligases participent à la reconnaissance spécifique du substrat. Les pointillés rouges retracent le cheminement de la molécule d'ubiquitine vers son substrat. La figure est adaptée de (Hochstrasser, 2006).

al, 2008; Dye & Schulman, 2007; Hochstrasser, 2007; Starita & Parvin, 2006)). Chez les eucaryotes supérieurs, on retrouve 2 enzymes E1, une quarantaine de E2 et plus de 600 E3 ligases et l'on sait que le couple E2-E3 définit la spécificité de reconnaissance du substrat à ubiquitine. Notons que le nombre des E3 ligases est voisin de celui des kinases ce qui illustre le fait que l'ubiquitination et la phosphorylation sont deux modifications aussi largement répandues l'une que l'autre (Komander, 2009).

Une seule molécule d'ubiquitine, plusieurs monomères d'ubiquitine, ou une chaîne de polyubiquitine peuvent être associées à des protéines. Ces différents types d'ubiquitination sont appelés respectivement ; monoubiquitination, multiubiquitination et polyubiquitination (Behrends & Harper, 2011). Les chaînes de polyubiquitine sont ainsi formées par des molécules successives d'ubiquitine qui sont elles-mêmes ubiquitinées. Les 7 lysines d'une molécule d'ubiquitine ainsi que son extrémité N-terminale peuvent être utilisées pour former plusieurs types de chaînes de polyubiquitine, qui présentent des conformations



**Figure 19. Conformation adoptée par les différents types de chaînes de polyubiquitine.**

Représentation de la conformation de chaînes de quatre ubiquitines liées par les Met1, Lys48, Lys63, et Lys11. Les chaînes liées par les Lys11 et Lys48 adoptent des structures compactes alors que les chaînes linéaires ou liées par la Lys63 sont allongées. Pour ces deux derniers types de chaînes, l'ubiquitine proximale (celle dont l'extrémité C-terminale est libre ou liée à un substrat) et l'ubiquitine distale (à l'opposé de l'ubiquitine proximale) sont représentées. Les références PDB des différentes structures sont : Met1, 2W9N ; Lys48, 2O6V ; Lys63, 2JF5 et Lys11, 2XEW). La figure est adaptée de : (Behrends & Harper, 2011).

différentes et qui sont associées à divers signaux fonctionnels comme nous allons le voir (Figure 19) (Behrends & Harper, 2011; Bremm & Komander, 2011). De plus, au sein d'une chaîne de polyubiquitine, toutes les molécules d'ubiquitines peuvent être liées sur la même lysine pour former une chaîne homogène mais ce n'est pas une obligation. En effet, il existe des chaînes de polyubiquitine qui comportent différents types de liaisons ce qui donne naissance à des chaînes hétérogènes linéaires ou à des chaînes branchées (Bremm & Komander, 2011; Komander, 2009).

Même si le monde de l'ubiquitine a été considérablement étudié depuis 30 ans, une comparaison entre la vision que l'on avait de ce domaine de recherche en 2000 et celle que l'on a aujourd'hui indique que de nombreuses avancées ont été réalisées durant la dernière décennie. La comparaison est adaptée de : (Grabbe et al, 2011).

**En 2000 :**

- Il était admis que les génomes eucaryotes codaient une seule enzyme E1, une vingtaine d'E2 et une centaine d'E3 ligases.
- La fonction des domaines de type RING et HECT était associée avec l'ubiquitination.
- On pensait que la spécificité de substrat était principalement réalisée par les E3 ligases.
- Les seuls domaines de liaison à l'ubiquitine qui étaient bien caractérisés étaient les domaines UBA (ubiquitin-associated).

- Seules des liaisons impliquant les Lys6, Lys11, Lys29, Lys48 and Lys63 de l'ubiquitine avaient été observées.
- La famille des molécules ressemblant à l'ubiquitine contenait 8 membres : NEDD8, SUMO1 (small ubiquitin-related modifier 1), SUMO2, SUMO3, ISG15, Atg8 (autophagy-related 8), Atg12 and Fub1.

**En 2010 :**

- On estime que les eucaryotes supérieurs possèdent deux enzymes E1, environ 40 E2 et plus de 600 E3 ligases (Komander, 2009).
- On pense que la spécificité de substrat est réalisée par le couple E2-E3, et que d'autres partenaires ou des modifications post-traductionnelles peuvent réguler l'ubiquitination de substrats (Komander, 2009).
- Un très grand nombre de domaines pouvant lier l'ubiquitine a été caractérisé et ceux-ci reconnaissent spécifiquement différentes surfaces de l'ubiquitine ou différents types de chaînes de polyubiquitine (Dikic et al, 2009).
- On sait que l'ensemble des lysines de l'ubiquitine (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63) peuvent être ubiquitinées. D'autre part, des chaînes linéaires, qui sont formées entre la glycine 76 et la méthionine 1 de deux molécules d'ubiquitine, ainsi que des chaînes présentant des branchements ou combinant différents types de liaisons ont été décrites.
- La famille des molécules ressemblant à l'ubiquitine s'est agrandie et elle contient 3 nouveaux membres qui sont : Urm1 (ubiquitin-related modifier 1), UFM (ubiquitin-fold modifier) et FAT10.

## **II. Une multitude de fonctions exercées par l'ubiquitine**

La fixation d'une petite protéine plutôt que d'un groupement chimique lors de l'ubiquitination d'une protéine permet de créer une surface importante sur laquelle, différents récepteurs spécifiques peuvent interagir. De plus, la combinaison de la monoubiquitination avec les différents types de polyubiquitination permet à l'ubiquitine de contrôler de multiples voies fonctionnelles (Wilkinson, 2009).

Le type d'ubiquitination qui a été le plus étudié et qui est le mieux compris est le marquage de protéines par une chaîne de polyubiquitine liées par la Lys48, la Lys 29 ou la Lys11 en vue de leur dégradation par le protéasome. La découverte du lien entre ces

chaînes de polyubiquitine et la dégradation des protéines par le protéasome a valu à Avram Hershko, à Aaron Ciechanover et à Irwin Rose de recevoir le Prix Nobel de chimie en 2004 (Giles, 2004; Wilkinson, 2009).

D'autre part, il a été montré que d'autres types d'ubiquitination ne correspondaient à pas un signal de dégradation mais jouaient un rôle de signalisation dans diverses voies fonctionnelles. Par exemple, des chaînes de polyubiquitine liées par la Lys63 sont impliquées dans la réparation de l'ADN, l'activation de kinase, la régulation de l'endocytose, ou la voie de signalisation de NF- $\kappa$ B (Behrends & Harper, 2011; Pickart & Fushman, 2004) et de façon intéressante, dans le cas de la voie de signalisation de NF- $\kappa$ B d'autres types de chaînes d'ubiquitines jouent un rôle dans la transduction du signal (Behrends & Harper, 2011; Grabbe et al, 2011). Au sujet de la polyubiquitination, on peut noter que des études récentes ont mis en évidence que des chaînes présentant différents types de liaisons pouvaient être attachées à une protéine. La fonction associée à nombre de ces chaînes hétérogènes reste encore à déterminer mais des progrès dans ce domaine ont eu lieu, notamment concernant des chaînes liées à la fois par la Lys11 et par une autre lysine (Bremm & Komander, 2011).

Par ailleurs, la monoubiquitination de protéines n'est jamais un signal de dégradation mais participe à la régulation de très nombreuses voies de signalisation. Par exemple, la monoubiquitination des histones H2A et H2B influence la structure de la chromatine ainsi que la transcription. En effet, la monoubiquitination de H2B est mise en place sur le corps des gènes transcrits, alors que la monoubiquitination de H2A est synthétisée par le complexe PRC1 (Polycomb Repressive Complex 1) et est une marque de répression de la transcription (cf. Partie sur l'ubiquitination de H2B) et (Weake & Workman, 2008). On peut enfin noter que de nombreuses évidences indiquent un rôle de l'ubiquitination dans la formation de mRNP (mRNA ribonucleoprotein) et le contrôle de leur qualité en vue de l'export des ARNm vers le cytoplasme (Babour et al, 2012).

## **B. Mécanismes de régulation des USP**

### **I. Cinq familles de déubiquitinasés**

L'activité E3 ligase catalyse la formation d'une liaison isopeptidique entre le groupement carboxyl de la glycine C-terminale de l'ubiquitine et le groupement  $\epsilon$ -amine d'une lysine, mais l'ubiquitination est un signal très dynamique et très peu stable. En effet, l'activité de déubiquitination permet le clivage de ce type de liaison isopeptidique (Nijman et al, 2005). Une analyse bioinformatique du génome humain a permis de mettre en évidence près de 80 gènes codant des déubiquitinasés (DUB) qui peuvent être regroupés en 5 familles de gènes : les Ubiquitin Specific Proteases (USP), les Ubiquitin C-terminal Hydrolases (UCH), les Ovarian Tumor Proteases (OTU), les Josephins et enfin les JAB1/MPN/MOV34 métalloenzymes (JAMM) (Komander et al, 2009; Nijman et al, 2005). La famille USP qui contient le plus de membres, ainsi que les familles UCH, OTU et Josephins sont des protéases à cystéine alors que les DUB de type JAMM sont des métalloprotéases à zinc (Komander et al, 2009; Nijman et al, 2005). Bien que la fonction de chacune de ces DUB ne soit pas connue avec précision, beaucoup d'études ont démontré l'importance fonctionnelle d'un grand nombre de ces enzymes. Par exemple, des protéines polyubiquitinées et envoyées vers le protéasome peuvent être déubiquitinées de façon spécifique par certaines DUB ce qui va éviter leur dégradation. D'autres DUB vont contrôler la quantité d'ubiquitine libre dans la cellule, en déubiquitinant des protéines au moment de leur dégradation pour permettre le recyclage de l'ubiquitine, ou en désassemblant les chaînes de polyubiquitine qui ne sont plus conjuguées à un substrat. En plus d'un rôle dans le taux de renouvellement des protéines ou dans l'homéostasie de l'ubiquitine, il a été montré que la déubiquitination de certaines protéines était liée à la régulation de processus cellulaires comme la transcription, la réparation de l'ADN, le contrôle du cycle cellulaire ou l'endocytose. Pour des revues sur la fonction des USP cf : (Atanassov et al, 2011; Nijman et al, 2005; Reyes-Turcu et al, 2009; Ventii & Wilkinson, 2008; Wilkinson, 2009).

Durant ma thèse, j'ai mis en évidence que USP22, la DUB du complexe SAGA humain (Zhang et al, 2008; Zhao et al, 2008), était régulée par des interactions avec d'autres sous-unités de SAGA (cf. Partie Résultats). Dans cette partie, j'ai voulu décrire différents mécanismes de régulation qui permettent à des enzymes de la famille USP

d'exercer de façon optimale leur fonction en déubiquitinant efficacement et spécifiquement leur substrat.

## **II. Caractéristiques générales des USP**

### **1. Diversification au cours de l'évolution**

Avec près de 60 membres, la classe des USP représente environ les deux tiers des DUB codées par le génome humain. De façon intéressante, seules une quinzaine d'USP sont présentes chez *S. cerevisiae* (appelées généralement Ubp dans cet organisme) ce qui indique qu'une forte augmentation du nombre des USP a eu lieu au cours de l'évolution. Une telle tendance est également observée pour les E3 ligases (le troisième facteur dans la cascade enzymatique de l'ubiquitination qui définit la spécificité de substrat) ce qui suggère une co-évolution entre ces deux groupes d'enzymes (Nijman et al, 2005; Semple, 2003). Cette diversification des USP illustre également une probable spécificité fonctionnelle de ces enzymes et suggère qu'une USP peut reconnaître spécifiquement un substrat donné pour le déubiquitiner dans des conditions spatio-temporelles définies. Nous allons voir que des domaines additionnels au domaine catalytique ainsi que des partenaires protéiques participent à rendre l'activité des USP spécifique.

### **2. Association à des complexes protéiques**

En 2009, Sowa et coll. ont analysé par spectrométrie de masse des purifications systématiques de l'ensemble des DUB codées par le génome humain ce qui a permis de démontrer que 75% d'entre elles étaient associées de façon stable à des partenaires protéiques (Sowa et al, 2009). Une étude similaire de protéomique a montré que la majorité des 20 DUB de la levure *S. pombe* sont également retrouvées au sein de complexes ce qui est requis pour leur permettre d'exercer leur fonction (Kouranti et al, 2010). Cette intégration au sein de complexes est en effet cruciale pour stimuler l'activité de plusieurs USP et/ou pour réguler la localisation cellulaire de certaines DUB comme nous le verrons par la suite (cf. B. III. 1. et 2., et IV. 1.).

### **3. Présence de domaines additionnels**

En plus d'un domaine catalytique, une très large majorité d'USP possède un ou plusieurs domaines additionnels qui interagissent, par exemple, avec des partenaires protéiques ou des substrats de ces DUB. Notons que de façon intéressante, de nombreux domaines connus pour lier l'ubiquitine comme des domaines UIM (ubiquitin-interacting motif),

des domaines UBA (Ubiquitin-associated domain) ou encore certains domaines ZnF-UBP (zinc finger ubiquitin-specific protease domain) sont retrouvés au sein de nombreuses USP (cf. B. IV. 2.). Pour des revues sur les domaines de liaison de l'ubiquitine cf : (Dikic et al, 2009; Harper & Schulman, 2006; Hurley et al, 2006; Winget & Mayor, 2010). D'autre part, près d'un tiers des USP présente un ou plusieurs domaines Ubl (Ubiquitin-like fold). La conservation en séquence entre ces différents domaines Ubl est relativement faible, mais ils adoptent tous une structure tridimensionnelle voisine de celle de l'ubiquitine d'où leur nom. Cependant, l'absence d'un motif Gly-Gly au niveau de leur extrémité C-terminale protège ces domaines Ubl d'un clivage aspécifique par une DUB. Plusieurs études récentes ont mis en évidence que l'activité catalytique de certaines USP était contrôlée par leurs domaines Ubl et nous verrons dans un prochain paragraphe que les mécanismes de régulation exercés par ces domaines peuvent varier d'une USP à l'autre (cf B. III.) et (Komander et al, 2009).

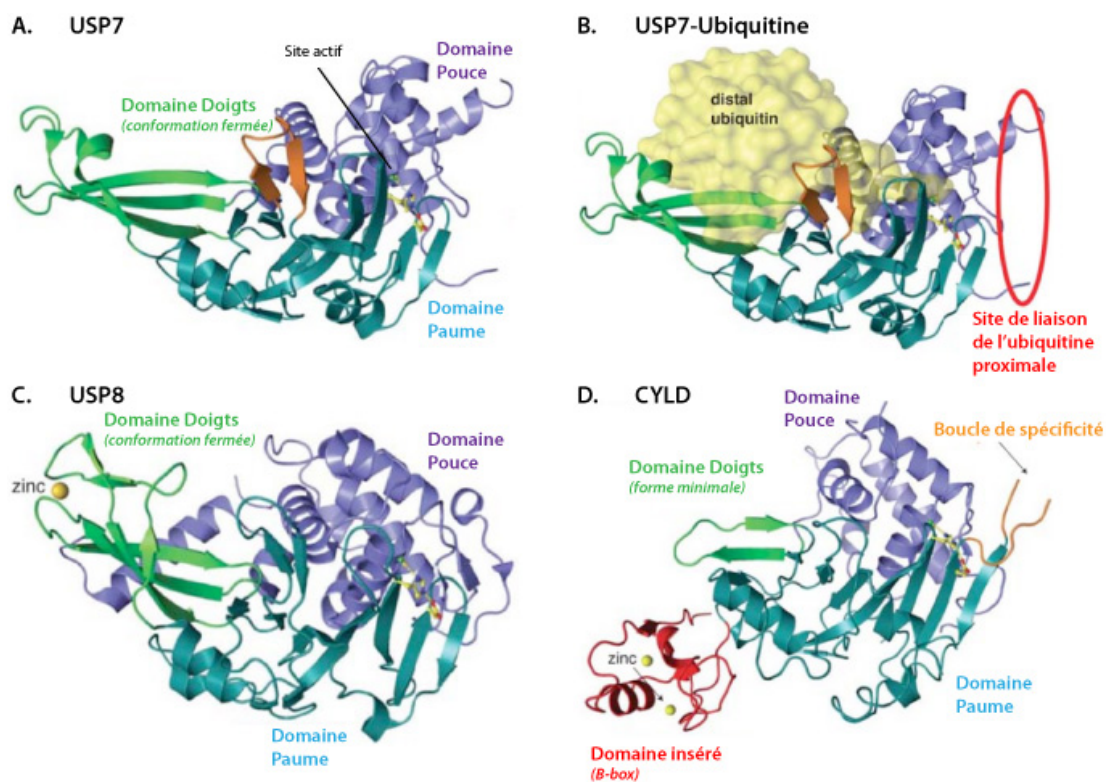
Dans la séquence des USP, les domaines additionnels peuvent se situer en N ou en C-terminal du domaine catalytique mais un quart des domaines additionnels sont insérés directement au sein du domaine catalytique. En générant un alignement multiple à partir des séquences des USP codées par les génomes humains et de *S. cerevisiae* et en utilisant des informations structurales sur le domaine catalytique USP, une étude a indiqué que l'ensemble de ces domaines USP pouvait être découpé en 6 motifs conservés. De façon intéressante, des insertions de séquence peuvent avoir lieu entre ces 6 motifs sans altérer le repliement structural du domaine USP ce qui permet sans doute de positionner de façon optimale les domaines additionnels par rapport au domaine catalytique des USP en fonction du rôle qu'ils ont à jouer (Ye et al, 2009).

#### **4. Description structurale du domaine USP**

##### Repliement structural conservé

Le domaine catalytique des USP est caractérisé par deux petits motifs très conservés qui sont appelés boîtes Cys et His et qui comprennent les résidus catalytiques ; une cystéine, une histidine et un aspartate (Nijman et al, 2005). Depuis la cristallisation du domaine catalytique de USP7 en 2002, la structure de nombreux autres domaines catalytiques d'USP a été résolue et il en ressort que tous ces domaines possèdent un repliement structural commun [USP7 (Hu et al, 2002), USP14 (Hu et al, 2005), USP8 (Avvakumov et al, 2006), USP2 (Renatus et al, 2006), CYLD (Komander et al, 2008), Ubp8 (Kohler et al,

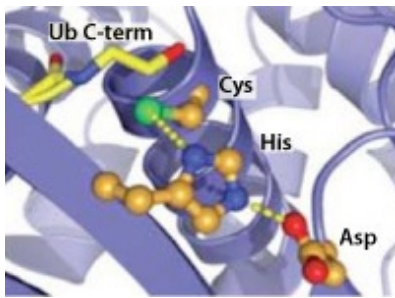
2010; Samara et al, 2010), USP4 (Luna-Vargas et al, 2011), USP21 (Ye et al, 2011), USP5 (Avvakumov et al, 2012)]. Une analogie avec la structure de la main permet de décrire ce repliement structural en définissant trois régions appelées Paume, Pouce et Doigts (Figure 20). Le site catalytique se situe à l'interface entre la Paume, qui contient la boîte Cys, et le Pouce, qui contient la boîte His, alors que les Doigts forment une poche dans laquelle vient se fixer l'ubiquitine proximale qui est liée sur une lysine d'un substrat à déubiquitiner (Figure 20 B.) (Hu et al, 2002). On peut noter que la Paume et le Pouce présentent une homologie structurale avec la classe des protéases à cystéines ressemblant à la papaïne, cependant, l'insertion des Doigts donne une spécificité



**Figure 20. Structure du domaine catalytique de différentes USP.**

Les domaines Pouce, Paume et Doigts de diverses USP sont représentés respectivement en mauve, en bleu et en vert. Les chaînes latérales des résidus catalytiques ainsi que les éventuels atomes de zinc (sphères jaunes) sont mis en évidence. **A.** La structure de USP7 est présentée sans ou avec **B.** de l'ubiquitine (référence PDB 1nbf). L'ubiquitine distale se situe dans le site catalytique et interagit principalement avec le domaine Doigts. Le site putatif de liaison de l'ubiquitine proximale est indiqué en rouge. **C.** La structure du domaine catalytique de USP8 indique que les Doigts sont repliés sur eux-mêmes vers le domaine Pouce, ce qui bloque la venue de l'ubiquitine dans le site catalytique de l'enzyme (référence PDB 2gfo). Notons que les domaines Doigts d'un certain nombre d'USP sont stabilisés par un atome de zinc. **D.** Structure du domaine catalytique de CYLD, la seule USP connue pour être spécifique des chaînes liées par la Lys63 ou des chaînes linéaires (référence PDB 2vhf). L'absence d'un domaine Doigts classique permet à CYLD de cliver des chaînes liées par la Lys63 en leur milieu, et une longue boucle (en orange) située non loin du site actif participe à la reconnaissance spécifique du substrat. Contrairement à d'autres domaines catalytiques d'USP, celui de CYLD présente une conformation active en l'absence d'ubiquitine. CYLD contient également un doigt de zinc B-box qui est inséré dans le domaine catalytique et dont la fonction serait de participer à la localisation cytoplasmique de l'enzyme. La figure est adaptée de : (Komander et al, 2009)





**Figure 21.**

**Vue détaillée du site catalytique de USP7.**

Les atomes des chaînes latérales des résidus catalytiques sont représentés comme des boules et les interactions entre les trois résidus catalytiques sont indiquées par des pointillés jaunes. L'extrémité C-terminale de l'ubiquitine est dessinée en jaune (référence PDB 1nbf). La figure est adaptée de : (Komander et al, 2009)

structurale au domaine catalytique des USP. Notons également que les Doigts dérivent d'un repliement de type « zinc ribbon » mais que beaucoup d'entre eux ne sont plus stabilisés par un atome de zinc (Krishna & Grishin, 2004).

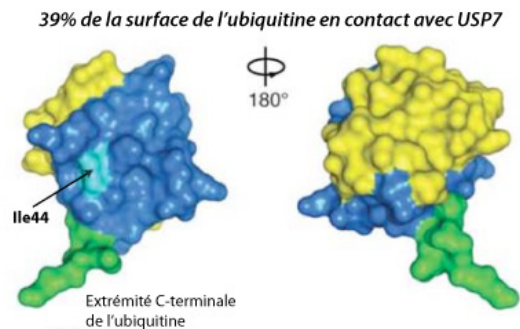
Mécanisme de catalyse

Les USP sont des cystéines protéases et la catalyse repose principalement sur trois résidus cruciaux qui forment une triade catalytique (Figure 21). Le principe de la

catalyse repose sur l'abaissement du pKa de la cystéine catalytique par la chaîne latérale d'une histidine voisine ce qui lui permet de réaliser une attaque nucléophile sur la liaison isopeptidique à cliver. Un troisième résidu, et dans certains cas un quatrième, (habituellement une asparagine et/ou un aspartate) permet de bien positionner et de polariser l'histidine catalytique. (Komander et al, 2009; Samara et al, 2010).

Mode de liaison de l'ubiquitine

On peut remarquer que près de 40% de la surface de l'ubiquitine, comprenant la zone hydrophobe autour de l'isoleucine 44, sont impliqués dans l'interaction avec le domaine USP (Figure 22) alors que seuls 10% de la surface de l'ubiquitine sont généralement reconnus par les domaines de liaison à l'ubiquitine (Komander et al, 2009; Winget & Mayor, 2010). D'autre part, la région C-terminale de l'ubiquitine se retrouve maintenue fermement dans un canal menant vers le site actif des USP

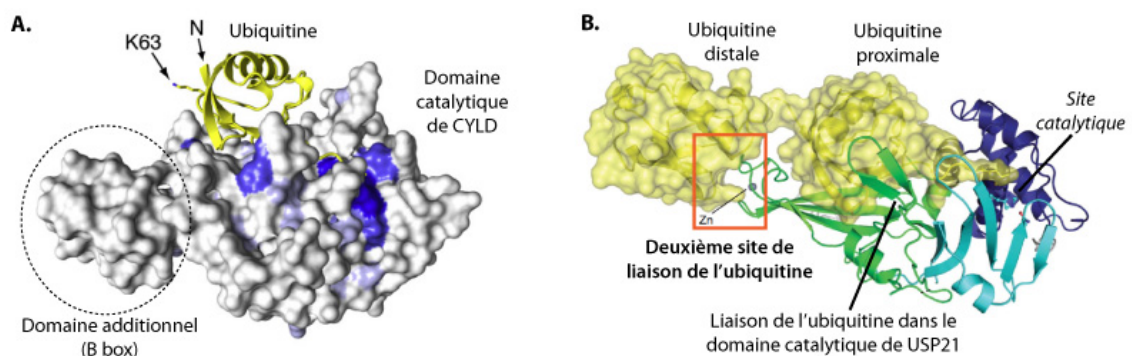


**Figure 22. Interaction de l'ubiquitine avec USP7.**

Les résidus de l'ubiquitine qui interagissent directement avec le domaine catalytique de USP7 sont colorés en bleu. Les quatre résidus de l'extrémité C-terminale de l'ubiquitine sont en verts. L'image de gauche est centrée sur le patch hydrophobe de l'isoleucine 44 (en bleu clair) alors que l'image de droite montre la face opposée de l'ubiquitine. Ceci illustre que le patch hydrophobe de l'Ile 44 est impliqué dans l'interaction avec le domaine catalytique de USP7. On peut noter que l'interaction entre l'ubiquitine et USP7 implique près de 40% de la surface de l'ubiquitine alors que de nombreux domaines de liaison à l'ubiquitine n'interagissent qu'avec environ 10% de la surface de la molécule. La détermination des résidus de l'ubiquitine en contact avec USP7 a été réalisée grâce au serveur PISA ([http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)). La figure a été adaptée de : (Komander et al, 2009)

ce qui est important pour la lyse de la liaison isopeptidique. Ces interactions entre l'extrémité C-terminale de l'ubiquitine proximale (l'ubiquitine liée au substrat) et le canal menant vers le site actif sont aussi importantes pour discriminer l'ubiquitine d'autres molécules comme par exemple SUMO (small ubiquitin-like modifier) ou NEDD8 (neuronal precursor cell expressed developmentally downregulated protein 8) qui possèdent le même repliement structural que l'ubiquitine mais des séquences C-terminales différentes (Komander et al, 2009; Ye et al, 2011).

Il apparaît également que la poche de liaison de l'ubiquitine d'un domaine USP est un élément important pour expliquer que certaines USP, par exemple CYLD (Cylindromatosis-associated DUB), possèdent une activité endo-DUB, c'est à dire qu'elles peuvent cliver une chaîne de polyubiquitine en son milieu, alors que d'autres USP, comme USP14, possèdent une activité exo-DUB car elles ne peuvent couper que l'ubiquitine distale (cette ubiquitine se situe à l'opposé de l'ubiquitine proximale dont l'extrémité C-terminale est liée au substrat). Dans le cas de USP14, les Doigts du domaine catalytique entrent en contact avec près de 40% de la surface de l'ubiquitine incluant les lysines 48 et 63 et empêchant donc USP14 d'interagir avec une molécule d'ubiquitine située au milieu d'une chaîne (Hu et al, 2005; Komander et al, 2009). Par contre, CYLD possède des Doigts beaucoup plus petits que ceux des autres USP et un modèle suggère que la lysine 63 (mais pas la lysine 48) d'une molécule d'ubiquitine



**Figure 23. Liaison de l'ubiquitine par deux endo-DUB.**

**A.** Modèle du complexe formé entre une molécule d'ubiquitine (en jaune) et le site catalytique de l'USP CYLD (en blanc, les parties bleues correspondant à des zones conservées avec d'autres USP). Dans ce modèle, la Lys 63 de l'ubiquitine liée au site catalytique de CYLD est accessible, ce qui permet à l'enzyme d'interagir et de cliver en son milieu une chaîne de polyubiquitine liée par la Lys 63 (activité endo-DUB). **B.** La structure du domaine catalytique de USP21 (domaine Doigts (en vert), Paume (en bleu ciel) et Pouce (en bleu marine)) avec une chaîne linéaire de diubiquitine (en jaune) révèle deux sites de liaison de l'ubiquitine sur ce domaine catalytique. En plus du site de liaison de l'ubiquitine présent au sein du domaine catalytique de toutes les USP, l'extrémité des Doigts de USP21 peut interagir avec une deuxième molécule d'ubiquitine ce qui permet à l'enzyme d'exercer une activité endo-DUB. La figure est adaptée de : (Komander et al, 2008; Ye et al, 2011).

interagissant avec le domaine catalytique de CYLD est accessible, ce qui permet à l'enzyme de lier et de cliver une chaîne de polyubiquitine de type K63 en son milieu (Figure 23 A.) (Komander et al, 2009; Komander et al, 2008).

Enfin, une structure récente indique que USP21 possède une activité endo-DUB car son domaine catalytique peut interagir avec deux molécules d'ubiquitines. En effet, une structure de ce domaine catalytique obtenue avec une chaîne linéaire de deux molécules d'ubiquitines indique qu'en plus du domaine classique de liaison de l'ubiquitine, le domaine catalytique de USP21 possède un second site pour lier une molécule d'ubiquitine qui est située au niveau de l'extrémité du doigt de zinc des Doigts (Figure 23 B.) (Ye et al, 2011).

### **5. Equilibre entre formes actives et inactives**

De façon intéressante, plusieurs études structurales ont révélé que les cristaux des domaines catalytiques de différentes USP obtenus en l'absence d'ubiquitine correspondaient à un état inactif de ces enzymes. Dans le cas de USP7, cela est dû à une mauvaise conformation des résidus catalytiques (Hu et al, 2002) alors que pour USP14, les résidus catalytiques sont bien orientés mais l'état inactif est causé par la présence de boucles qui viennent obstruer le site catalytique de l'enzyme (Hu et al, 2005). D'autre part, pour USP8, le repliement des Doigts sur eux même empêche l'ubiquitine du substrat de venir interagir avec le domaine catalytique (Avvakumov et al, 2006). Enfin, on peut noter que dans le cas de USP4, le site catalytique de l'enzyme est obstrué à la fois par trois boucles ainsi que par un repliement des Doigts sur eux mêmes (Luna-Vargas et al, 2011).

Il est important de noter que ces formes inactives ne sont pas figées et que la conformation des Doigts ainsi que les boucles obstruant le site actif présentent une certaine flexibilité (Luna-Vargas et al, 2011). De plus, la co-cristallisation des domaines catalytiques de USP7 ou USP14 avec de l'ubiquitine aldéhyde, un substrat suicide qui peut interagir de façon stable avec le domaine catalytique des USP et qui mime le comportement de l'ubiquitine, permet de stabiliser ces domaines USP dans un état actif. Un équilibre entre les formes actives et inactives de ces enzymes doit donc exister et la stabilisation de l'un ou l'autre de ces états offre un moyen de réguler l'activité de certaines USP (Faesen et al, 2011; Hu et al, 2002; Hu et al, 2005).

Notons cependant que l'existence de conformation inactive n'est pas une règle générale pour toutes les USP car par exemple, la structure du domaine catalytique de

l'USP CYLD en absence d'ubiquitine indique que celui-ci se situe dans une conformation active (Komander et al, 2008).

## **6. Divers modes de régulations**

L'activité de nombreuses USP est stimulée de façon allostérique par des interactions avec des partenaires spécifiques et je vais présenter dans les prochains chapitres plusieurs exemples qui illustrent ce type de régulation (cf B. III. et IV.) et (Komander et al, 2009). Cependant, il existe d'autres mécanismes de régulation que je voudrais rapidement lister.

Tout d'abord, certaines USP possèdent des éléments structuraux qui leur permettent de cliver spécifiquement un type de chaîne d'ubiquitines donné. Par exemple, l'USP CYLD présente une boucle à proximité de son site actif, qui n'est pas retrouvée dans d'autres USP et qui est nécessaire à la lyse efficace de chaînes de polyubiquitine de type K63 *in vitro* (Figure 20 D.) (Komander et al, 2008). Cependant, il semblerait que ce mécanisme soit plus une exception qu'une généralité car *in vitro* une grande majorité d'USP est capable de cliver différents types de chaînes de polyubiquitine (Komander et al, 2009).

Un autre mode de régulation vise à contrôler la localisation cellulaire des USP. Par exemple, un doigt de zinc de type B-box permet de maintenir l'enzyme CYLD dans le cytoplasme (Komander et al, 2008). Des modifications post-traductionnelles peuvent également réguler l'activité enzymatique de certaines USP. Il a notamment été montré que la phosphorylation de CYLD inhibe sa capacité à déubiquitiner TRAF2 (TNF receptor-associated factor 2) ce qui supprime la voie de signalisation NF- $\kappa$ B (Komander et al, 2009).

## **III. Différentes régulations par les domaines Ubl**

De nombreuses USP possèdent un domaine Ubl et nous allons voir que ceux-ci peuvent réguler de différentes manières l'activité catalytique de ces DUB. Tout d'abord, je vais décrire les résultats montrant que le domaine Ubl de USP14 est nécessaire pour que l'enzyme interagisse avec le protéasome et soit activée. Puis nous verrons que les 5 domaines Ubl de USP7 stabilisent la conformation active de l'enzyme et enfin nous verrons que le domaine Ubl de USP4 inhibe l'enzyme en entrant en compétition avec le

substrat pour interagir avec la poche de liaison de l'ubiquitine de son domaine catalytique.

### **1. Régulation de Ubp6/USP14 par le protéasome**

#### Le protéasome

Chez les eucaryotes, le marquage d'une protéine par une chaîne de plus de quatre ubiquitines liées par la lysine 48 est un signal pour la dégradation de cette protéine par le protéasome, une grosse machinerie cellulaire d'une taille d'environ 1,5 MDa ou 26S. Le protéasome, qui contient une trentaine de sous-unités, est constitué par un cœur structural cylindrique, appelé particule de protéolyse 20S, qui est entouré à chaque extrémité par une particule de régulation 19S. La particule 20S possède les activités protéolytiques qui vont digérer les protéines alors que la particule de régulation va interagir avec les substrats du protéasome via leur chaîne de polyubiquitine, puis va les déubiquitiner, les déplier de façon ATP-dépendante et finalement les transférer dans le cylindre central en vue de leur dégradation (Finley, 2009).

#### Identification de USP14/Ubp6 comme une sous-unité du protéasome

Un membre de la famille des USP est associée au protéasome : il s'agit de Ubp6 chez la levure ou de USP14 chez les mammifères. Comme l'interaction entre ces déubiquitinasés et le protéasome est de relativement faible intensité, Ubp6 et USP14 n'ont pas été identifiées lors des premières caractérisations du protéasome, mais des approches de purifications moins stringentes ont permis dans un second temps de les définir comme des sous-unités du protéasome (Borodovsky et al, 2001; Leggett et al, 2002). On peut noter que chez les mammifères, c'est l'utilisation d'ubiquitine vinyl sulfone (UbVS) qui a permis de montrer que USP14 est une sous-unité du protéasome. L'UbVS est un substrat suicide qui peut inhiber de façon irréversible la forme active de nombreuses déubiquitinasés des familles USP ou UCH en se liant covalamment à leur site actif. Après une électrophorèse de type SDS-PAGE, la liaison d'une déubiquitinasé à l'UbVS peut être révélée par un déplacement vers le haut du gel d'environ 8 kDa (la taille de l'ubiquitine) de la forme liée à l'UbVS par rapport à la forme libre de l'enzyme (Borodovsky et al, 2001). Dans le cas de USP14, un extrait protéique de cellules murines a été fractionné par une chromatographie d'exclusion - inclusion, puis chaque fraction a été incubée avec de l'UbVS radioactif et la radioactivité associée avec des déubiquitinasés a été détectée après un gel SDS-PAGE. De façon intéressante, dans plusieurs fractions contenant des

complexes de très haut poids moléculaires, deux bandes correspondant à des déubiquitinasés de petite taille ont été détectées suggérant que ces enzymes étaient associées avec un gros complexe. La bande la plus intense correspond à la déubiquitinasé UCH37 qui était déjà connue pour être une sous-unité du protéasome chez les mammifères (Lam et al, 1997). Des analyses supplémentaires ont indiqué que la seconde bande correspondait à USP14 et que cette déubiquitinasé était également une sous-unité du protéasome (Borodovsky et al, 2001). Cette étude décrit la première utilisation de l'UbVS, un composé que nous avons également utilisé pour analyser l'activité enzymatique de USP22 (cf. Partie Résultats).

En plus de leur domaine catalytique, Ubp6 et USP14 possèdent un domaine Ubl dans leur région N-terminale (Figure 24 A.) qui leur permet d'interagir avec la sous-unité Rpn1 de la particule 19S du protéasome (Hu et al, 2005; Leggett et al, 2002). D'autre part, des souches de levure délétées de la protéine Ubp6 entière ou seulement du domaine Ubl présentent les mêmes phénotypes ce qui montre que ce domaine est requis pour la fonction de la protéine Ubp6 *in vivo* (Chernova et al, 2003; Leggett et al, 2002).

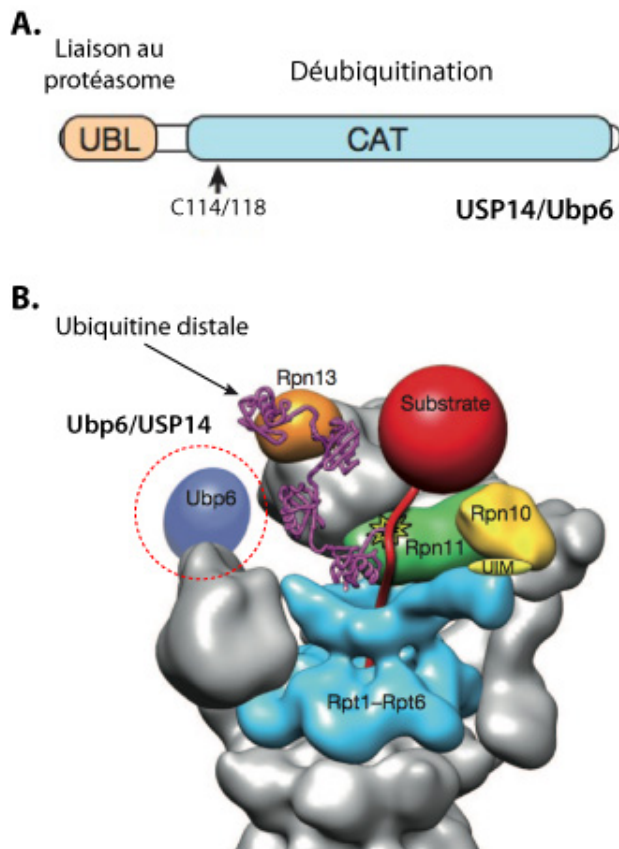
#### Régulation enzymatique de USP14 par le protéasome

Remarquons que les résultats de plusieurs tests de déubiquitination *in vitro* montrent que la liaison au protéasome de Ubp6 ou USP14 par leur domaine Ubl permet de stimuler l'activité de ces enzymes de façon très importante (Hu et al, 2005; Lee et al, 2010; Leggett et al, 2002). Ces résultats fonctionnels sont en plein accord avec l'étude cristallographique qui a déterminé la structure tridimensionnelle du domaine catalytique de USP14 (Hu et al, 2005). En effet, deux structures de ce domaine catalytique ont été obtenues : l'une en l'absence et l'autre en présence d'ubiquitine aldéhyde. La première structure illustre une forme inactive de l'enzyme alors que la seconde correspond à une forme active de USP14. La comparaison de ces deux structures a donc permis d'identifier les différences structurales qui existent entre les formes active et inactive de l'enzyme. Dans les deux cas, les résidus catalytiques sont positionnés de façon optimale, cependant, en l'absence d'ubiquitine aldéhyde, deux boucles viennent obstruer le canal menant au site actif dans lequel doit se fixer l'extrémité C-terminale de l'ubiquitine (Hu et al, 2005). Les auteurs de l'étude spéculent donc que la liaison de Ubp6/USP14 au protéasome puisse faciliter le déplacement de ces

boucles et ainsi permettre à l'enzyme de lier de façon beaucoup plus rapide ses substrats (Hu et al, 2005).

Fonction de USP14 au sein du protéasome

En plus de USP14, la particule de régulation 19S contient la déubiquitinase RPN11, qui



**Figure 24. Localisation de Ubp6/USP14 au sein de la particule 19S du protéasome.**

**A.** Organisation en domaines de USP14 ou de son homologue Ubp6 chez la levure *S. cerevisiae*. **B.** Modèle représentant l'interaction entre un substrat polyubiquitiné et la particule 19S du protéasome de levure. La structure de la particule 19S a été obtenue par microscopie électronique et celle de la chaîne de 4 ubiquitines liées par la Lys 48 (en mauve) par cristallographie (référence PDB 2KDE). Cette chaîne de polyubiquitine est fixée sur un substrat (en rouge) et interagit avec Rpn13, un récepteur à ubiquitine du protéasome (en orange). Ubp6 (en bleu marine et entourée en pointillés rouges) se situe à proximité de l'ubiquitine distale de la chaîne et peut donc cliver celle-ci pour éventuellement permettre le relargage du substrat. Dans le cas inverse, la DUB Rpn11 (en vert, son site actif est représenté par une étoile) peut cliver la liaison isopeptidique entre le substrat et sa chaîne de polyubiquitine et celui-ci va être déplié au niveau de la base de la particule 19S (bleu ciel) avant d'être dégradé dans la particule 20S. La figure est adaptée de : (Lander et al, 2012).

appartient à la famille des JAMM protéases, et dont la fonction est de cliver la liaison isopeptidique entre le substrat et sa chaîne de polyubiquitine au moment où celui-ci est déplié par des ATPases puis transféré dans la particule 20S en vue de sa dégradation (Lee et al, 2011b). Enfin, on peut noter que chez les métazoaires, une troisième déubiquitinase, UCH37, de la famille des Ubiquitine C-terminale Hydrolases, est aussi retrouvée au sein de la particule 19S (Lam et al, 1997).

La perte de la fonction de USP14 ou de son homologue Ubp6 cause de nombreux phénotypes, mais la fonction exacte jouée par ces enzymes au niveau du protéasome n'est pas encore clairement comprise. Parmi les phénotypes observés, il a été montré que la délétion de Ubp6 chez *S. cerevisiae* provoque une diminution de la quantité d'ubiquitine libre dans la cellule

(Chernova et al, 2003; Hanna et al, 2003). D'autre part, des souris homozygotes pour une mutation causant une diminution de l'expression de USP14 présentent un défaut dans la transmission synaptique et possèdent moins d'ubiquitine libre dans de nombreux tissus (Anderson et al, 2005; Wilson et al, 2002). On peut également noter qu'une étude suggère que USP14 pourrait réguler le degré d'ouverture de la particule 20S du protéasome (Peth et al, 2009). De façon intéressante, il a aussi été observé que la délétion de Ubp6 chez la levure ou qu'une inhibition de l'activité de USP14 par une drogue spécifique ou par un ARN interférent accélèrent la dégradation de protéines polyubiquitinées *in vivo*, ce qui suggère que USP14 agirait comme un inhibiteur du protéasome (Hanna et al, 2006; Koulich et al, 2008; Lee et al, 2010). Ces résultats ont donc permis de proposer un modèle qui suggère que USP14 pourrait raccourcir les chaînes de polyubiquitine de certaines protéines liées au protéasome depuis leur extrémité distale. La diminution de la taille de ces chaînes conduirait alors à un affaiblissement de l'interaction entre ces chaînes et les récepteurs à ubiquitine de la particule 19S ce qui favoriserait leur dissociation du protéasome (Lee et al, 2011b). Notons enfin que la structure obtenue par microscopie électronique de la particule 19S du protéasome de levure indique que Ubp6 est localisée au niveau de la périphérie de cette particule ce qui semble en cohérence avec une activité dirigée contre l'extrémité distale des chaînes de polyubiquitine (Figure 24 B.) (Lander et al, 2012).

Cette fonction présumée de USP14 explique donc l'importance de la régulation de USP14 par le protéasome. En effet, cette enzyme doit uniquement pouvoir réduire la taille des chaînes de polyubiquitine au niveau du protéasome afin de favoriser le relargage ou la dégradation de substrats par le protéasome. Il est probable que UCH37 exerce une fonction similaire à celle de USP14 et il a été montré que l'activité enzymatique de UCH37 était aussi régulée par le protéasome (Qiu et al, 2006). Des études additionnelles seront toutefois nécessaires afin de déterminer quels types de chaîne de polyubiquitine sont clivés de façon préférentielle par USP14 ou UCH37.

## **2. Activation allostérique de USP7 par ses domaines Ubl**

### USP7 régule p53

USP7, également appelée HAUSP (herpes virus-associated ubiquitin specific protease), participe entre autres à la régulation du niveau de la protéine p53, un des plus importants suppresseurs de tumeur de la cellule. En réponse à une perte de l'intégrité

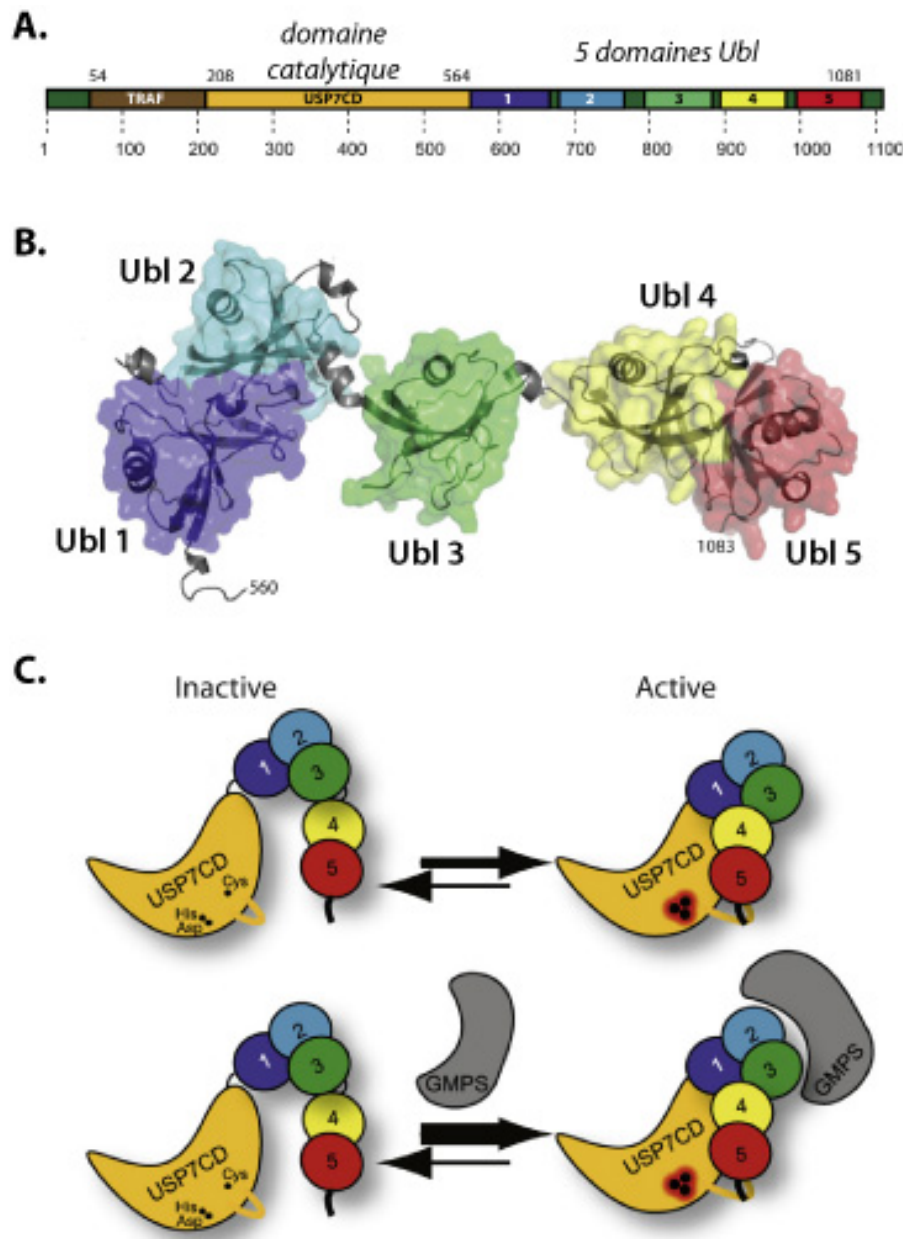


du génome, p53 peut en effet activer des programmes transcriptionnels qui vont mener au blocage du cycle cellulaire et si nécessaire, induire l'apoptose de ces cellules (Ventii & Wilkinson, 2008). La quantité de p53 doit donc être finement contrôlée et cette régulation est exercée par différents acteurs. Tout d'abord, l'E3 ligase Mdm2 (murine double minute 2) peut réduire la quantité de p53 en polyubiquitinant cette protéine en vue de sa dégradation par le protéasome. Au contraire, la déubiquitinase USP7 peut interagir avec p53 pour la déubiquitiner et ainsi empêcher sa dégradation par le protéasome. Cependant, USP7 peut aussi déubiquitiner et stabiliser l'E3 ligase Mdm2 ce qui suggère que USP7 se trouve au cœur d'un système de régulation complexe en agissant la fois sur la stabilisation de p53 mais aussi sur sa dégradation en facilitant son ubiquitination par Mdm2 (Ventii & Wilkinson, 2008). Il apparaît donc crucial que l'activité enzymatique de USP7 soit finement contrôlée pour permettre le maintien de quantités cellulaires adéquates de p53 et de Mdm2. On peut d'ailleurs noter que plusieurs études récentes ont démontré que l'activité catalytique de USP7 est régulée de façon allostérique (Faesen et al, 2011; Fernandez-Montalvan et al, 2007; Hu et al, 2002; Ma et al, 2010).

#### Activation de USP7 par ses domaines Ubl

La résolution de la structure tridimensionnelle du domaine catalytique de USP7 seul a démontré que ce domaine catalytique se trouve dans une conformation inactive car la distance entre la cystéine et l'histidine catalytique est trop grande pour permettre à l'enzyme de cliver une liaison isopeptidique. Par contre, la co-cristallisation du domaine catalytique de USP7 avec le substrat suicide ubiquitine aldéhyde a permis de stabiliser l'enzyme dans sa forme active dans laquelle la triade catalytique est positionnée de façon optimale (Hu et al, 2002).

Plusieurs études structurales ou fonctionnelles ont récemment démontré que l'équilibre entre les formes actives et inactives de l'enzyme était notamment régulé par les domaines additionnels de USP7 (Faesen et al, 2011; Fernandez-Montalvan et al, 2007; Ma et al, 2010). En effet, en plus de son domaine catalytique, USP7 possède au niveau de son extrémité N-terminale un domaine TRAF (tumor necrosis factor-receptor associated factor)/Math (meprin and TRAF homology) ainsi que 5 domaines Ubl dans sa région C-terminale (Figure 25 A.) (Faesen et al, 2011)). Bien que l'on retrouve des domaines Ubl dans plus d'une quinzaine d'autres USP, USP7 est la seule enzyme à posséder un nombre aussi important de domaines Ubl consécutifs.



**Figure 25. Mécanisme d'activation de USP7 par ses domaines Ubl.**

**A.** USP7 contient un domaine TRAF/Math (brun), un domaine catalytique (orange) et 5 domaines Ubl (arc-en-ciel). **B.** Représentation de la structure des 5 domaines Ubl de USP7. Les domaines 1 et 2 ainsi que les domaines 4 et 5 forment deux paires alors que le domaine 3, au centre, est plus libre (référence PDB : 2YLM). **C.** Schéma représentant le mécanisme d'activation du domaine catalytique de USP7 par ses domaines Ubl. Dans l'état inactif, les résidus du site catalytique ne sont pas correctement positionnés ce qui empêche l'enzyme de déubiquitiner un substrat. Cependant, l'interaction des domaines Ubl 4 et 5 avec le domaine catalytique de USP7 provoque un changement de conformation de la « boucle de permutation » ce qui permet de rendre l'enzyme active en positionnant les résidus catalytiques de façon optimale (schéma du dessus). La liaison de GMPS sur les domaines Ubl 1, 2 et 3 va ensuite pouvoir stabiliser la forme active de USP7 en augmentant l'affinité des domaines Ubl 4 et 5 pour le domaine catalytique de USP7 (schéma du dessous). La figure est adaptée de : (Faesen et al, 2011).

Les résultats de plusieurs études montrent que le domaine N-terminal TRAF de USP7 interagit directement avec des substrats de l'enzyme comme p53 ou Mdm2 (Hu et al, 2006; Hu et al, 2002; Saridakis et al, 2005; Sarkari et al, 2010; Sheng et al, 2006).

Cependant, il a été montré que le domaine TRAF ne régule pas de façon directe l'activité enzymatique de USP7 car sa délétion n'altère pas la capacité de l'enzyme à cliver un substrat synthétique (Faesen et al, 2011; Fernandez-Montalvan et al, 2007; Ma et al, 2010). Par contre, la délétion de la région C-terminale contenant les domaines Ubl provoque une diminution très importante de l'activité catalytique de USP7 sur des substrats synthétiques (comme par exemple l'Ub-AMC (Ubiquitine-7-amino-4-méthylcoumarin)) (Faesen et al, 2011; Fernandez-Montalvan et al, 2007; Ma et al, 2010). Récemment, la structure tridimensionnelle des 5 domaines Ubl de USP7 a été résolue et celle-ci montre que les domaines Ubl 1 et 2 ainsi que les domaines Ubl 4 et 5 sont appariés deux à deux, alors que le domaine Ubl 3 au centre est plus libre (Figure 25 B.). Des mesures biophysiques suggèrent d'ailleurs que la chaîne de domaines Ubl pourrait se replier sur elle-même en formant un coude au niveau du domaine Ubl 3 ce qui permettrait aux domaines Ubl 4 et 5 de se trouver à proximité du domaine catalytique de USP7 (Figure 25 C.) (Faesen et al, 2011). En effet, dans cette même étude, il a été montré que les domaines Ubl 4 et 5 pouvaient interagir avec le domaine catalytique de USP7 et qu'ils étaient nécessaires et suffisants pour activer l'enzyme. Enfin, il apparaît qu'une forme tronquée de USP7 ne possédant pas les domaines Ubl 4 et 5 possède une affinité plus faible pour l'ubiquitine qu'une forme entière de la protéine (Faesen et al, 2011).

#### Mécanisme de la régulation allostérique

Nous venons de voir que le passage d'une forme inactive à une forme active de USP7 implique un repositionnement des résidus de la triade catalytique. La comparaison précise des structures correspondant aux deux formes de l'enzyme indique qu'en fait, l'ensemble du site actif de l'enzyme est remodelé et qu'en particulier, la conformation d'une petite boucle se situant à proximité des résidus catalytiques est modifiée (Faesen et al, 2011; Hu et al, 2002). En mutant des résidus de cette boucle dite « boucle de permutation », Suxma et coll. ont montré que l'activation *in vitro* de USP7 par les domaines UBL 4 et 5 était perdue. De plus, il apparaît que la mutation combinée de deux résidus de cette « boucle de permutation » provoque une perte très importante de l'affinité du domaine catalytique de USP7 pour l'ubiquitine. De façon remarquable, une délétion de l'extrémité flexible C-terminale du domaine Ubl 5 provoque également une perte de l'activité enzymatique de USP7. Enfin, une série de mutations ponctuelles au sein de cette extrémité C-terminale du domaine Ubl 5 a mis en évidence le rôle crucial

joué par une isoleucine dans l'activation enzymatique de USP7. Les auteurs proposent donc un modèle dans lequel l'extrémité C-terminale du domaine Ubl 5 pourrait interagir avec la « boucle de permutation » afin de stabiliser la forme active du domaine catalytique de USP7 (Figure 25 C.) (Faesen et al, 2011).

D'autres études ont montré que la GMP-synthétase (GMPS), une enzyme du métabolisme des nucléotides, interagit avec USP7 chez la drosophile (van der Knaap et al, 2005) et dans des cellules humaines (Sarkari et al, 2009). De plus, cette interaction stimule l'activité catalytique de USP7 sur certains de ses substrats comme p53 ou H2B chez la drosophile (van der Knaap et al, 2005). L'étude réalisée par Suxma et coll. montre que l'interaction entre GMPS et USP7 a lieu au niveau des trois premiers domaines Ubl de la déubiquitinase et que cette interaction favorise la liaison des domaines Ubl 4 et 5 sur le domaine catalytique de USP7. Le modèle final propose donc que la liaison de GMPS sur les trois premiers domaines Ubl de USP7 déplace l'équilibre de la forme inactive vers la forme active de l'enzyme en stabilisant l'interaction entre l'extrémité du domaine Ubl 5 et la « boucle de permutation » du domaine catalytique (Figure 25 C.) (Faesen et al, 2011).

### **3. Inhibition de USP4 par un domaine Ubl**

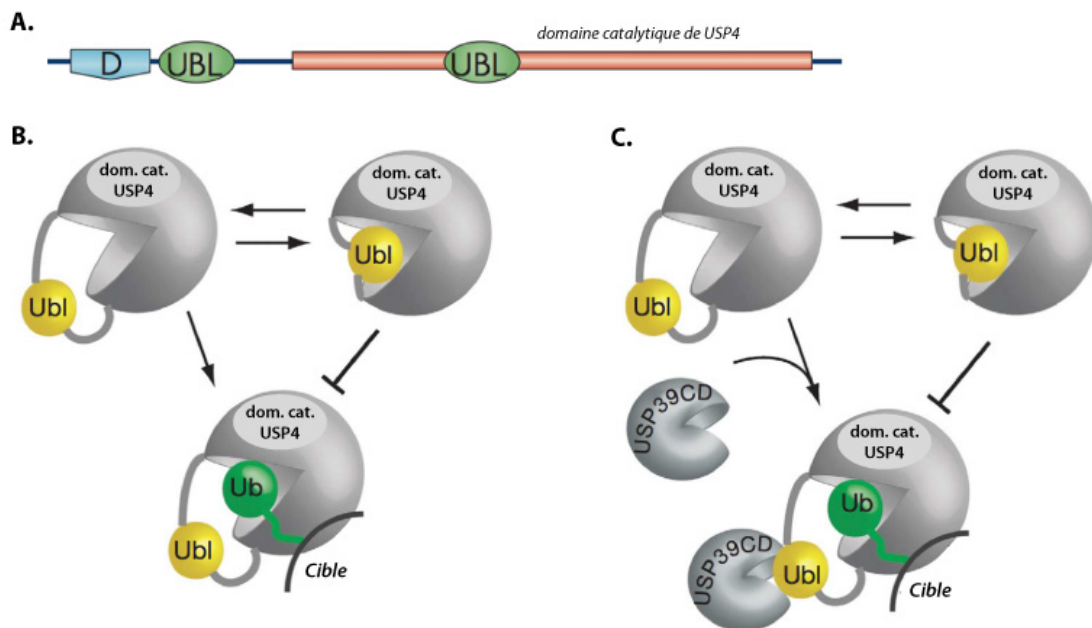
USP4 peut être recrutée au niveau du spliceosome par la protéine Sart3 pour déubiquitiner de façon préférentielle des chaînes de polyubiquitines liées par la K63 sur la sous-unité Prp3 de la particule U4 (Song et al, 2010).

USP4 possède deux domaines Ubl ; le premier se situe dans la région N-terminale de l'enzyme alors que le second est inséré dans le domaine catalytique (Figure 26 A.). Récemment, une étude présentant la structure cristallographique du domaine catalytique de USP4. Cette structure révèle une forme particulièrement inactive de USP4 car non seulement trois boucles viennent obstruer le site de liaison à l'ubiquitine mais en plus, l'extrémité du domaine Doigts est repliée sur elle même (Luna-Vargas et al, 2011). D'autre part, une analyse des propriétés fonctionnelles de USP4 indique que la délétion du domaine Ubl interne augmente la vitesse de clivage d'un substrat synthétique par l'enzyme *in vitro*. Les auteurs démontrent également que le domaine Ubl interne de USP4 peut interagir avec la poche de liaison de l'ubiquitine du domaine catalytique de l'enzyme et que la présence du domaine Ubl interne diminue l'affinité du domaine catalytique pour l'ubiquitine. Ces résultats suggèrent donc un mécanisme d'auto-régulation de USP4 par son domaine Ubl interne qui inhiberait l'activité de

l'enzyme en entrant en compétition avec le substrat pour lier le domaine catalytique (Figure 26 B.) (Luna-Vargas et al, 2011).

Il apparaît donc que l'activité enzymatique de USP4 est particulièrement contrôlée, non seulement par des éléments structuraux du domaine catalytique mais en plus par une inhibition exercée par le domaine Ubl interne. Un mécanisme pour lever cette seconde inhibition est proposé par les auteurs de l'étude qui suggèrent une implication de USP39, un composant du spliceosome qui possède un domaine catalytique USP mais dans lequel les résidus catalytiques sont absents (Song et al, 2010; Sowa et al, 2009). Il est suggéré que le domaine Ubl de USP4 pourrait interagir avec le domaine USP de USP39 ce qui diminuerait le temps de présence du domaine Ubl de USP4 dans le domaine catalytique de l'enzyme. En accord avec ce modèle, l'ajout de USP39 permet d'augmenter la vitesse de lyse d'un substrat synthétique par le domaine catalytique de USP4 contenant le domaine Ubl (Figure 26 C.) (Luna-Vargas et al, 2011).

Enfin, les auteurs spéculent sur la possibilité que des modifications post-



**Figure 26. Mécanisme d'inhibition de USP4 par son domaine Ubl interne.**

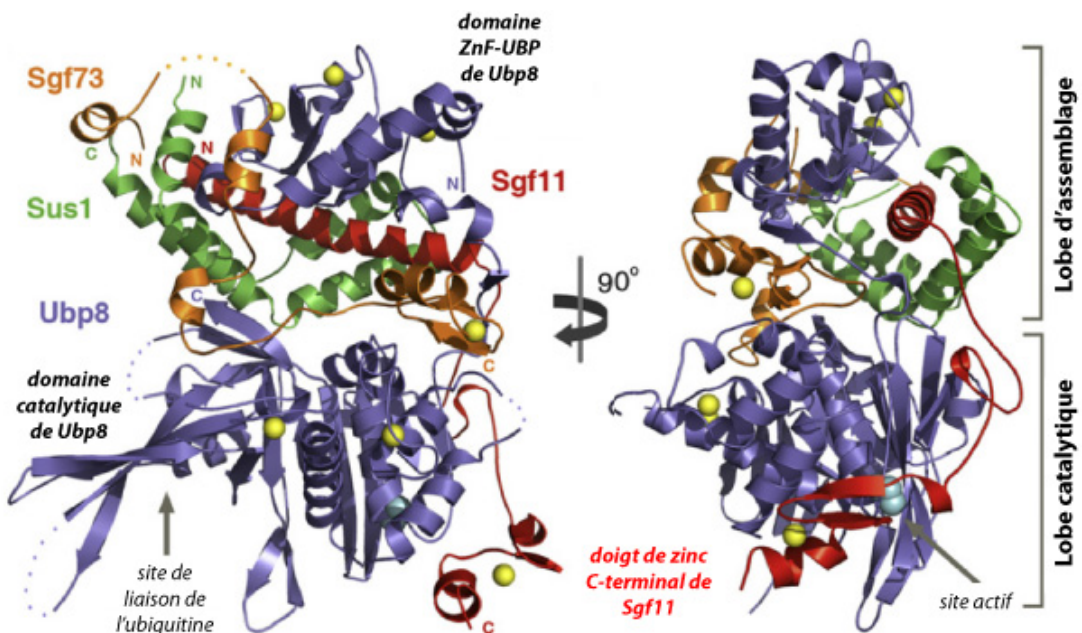
**A.** En N-terminal par rapport à son domaine catalytique (en rouge), USP4 possède un domaine DUSP (Domain in USP) (en bleu) et un domaine Ubl (en vert). De plus, USP4 présente un second domaine Ubl (en vert) qui est inséré au sein du domaine catalytique de l'enzyme. **B.** Modèle schématisé illustrant le mécanisme d'auto-inhibition exercé par le domaine Ubl interne (en jaune) sur le domaine catalytique (en gris) de USP4. La présence du domaine Ubl interne dans le domaine catalytique de USP4 empêche la venue d'un substrat. **C.** D'autres USP, comme USP39 pourrait participer à la levée de l'inhibition du domaine Ubl interne de USP4 en interagissant avec lui. La figure est adaptée de : (Komander et al, 2009; Luna-Vargas et al, 2011).

traductionnelles du domaine Ubl pourraient affecter la liaison de ce domaine au domaine catalytique de USP4 ce qui ajouterait un niveau additionnel de régulation de l'enzyme. De façon intéressante, USP4 n'est pas la seule DUB à posséder un domaine Ubl inséré au sein de son domaine catalytique. Parmi les 7 USP qui possèdent un tel domaine Ubl, USP11 et USP15 présentent une structure primaire similaire à celle de USP4 ce qui suggère que ces deux enzymes au moins pourraient être inhibées par leur domaine Ubl interne (Luna-Vargas et al, 2011).

#### IV. Activations allostériques réalisées par des ZnF-UBP

##### 1. Régulation de Ubp8 par le DUBm de SAGA

Nous avons vu que le complexe co-activateur de la transcription SAGA possède une activité de déubiquitination des histones H2B portée par Ubp8 chez la levure *S. cerevisiae*. De plus, Ubp8 exerce son activité au sein d'un module de SAGA appelé module de déubiquitination (DUBm) qui contient trois autres sous-unités ; Sgf73, Sgf11 et Sus1 (cf. Partie sur le complexe SAGA) et (Daniel et al, 2004; Henry et al, 2003; Ingvarsdottir et al, 2005; Kohler et al, 2006; Köhler et al, 2008; Lee et al, 2005; Powell et al, 2004). En



**Figure 27. Structure du module de déubiquitination de SAGA chez *S. cerevisiae*.**

Deux vues orthogonales du DUBm contenant Ubp8 (en violet), Sgf11 (en rouge), Sus1 (en vert) et le domaine N-terminal de Sgf73 (en orange) sont présentées. Le lobe d'assemblage contient le ZnF-UBP de Ubp8, Sus1 et l'hélice N-terminale de Sgf11 alors que le lobe catalytique contient le domaine catalytique de Ubp8 et le doigt de zinc C-terminal de Sgf11. Le domaine N-terminal de Sgf73 permet de maintenir les deux lobes à proximité l'un de l'autre. Les régions non structurées apparaissent en pointillés. Les atomes de zinc sont représentés par des sphères jaunes et la chaîne latérale de la cystéine catalytique est colorée en bleu ciel (référence PDB : 3M99). La figure est adaptée de : (Kohler et al, 2010).

reconstituant un DUBm recombinant, Kohler et coll. ont également démontré que la présence des quatre sous-unités du module était nécessaire pour que Ubp8 clive efficacement un substrat synthétique *in vitro* (Köhler et al, 2008). Récemment, deux équipes ont déterminé la structure tridimensionnelle du DUBm de la levure *S. cerevisiae* comprenant le domaine N-terminal de Sgf73 ainsi que les protéines entières Ubp8, Sgf11 et Sus1. Ces informations structurales ont permis de déchiffrer le mécanisme d'activation de Ubp8 par ses partenaires du DUBm (Kohler et al, 2010; Samara et al, 2010).

#### Formation d'un module structural

La structure du DUBm révèle que celui-ci est organisé en deux sous-modules, appelés lobe d'assemblage et lobe catalytique (Figure 27). Le premier s'organise autour du ZnF-UBP de Ubp8 alors que le second est formé autour du domaine catalytique de l'enzyme. De plus, il apparaît que le DUBm est stabilisé par un réseau très complexe d'interactions dans lequel chaque sous-unité du module réalise des contacts avec ses trois autres partenaires. Notons également que 8 atomes de zinc sont retrouvés au sein du DUBm (Kohler et al, 2010; Samara et al, 2010).

En plus du domaine ZnF-UBP de Ubp8, le lobe d'assemblage contient la région N-terminale de Sgf11 et Sus1. La structure du dimère formé par cette région N-terminale de Sgf11 et Sus1 était déjà connue et montre que Sus1 s'agrippe autour d'une longue hélice formée par Sgf11. D'ailleurs, deux molécules de Sus1 interagissent de la même manière autour d'une longue hélice de Sac3 au sein du complexe TREX2, un complexe d'export des ARNm dont Sus1 est une sous-unité (Ellisdon et al, 2010; Jani et al, 2009). Dans la structure du DUBm, on retrouve la longue hélice de Sgf11 sur laquelle non seulement Sus1 mais aussi le ZnF-UBP de Ubp8 viennent se fixer (Figure 27) (Kohler et al, 2010; Samara et al, 2010). En effet, contrairement à ce qui est observé pour le ZnF-UBP de USP5, le ZnF-UBP de Ubp8 ne peut pas interagir avec l'extrémité C-terminale libre de l'ubiquitine, mais par contre, il réalise des interactions qui sont à la base de la formation structural du DUBm de SAGA (cf. B. IV. 2. et Partie Résultats).

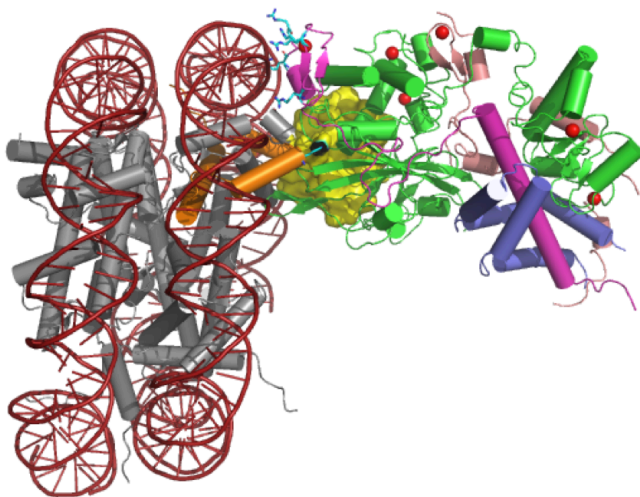
Il est en effet crucial que la région N-terminale de Sgf11 soit figée entre le ZnF-UBP de Ubp8 et Sus1, car cela permet de positionner de façon optimale le doigt de zinc de l'extrémité C-terminale de Sgf11 par rapport au domaine catalytique de Ubp8. En effet, cette extrémité C-terminale de Sgf11 forme avec le domaine enzymatique de Ubp8 le lobe catalytique du DUBm au sein duquel, le doigt de zinc de Sgf11 joue un rôle crucial

pour activer l'enzyme, comme nous allons le voir (Kohler et al, 2010; Samara et al, 2010).

Enfin, le domaine N-terminal de Sgf73, qui est formé par une longue région dépliée et qui se termine par un petit doigt de zinc va réaliser des interactions avec les deux lobes structuraux afin de les maintenir de façon stable à proximité l'un de l'autre dans une conformation donnée (Kohler et al, 2010; Samara et al, 2010).

#### Activation allostérique et reconnaissance spécifique du substrat

La comparaison de la structure du DUBm obtenue avec ou sans ubiquitine aldéhyde ne révèle pas de changements conformationnels importants liés à la présence d'ubiquitine aldéhyde. Il est cependant très vraisemblable que les interactions réalisées par Ubp8 au sein du DUBm maintiennent l'enzyme dans une conformation active ce qui expliquerait pourquoi l'enzyme isolée est beaucoup moins active que l'enzyme au sein du module (Samara et al, 2010).



**Figure 28. Modèle de la liaison du DUBm avec un nucléosome ubiquitiné sur H2B.**

Le DUBm (référence PDB : 3MHS) est composé par Ubp8 (en vert), Sgf11 (en rose foncé), Sus1 (en violet), et par le domaine N-terminal de Sgf73 (en rose pâle). Les atomes de zinc stabilisant la structure sont représentés en rouge. L'ubiquitine (en jaune) est fixée sur une molécule de H2B (en orange) et interagit avec le domaine catalytique de Ubp8. Les autres histones sont colorées en gris et l'ADN nucléosomal est représenté en rouge. Les arginines du doigt de zinc de Sgf11 (chaîne latérale en bleu ciel) qui sont nécessaires pour que le DUBm déubiquitine un nucléosome, sont en contact avec les groupements phosphates de l'ADN du nucléosome. Ces interactions doivent permettre au DUBm de reconnaître spécifiquement son substrat. La figure est adaptée de : (Samara et al, 2010).

Au sein du lobe catalytique, le petit doigt de zinc de Sgf11 se retrouve à proximité du site catalytique de Ubp8. De façon très intéressante, la délétion de ce seul doigt de zinc n'altère pas la stabilité du DUBm mais provoque une diminution importante de l'activité enzymatique de Ubp8. De plus, des expériences de mutagenèse dirigées ont révélé que des résidus hydrophobes du doigt de zinc de Sgf11 interagissant avec le domaine catalytique de Ubp8 sont nécessaires pour permettre à l'enzyme de cliver efficacement un substrat synthétique *in vitro*. Ceci démontre donc que

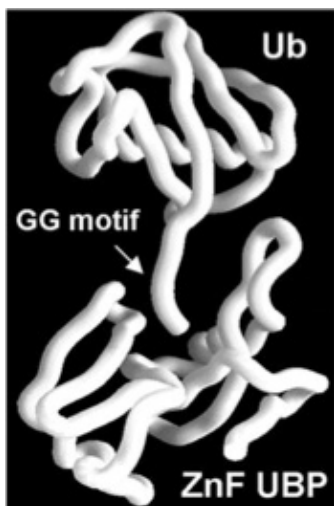


l'interaction du doigt de zinc de Sgf11 avec le domaine catalytique de Ubp8 permet d'activer ce dernier (Kohler et al, 2010).

D'autre part, la mutation d'arginines du doigt de zinc de Sgf11, qui sont exposées vers le solvant et qui sont strictement conservées au cours de l'évolution n'affecte pas la capacité de Ubp8 à déubiquitiner un substrat synthétique *in vitro*. Par contre, la mutation de ces résidus empêche le DUBm de déubiquitiner H2B dans le contexte d'un nucléosome, tant *in vitro* que *in vivo* (Kohler et al, 2010). Ceci suggère donc que ces arginines de Sgf11 pourraient interagir avec les groupements phosphates de l'ADN du nucléosome, et que ces interactions seraient nécessaires pour permettre au DUBm de SAGA de reconnaître spécifiquement et de déubiquitiner H2Bub dans le contexte d'un nucléosome (Figure 28) (Samara et al, 2010).

## 2. USP5 reconnaît spécifiquement des chaînes de polyubiquitine libres

USP5 (également appelée Isopeptidase T) et son homologue chez la levure Ubp14, sont responsables du désassemblage des chaînes de polyubiquitine qui ne sont pas conjuguées à une protéine cible (Amerik et al, 1997; Reyes-Turcu et al, 2008). En effet, la délétion de Ubp14 chez la levure ainsi que l'élimination de USP5 par un traitement avec un ARN interférent dans des cellules mammifères causent une accumulation de ces

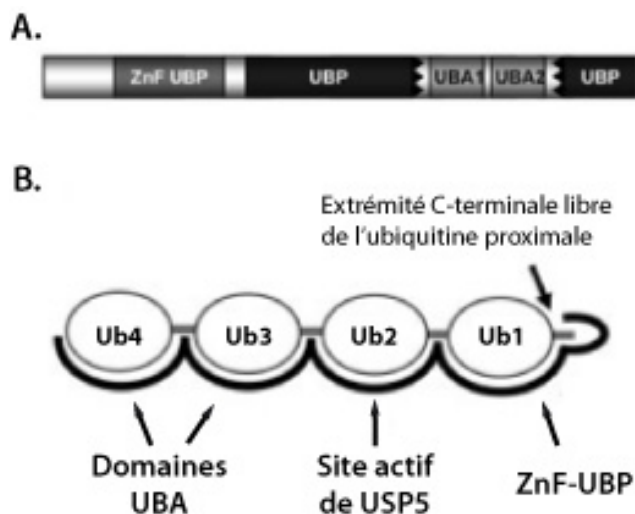


**Figure 29. Interaction du domaine ZnF-UBP de USP5 avec l'ubiquitine.** Les deux glycines (motif GG) de l'extrémité C-terminale de l'ubiquitine libre lient la poche du domaine ZnF-UBP de USP5 (référence PDB : 2G45). La figure est adaptée de : (Reyes-Turcu et al, 2006).

chaînes d'ubiquitines libres. Celles-ci sont notamment générées au moment où une chaîne de polyubiquitine est enlevée de sa protéine cible lors la dégradation de cette dernière par le protéasome. D'ailleurs, l'accumulation de ces chaînes libres provoque une baisse de l'activité du protéasome, sans doute à cause d'une compétition entre les chaînes d'ubiquitines libres et les chaînes liées à un substrat pour se fixer sur les récepteurs à ubiquitine du protéasome. Il est donc très important que la dégradation de ces chaînes libres soit réalisée de façon efficace pour ne pas inhiber le protéasome mais aussi pour permettre le maintien d'un niveau cellulaire adéquat d'ubiquitine libre (Amerik et al, 1997; Reyes-Turcu et al, 2008).

En plus d'un domaine catalytique, USP5 possède trois domaines de liaison à l'ubiquitine qui lui permettent de dégrader des chaînes de polyubiquitine libres de façon spécifique et efficace (Figure 30). Tout d'abord, le domaine N-terminal de la protéine, un domaine ZnF-UBP, peut lier le motif Gly-Gly de l'extrémité C-terminale d'une molécule d'ubiquitine libre (Figure 29). De façon intéressante, cette interaction est nécessaire pour stimuler l'activité enzymatique de USP5 ce qui explique la spécificité de USP5 pour des chaînes de polyubiquitine libres (Reyes-Turcu et al, 2006) et (cf. Partie Résultats). D'autre part, USP5 possède deux domaines UBA qui peuvent aussi lier l'ubiquitine. Une étude a montré que ces domaines participaient à la liaison de substrats de USP5 et que l'affinité de l'enzyme augmentait pour de longues chaînes de polyubiquitine (Reyes-Turcu et al, 2008).

Ces résultats suggèrent donc que l'extrémité proximale de la chaîne d'ubiquitine libre soit reconnue par le ZnF-UBP de USP5 ce qui activerait l'enzyme, que la seconde molécule d'ubiquitine se trouve dans le site actif de USP5 et enfin que les troisièmes et



**Figure 30. Domaines de liaison à l'ubiquitine de USP5.**

**A.** USP5 possède 4 domaines de liaison à l'ubiquitine : un domaine ZnF-UBP, un domaine catalytique USP, et deux domaines UBA. **B.** Liaison putative d'une chaîne de 4 ubiquitines sur USP5. L'extrémité libre de l'ubiquitine proximale occupe la poche du domaine ZnF-UBP, ce qui active l'enzyme, et les autres ubiquitines occupent les deux domaines UBA ainsi que le domaine catalytique de USP5. L'ensemble de ces sites de liaison de l'ubiquitine doit permettre à USP5 de cliver spécifiquement et efficacement des chaînes libres de polyubiquitine. La figure est adaptée de : (Reyes-Turcu et al, 2008).

quatrième molécules d'ubiquitines interagissent avec les domaines UBA, pour stabiliser la liaison du substrat sur USP5 (Figure 30). La présence de 4 domaines de liaison de l'ubiquitine au sein de USP5 suggère donc une optimisation de l'enzyme pour la dégradation de longues chaînes de polyubiquitine libre afin d'éviter toute interférence de celles-ci avec le protéasome (Reyes-Turcu et al, 2006; Reyes-Turcu et al, 2008).

## **Partie 4 : l'ubiquitination de H2B**

Cette partie de l'introduction de ma thèse sur l'ubiquitination de H2B se présente sous la forme d'un manuscrit de revue que nous souhaitons à terme soumettre à publication. Cette version du manuscrit est encore une version préliminaire sur laquelle nous allons continuer à travailler mais le message que nous souhaitons faire passer apparaît déjà très clairement.

Notre but a été de reprendre un nombre important de publications sur le sujet de l'ubiquitination de H2B afin d'essayer de faire la part des choses entre des expériences indubitables et d'excellentes qualités et d'autres résultats plus douteux. Notre objectif n'est pas de réconcilier l'ensemble des résultats qui ont été publiés sur le sujet mais davantage de dégager un modèle cohérent faisant ressortir ce que l'on sait de la fonction de l'ubiquitination de H2B à l'heure actuelle.

Dans ce manuscrit, nous commençons par présenter les enzymes qui catalysent l'ubiquitination de l'histone H2B puis celles qui éliminent cette marque en déubiquitinant H2B. Nous discutons ensuite les expériences de ChIP qui ont permis, notamment chez les eucaryotes supérieurs, de démontrer que H2B est ubiquitiné sur la région transcrite des gènes exprimés. Enfin, nous essayons de faire le point sur les fonctions qui sont associées à l'ubiquitination de H2B en lien avec l'initiation ou l'élongation de la transcription.

## Histone H2B ubiquitination: signaling, but to whom and why?

Jacques Bonnet, Didier Devys and László Tora\*

Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie  
Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 964, Université de  
Strasbourg, BP 10142 - 67404 ILLKIRCH Cedex, France

\*To whom correspondence should be addressed:

IGBMC, 1 rue Laurent Fries, BP10142

67404 Illkirch Cedex, France

Tel: 33 388 65 34 44

Fax 33 388 65 32 01

Email: [laszlo@igbmc.fr](mailto:laszlo@igbmc.fr)

### **Abstract**

Recently, the genome wide distribution of histone H2B ubiquitination in different organisms has been reported. In this review we discuss the described functions of H2Bubiquitination as well as on the enzymes responsible for setting and erasing this mark at the light of these recent studies. In particular we are describing novel concepts on the role of histone H2B ubiquitination and discussing the molecular mechanisms in which it participates.

## Introduction

Post-translational modifications of histones regulate chromatin states and thus they influence all chromatin dependent processes. Histone ubiquitination can regulate other modifications (e.g. methylation and phosphorylation and vice versa) in order to define precise and functional chromatin environments and also for setting the stage for other gene regulatory processes.

Ubiquitin, a 76-residue protein, is covalently linked to a target protein in a multistep process leading to the formation of an isopeptide bond between the C-terminal glycine residue of ubiquitin and the  $\epsilon$ -amino group on a lysine residue of the target protein. After binding to an ubiquitin-activating enzyme (E1), ubiquitin is transferred to one of a few dozen ubiquitin-conjugating enzymes (E2) and finally conjugated to its substrate by one of several hundred ubiquitin ligases (E3), which confer substrate specificity<sup>1</sup>. Ubiquitin has seven lysine residues and all of them can be used for the formation of polyubiquitin chains. The best understood function of ubiquitination is to target protein for degradation through the recognition by the proteasome of lysine-48-polyubiquitin chains. However, more diverse functions were reported for other modifications such as polyubiquitination with different linkages or also monoubiquitination, the conjugation of a single ubiquitin to a lysine residue. Like other proteins modifications, ubiquitination is highly controlled and can be reversed by the action of deubiquitinating enzymes or deubiquitinases<sup>2</sup>. There are about 100 human deubiquitinases and it has been proposed that substrate specificity relies on their association with protein partners<sup>3,4</sup>.

Histone H2A was the first protein shown to be ubiquitinated and later on ubiquitination of histone H2B was also shown to occur *in vivo*<sup>5,6</sup>. Histone H2A is monoubiquitinated (H2Aub) on lysine 119 exclusively in higher eukaryotes whereas histone H2B is monoubiquitinated (H2Bub) on lysine 123 in yeast or the corresponding lysine 120 in mammals. H2Aub and H2Bub are the most abundant ubiquitin conjugates in eukaryotes corresponding to ~10% of H2A and ~1% of H2B in mammalian cells whereas ~10% of H2B is ubiquitinated in yeast, which contains no detectable levels of H2Aub. Although polyubiquitination of histone H2A and H2B has also been reported, monoubiquitination is by far the most abundant form of ubiquitinated H2A or H2B. Besides H2A and H2B, histones H1,

H2AZ, H3 and H4 were also reported as substrates for ubiquitin modification but at lower levels than H2A and H2B. Since a number of recent studies appeared reporting the genome wide distribution of H2B monoubiquitination, in this review we will focus on this modification and on emerging mechanistic models involving the function of H2Bub in gene regulation and downstream events.

## I. Identification of H2Bub enzymes shed new light on the function of this modification.

### I.1 Histone H2B E2-E3 ligases

The first steps towards understanding the role of H2Bub were the identification of H2B ubiquitinating enzymes. In *S. cerevisiae* Rad6 was identified as the only H2B ubiquitin-conjugating enzyme (E2) and Bre1, a RING finger domain-containing protein, as the only H2B ubiquitin ligase (E3)<sup>7-9</sup>. Indeed deletion of either Rad6 or Bre1 led to a complete disappearance of H2Bub comparable to what is observed in a yeast strain carrying a non-ubiquitinable form of histone H2B. In addition, Lge1, a protein interacting with Bre1 was also shown to be required for H2B ubiquitination<sup>7,10</sup>.

In *S. pombe* a single Rad6 homolog, Rhp6, was found associated with two Bre1 homologs, Brl1 and Brl2, in a multisubunit complex that also contains the serine rich protein Shf1<sup>11,12</sup>. The two highly conserved human Rad6 homologs, RAD6A and RAD6B, can rescue Rad6 function in a *rad6Δ* yeast strain<sup>13</sup>. However, RAD6A appears to comprise the majority of the endogenous RAD6 activity in human cells and the function of these two E2 conjugating enzymes are not fully redundant<sup>13,14</sup>. Similarly to *S. pombe*, mammals have two Bre1 homologs, RNF20 and RNF40 that heterodimerize to form a protein complex that interacts with RAD6<sup>13,15,16</sup>. The two RING finger domain-containing E3, RNF20 and 40 similarly contribute to H2B ubiquitination<sup>13</sup>. The WW domain-containing Adaptor with Coiled-coil protein (WAC) was also identified as a subunit of the RNF20/RNF40 complex and shown to be required for H2B ubiquitination<sup>17</sup>. Although, histone H2B E2 and E3 proteins are highly conserved through evolution, the other associated proteins appear to be specific in different species as no sequence similarity could be found between Lge1, Shf1 and WAC. However, these proteins play a crucial role in the regulation of H2B E2 and E3 enzymes as H2B ubiquitination was suppressed in *LGE1* or *SHF1* mutants which display a phenotype that is comparable to that of *RAD6* or *BRE1* mutants and WAC depletion strongly reduced H2Bub levels. Thus, the role of these different Bre1- and RNF20/40-associated proteins in the regulation of H2B ubiquitination needs to be further investigated in the different model systems.

## **I.2 H2Bub deubiquitinases**

### **a) Deubiquitination activity of SAGA**

The dynamic nature of H2Bub was revealed by the identification of several specific deubiquitinases in different organisms. Ubp8, a subunit of the SAGA co-activator complex, was the first identified H2B deubiquitinase in *S. cerevisiae*<sup>18,19</sup>. The deubiquitination activity of the SAGA complex is highly conserved as revealed by the identification of Ubp8 orthologs in *S. pombe* (Ubp8), *D. melanogaster* (Nonstop) and human (USP22)<sup>20-23</sup>. Ubp8 was later shown to interact with three other SAGA subunits (Sgf73, Sgf11 and Sus1) to form a functional deubiquitination module of the SAGA complex. In addition, catalytic activities of either  $\gamma$ Ubp8 or hUSP22 were shown to be allosterically regulated by their interactions with the other subunits of this module<sup>24-28</sup>.

### **b) Other H2B deubiquitinases**

In *S. cerevisiae* another ubiquitin protease Ubp10, was also shown to deubiquitinate H2Bub. The steady state levels of H2Bub were increased in Ubp8 $\Delta$  or in Ubp10 $\Delta$  cells, but were further increased in a double Ubp8 Ubp10 mutant suggesting that these two deubiquitinases target different loci in the yeast genome<sup>29,30</sup>. Note however that sequence comparison analysis failed to identify a clear Ubp10 ortholog in *S. pombe* or in metazoans. In *D. melanogaster*, another protease, USP7, deubiquitinates H2Bub and its activity is stimulated through an association with the biosynthetic enzyme GMP synthetase (GMPS) *in vitro*<sup>31</sup>. It was then suggested that downregulation of USP7 induced an increase of H2Bub in *Drosophila* embryos and in human cells<sup>32,33</sup>. In addition, *Drosophila* Scrawny was also identified as an H2B deubiquitinase that plays a role in gene silencing<sup>34</sup>. Moreover, a study in *A. thaliana* suggested that the ubiquitin protease Ubp26 deubiquitinates H2Bub and mediates heterochromatin silencing<sup>35</sup>. In contrast with the SAGA deubiquitination activity that is highly conserved through evolution, Ubp10 does not have any definite ortholog in other species whereas *Drosophila* Scrawny or *Arabidopsis* Ubp26 have orthologs in mammals that were not reported as H2B deubiquitinases so far.

## **II Genome wide distribution of H2Bub**

Initial studies using ChIP experiments in *S. cerevisiae* analyzed H2B ubiquitination at the *GAL1* promoter only and suggested that H2B is transiently ubiquitinated during gene activation and is then deubiquitinated by Ubp8<sup>19,36</sup>. However, it was then proposed that



H2B is not only ubiquitinated at the promoter but also in the transcribed region of active genes. Indeed, H2Bub was detected by CHIP in the coding region of transcribed genes and Rad6, Bre1 or Paf1 complex mutants display transcription elongation defects as assessed by 6-azauracil sensitivity and synthetic lethality with transcription elongation factors<sup>36,37</sup>.

All early results explaining the role of H2Bub in transcription were mainly based on the analysis of individual model genes in yeast, but more recent genome wide studies in higher eukaryotes provide a different view of the role of H2Bub. The development of a ChIP grade antibody that is highly specific to human H2Bub<sup>38</sup> allowed to precisely map this modification genome wide at different gene regions in human, *Drosophila* and *Arabidopsis thaliana* cells<sup>39-41</sup>. These analyses demonstrate that H2Bub is exclusively found associated with active genes and is not detected at silent genes, intergenic regions or other heterochromatin territories. In most active genes, the highest levels of H2Bub were observed just downstream of the PolIII transcription start sites (TSSs) and H3K4me3 peaks and cover the whole transcribed regions. Interestingly, H2Bub better correlate with H3K36me3 and with H3K79me1 and H3K79me2 than with H3K4me3 in good agreement with its role in the regulation of Dot1 activity (see below). These different studies concordantly demonstrate that in higher eukaryotes, H2B is exclusively ubiquitinated in transcribed genes and predominantly downstream of the TSSs strongly supporting a role for H2Bub in transcription elongation.

In contrast with the genome-wide distribution of H2Bub that is found on most active genes, suppression of H2B ubiquitination appears to have a mild effect on gene expression. Indeed, mutation of the ubiquitination site of H2B in yeast results in an about twofold or higher change in expression of about 1.5% of the protein coding genes with 75% of these increasing<sup>42</sup>. A gene expression analysis of mutants of more than 150 chromatin machinery components showed that Rad6, Bre1 and Lge1 mutants share a similar pattern of gene expression changes and interestingly cluster together with components of the PAF, COMPASS and BUR complexes<sup>43</sup>. Similarly, only a small subset of genes was transcriptionally affected by RNF20 depletion in human cells<sup>44</sup>. Further analyses of genes downregulated by RNF20 suggested that RNF20 would interfere with the recruitment of the elongation factor TFIIS although further studies are needed to better understand this mechanism<sup>45</sup>.

### **III. Regulation and functions H2Bub**

### III.1 Mechanisms of trans-tail modifications of histone H3

A key step in the understanding of H2Bub function was the observation of its involvement in a trans-tail regulation of H3K4 and H3K79 methylation in *S. cerevisiae*. Indeed mutations of *RAD6* and *BRE1*, or loss of the H2B ubiquitination site suppressed H3K4 methylation by Set1 and H3K79 methylation by Dot1, two site specific methyltransferases<sup>46-49</sup>. Interestingly, H2B ubiquitination does not regulate Dot1 recruitment within gene bodies, but is required to generate the di- and tri-methylation states of H3K79 in vivo<sup>50</sup>. Further experiments indicate that Dot1 is a distributive enzyme that dissociates and reassociates to its substrate to perform the different rounds of methylation and that H2Bub stimulates its overall catalytic activity and thereby synthesis of all methylation states<sup>51</sup>. Experiments using nucleosomes containing H2Bub suggest that the ubiquitin moiety would be an allosteric activator of Dot1 stimulating its catalytic activity<sup>52</sup> (Figure 1).

Unlike Dot1, Set1 methylase is part of protein complex called COMPASS and appears to be a processive enzyme that once bound to its substrate can generate mono-, di- or trimethylation. Hence its regulation by H2Bub appears to be mediated by mechanisms that are highly divergent from Dot1 regulation involving other COMPASS subunits<sup>53,54</sup>. Two independent studies identified the COMPASS subunit Swd2 as a key factor for H3K4 trans-tail regulation although the underlying regulatory mechanisms described in the two studies are different. One report indicates that Swd2 is ubiquitinated by Rad6-Bre1, a modification required for the recruitment of the Spp1 subunit of COMPASS which promotes H3K4 trimethylation<sup>54</sup> (Figure 1). In contrast, the other study suggests that H2Bub directly or indirectly, recruits Swd2 to chromatin thereby facilitating its association with COMPASS<sup>53</sup>. In this model an inactive COMPASS complex that lacks Swd2 is recruited to the chromatin through an interaction with the Paf1 complex (see below) and Swd2 is independently recruited by H2Bub and then assemble with and activate COMPASS (Figure 1).

However, it should be noted that the role of H2Bub is not restricted to the cross-talk with H3 methylation. Indeed, in *S. pombe* the loss of H2B ubiquitination results in a phenotype and in transcriptional alterations that are not fully recapitulated in strain lacking H3K4 methylation<sup>11</sup>. A similar cross-talk was also reported in human cells in which depletion of RNF20 strongly reduced H2B ubiquitination and subsequent methylation of H3K4 and H3K79. Surprisingly, overexpression of RNF20 increased the levels of H2Bub and also of

H3K4 and H3K79 di- and tri-methylation<sup>15, 16</sup>. This latter result contrasts with the observation that elevated levels of H2Bub in an Ubp8Δ yeast strain do not influence the global levels of H3K4 methylation<sup>18, 55</sup>.

### **III.2 Rad6/Bre1 recruitment and role of H2Bub in transcription initiation**

Early studies analysed the role of H2Bub in the regulation of different model genes and focused on its establishment at the promoter region. Bre1 was shown to be required for Rad6 recruitment at the promoter of highly transcribed genes<sup>9</sup>. Surprisingly, some authors reported Rad6 recruitment exclusively at the promoter region of the highly expressed yeast *PMA1* gene<sup>9</sup> although others found equal levels of Rad6 associated with all analysed regions of 3 highly expressed genes including *PMA1*<sup>37</sup>. Furthermore, after gene induction, Rad6 was found to be recruited to all *GAL1* genic regions (promoter, 5' and 3' coding regions) in a Gal4 activator dependent manner<sup>36</sup> (Figure 1). Similarly, the human E3 protein RNF20 was shown to directly interact with p53 and to be recruited to promoters in a p53 dependant manner<sup>15</sup>.

### **III.3 Role of H2Bub in transcription elongation**

Further support of the trans-tail regulation was provided by a number of studies aimed at the identification of factors regulating histone H3K4 and H3K79 methylation. Indeed, two independent studies showed that some, but not all, subunits of the Paf1 complex are required for H3K4 and H3K79 methylation. Furthermore, the Paf1 complex which is known to associate with the elongating form of RNA polymerase II (Pol II) was shown to mediate the recruitment of COMPASS to the promoter and 5' region of transcribed genes<sup>56, 57</sup>. Further studies indicate that Paf1 is required for ubiquitination of H2B, strongly suggesting that Paf1 affects H3K4 and H3K79 methylation through its effect on H2B ubiquitination<sup>58, 59</sup>. Interestingly, deletion of Rtf1 and Paf1, two subunits of the Paf1 complex, suppressed H2B ubiquitination although the recruitment of Rad6-Bre1 was not affected suggesting that the Paf1 complex interacts with and activates Rad6-Bre1 (Figure 1).

Similarly to the Paf1 complex, BUR2, a cyclin subunit of the the Bur1/Bur2 cyclin-dependent protein kinase, was identified as a regulator of H3K4 methylation and further shown to be required for H2B ubiquitination<sup>60, 61</sup>. Furthermore, in a BUR2 deleted strain the recruitment of the PAF complex was shown to be impaired through an unknown mechanism.

Finally, one study reported that the Bur1/Bur2 complex can phosphorylate the conserved Ser120 of Rad6 and that the mutation of this Serine to Alanine results in a significant reduction of H2Bub<sup>61</sup>.

H2Bub was also shown to facilitate H2A/H2B dimer eviction and nucleosome reassembly in the wake of elongating RNA Pol II underscoring its role in transcription elongation<sup>62-64</sup>. Indeed, both the H2B ubiquitinating enzymes and FACT, a histone chaperone and elongation factor, were shown to interact with PAF and together these three complexes are required for transcription elongation. In addition, H2Bub was proposed to act as a barrier for the recruitment of Ctk1, the major kinase for the serine-2 of the Pol II CTD<sup>65</sup>. In this model, removal of H2Bub by SAGA following transcription initiation would be required for phosphorylation of the serine-2 of the Pol II CTD and further recruitment of the Set2 methyltransferase and consequent H3K36 methylation.

Such a profile of H2Bub distribution is supported by the need of the cyclin-dependent kinase 9 (CDK9) activity for H2B ubiquitination<sup>66</sup>. The Serine 2 (Ser2) of the PolIII C-terminal domain (CTD) is phosphorylated by CDK9, the kinase subunit of pTEF-B elongation factor, in the transcribed region and is required for post-initiation transcription events. Interestingly, inhibition of CDK9 activity suppresses phosphorylation of Ser2 and reduces the global levels of H2Bub. As it has been shown that WAC, a factor associated with RNF20/RNF40 and required for H2B ubiquitination interacts with the P-Ser2 of PolIII CTD, this suggest that RNF20/RNF40 would be recruited in the transcribed region by an interaction with the elongating PolIII<sup>17</sup>.

## **Conclusions**

Since the discovery of histone H2B monoubiquitination more than 30 years ago, a number of studies have identified different factors involved in the regulation of this modification. Progresses made in chromatin analysis and the development of specific tools led to a accurate picture of H2Bub distribution and provide a better understanding of its functions. However, the molecular mechanisms underlying the role of H2Bub in transcription initiation and elongation, in mRNA modification and export, in silencing or in DNA repair are still incompletely characterized. It is still not clear why this mark that is associated with the transcribed region of most active genes appears to regulate the transcription of only a subset of these genes. It remains also to determine whether the crosstalk between H2Bub

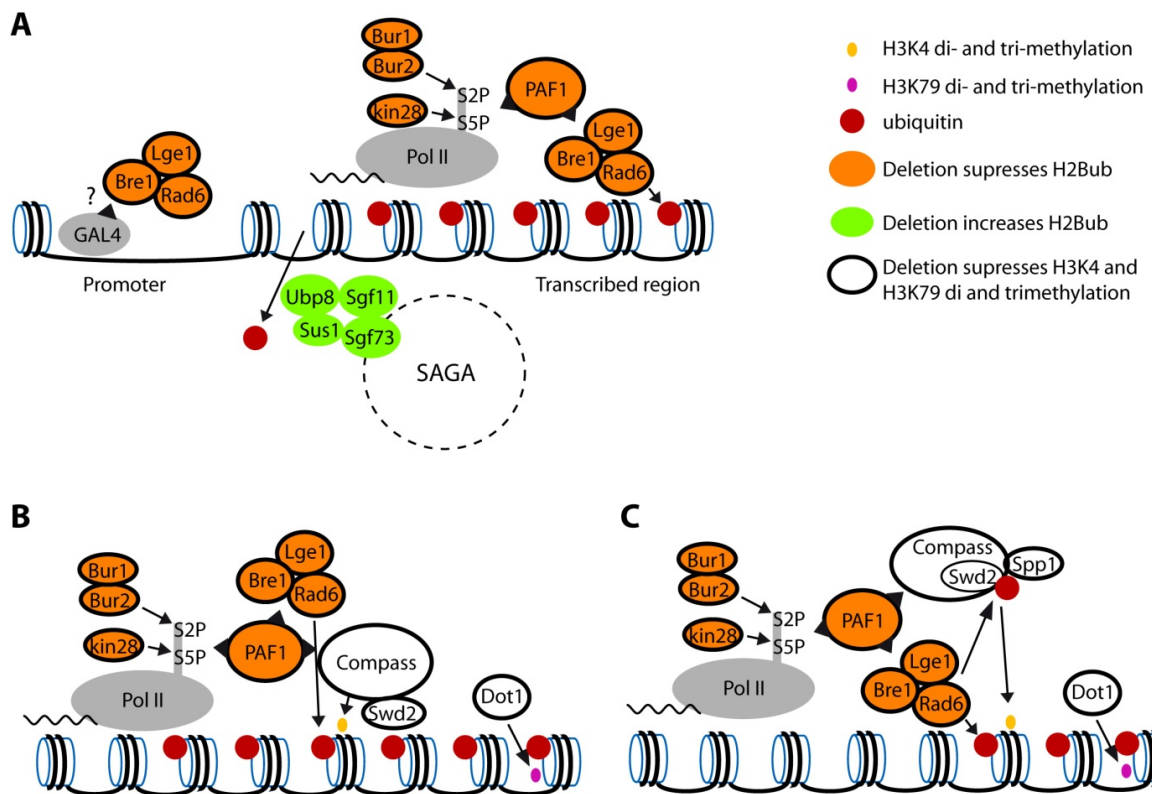
and H3K79 methylation that is involved in telomeric silencing in *S. cerevisiae* has an equivalent function in other species. An indirect link between H2B ubiquitination and both mRNA export and gene gating has been suggested based on the observation that loss of the deubiquitination activity of the SAGA complex induced nuclear accumulation of the GAL1 mRNA and a reduced association the GAL1 locus with the nuclear periphery after gene induction<sup>25</sup>. Finally, two recent studies have implicated H2Bub in repair of DNA double stand breaks. The role of H2BUB in both transcription elongation and DNA repair could suggest a unifying mechanism based on local chromatin relaxation as a result of the mere histone ubiquitination. Such a mechanism was recently proposed by Fierz et al. who showed that just the addition of a ubiquitin moiety to histone H2B is sufficient to interfere with compaction of the 30 nm chromatin fiber and lead to an open, biochemically accessible fiber conformation.

1. Komander, D. (2009) The emerging complexity of protein ubiquitination. *Biochem Soc Trans* 37, 937-953
2. Komander, D., *et al.* (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* 10, 550-563
3. Sowa, M.E., *et al.* (2009) Defining the human deubiquitinating enzyme interaction landscape. *Cell* 138, 389-403
4. Ventii, K.H., and Wilkinson, K.D. (2008) Protein partners of deubiquitinating enzymes. *Biochem J* 414, 161-175
5. Jason, L.J., *et al.* (2002) Histone ubiquitination: a tagging tail unfolds? *Bioessays* 24, 166-174
6. Osley, M.A. (2006) Regulation of histone H2A and H2B ubiquitylation. *Brief Funct Genomic Proteomic* 5, 179-189
7. Hwang, W.W., *et al.* (2003) A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell* 11, 261-266
8. Robzyk, K., *et al.* (2000) Rad6-dependent ubiquitination of histone H2B in yeast. *Science* 287, 501-504
9. Wood, A., *et al.* (2003) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* 11, 267-274
10. Pan, X., *et al.* (2006) A DNA integrity network in the yeast *Saccharomyces cerevisiae*. *Cell* 124, 1069-1081
11. Tanny, J.C., *et al.* (2007) Ubiquitylation of histone H2B controls RNA polymerase II transcription elongation independently of histone H3 methylation. *Genes Dev* 21, 835-847
12. Zofall, M., and Grewal, S.I. (2007) HULC, a histone H2B ubiquitinating complex, modulates heterochromatin independent of histone methylation in fission yeast. *J Biol Chem* 282, 14065-14072
13. Kim, J., *et al.* (2009) RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* 137, 459-471
14. Baarends, W.M., *et al.* (2007) Increased phosphorylation and dimethylation of XY body histones in the Hr6b-knockout mouse is associated with derepression of the X chromosome. *J Cell Sci* 120, 1841-1851
15. Kim, J., *et al.* (2005) The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol Cell* 20, 759-770
16. Zhu, B., *et al.* (2005) Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell* 20, 601-611
17. Zhang, F., and Yu, X. (2011) WAC, a functional partner of RNF20/40, regulates histone H2B ubiquitination and gene transcription. *Mol Cell* 41, 384-397
18. Daniel, J.A., *et al.* (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* 279, 1867-1871
19. Henry, K.W., *et al.* (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* 17, 2648-2663
20. Helmlinger, D., *et al.* (2008) The *S. pombe* SAGA complex controls the switch from proliferation to sexual differentiation through the opposing roles of its subunits Gcn5 and Spt8. *Genes Dev* 22, 3184-3195
21. Weake, V.M., *et al.* (2008) SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the *Drosophila* visual system. *EMBO J* 27, 394-405
22. Zhang, X.Y., *et al.* (2008) The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol Cell* 29, 102-111
23. Zhao, Y., *et al.* (2008) A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell* 29, 92-101

24. Kohler, A., *et al.* (2006) The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Mol Biol Cell* 17, 4228-4236
25. Kohler, A., *et al.* (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* 10, 707-715
26. Kohler, A., *et al.* (2010) Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell* 141, 606-617
27. Lang, G., *et al.* (2011) The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements. *Mol Cell Biol* 31, 3734-3744
28. Samara, N.L., *et al.* (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* 328, 1025-1029
29. Emre, N.C., *et al.* (2005) Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing. *Mol Cell* 17, 585-594
30. Gardner, R.G., *et al.* (2005) Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin. *Mol Cell Biol* 25, 6123-6139
31. van der Knaap, J.A., *et al.* (2005) GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7. *Mol Cell* 17, 695-707
32. Sarkari, F., *et al.* (2009) EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. *PLoS Pathog* 5, e1000624
33. van der Knaap, J.A., *et al.* (2010) Biosynthetic enzyme GMP synthetase cooperates with ubiquitin-specific protease 7 in transcriptional regulation of ecdysteroid target genes. *Mol Cell Biol* 30, 736-744
34. Buszczak, M., *et al.* (2009) Drosophila stem cells share a common requirement for the histone H2B ubiquitin protease scrawny. *Science* 323, 248-251
35. Sridhar, V.V., *et al.* (2007) Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. *Nature* 447, 735-738
36. Kao, C.F., *et al.* (2004) Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. *Genes Dev* 18, 184-195
37. Xiao, T., *et al.* (2005) Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol Cell Biol* 25, 637-651
38. Minsky, N., *et al.* (2008) Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* 10, 483-488
39. Kharchenko, P.V., *et al.* (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480-485
40. Roudier, F., *et al.* (2011) Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. *EMBO J* 30, 1928-1938
41. Roy, S., *et al.* (2010) Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* 330, 1787-1797
42. Mutiu, A.I., *et al.* (2007) The role of histone ubiquitylation and deubiquitylation in gene expression as determined by the analysis of an HTB1(K123R) *Saccharomyces cerevisiae* strain. *Mol Genet Genomics* 277, 491-506
43. Lenstra, T.L., *et al.* (2011) The specificity and topology of chromatin interaction pathways in yeast. *Mol Cell* 42, 536-549
44. Shema, E., *et al.* (2008) The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes Dev* 22, 2664-2676
45. Shema, E., *et al.* (2011) RNF20 inhibits TFIIIS-facilitated transcriptional elongation to suppress pro-oncogenic gene expression. *Mol Cell* 42, 477-488
46. Briggs, S.D., *et al.* (2002) Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* 418, 498
47. Dover, J., *et al.* (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 277, 28368-28371

48. Ng, H.H., *et al.* (2002) Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *J Biol Chem* 277, 34655-34657
49. Sun, Z.W., and Allis, C.D. (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418, 104-108
50. Shahbazian, M.D., *et al.* (2005) Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. *Mol Cell* 19, 271-277
51. Frederiks, F., *et al.* (2008) Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. *Nat Struct Mol Biol* 15, 550-557
52. McGinty, R.K., *et al.* (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* 453, 812-816
53. Lee, J.S., *et al.* (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* 131, 1084-1096
54. Vitaliano-Prunier, A., *et al.* (2008) Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. *Nat Cell Biol* 10, 1365-1371
55. Song, Y.H., and Ahn, S.H. (2010) A Bre1-associated protein, large 1 (Lge1), promotes H2B ubiquitylation during the early stages of transcription elongation. *J Biol Chem* 285, 2361-2367
56. Krogan, N.J., *et al.* (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* 11, 721-729
57. Ng, H.H., *et al.* (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* 11, 709-719
58. Ng, H.H., *et al.* (2003) The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *J Biol Chem* 278, 33625-33628
59. Wood, A., *et al.* (2003) The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J Biol Chem* 278, 34739-34742
60. Larabee, R.N., *et al.* (2005) BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. *Curr Biol* 15, 1487-1493
61. Wood, A., *et al.* (2005) The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. *Mol Cell* 20, 589-599
62. Fleming, A.B., *et al.* (2008) H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Mol Cell* 31, 57-66
63. Pavri, R., *et al.* (2006) Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* 125, 703-717
64. Batta, K., *et al.* (2011) Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev* 25, 2254-2265
65. Wyce, A., *et al.* (2007) H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Mol Cell* 27, 275-288
66. Pirngruber, J., *et al.* (2009) CDK9 directs H2B monoubiquitination and controls replication-dependent histone mRNA 3'-end processing. *EMBO Rep* 10, 894-900





**Figure Legend:**

**Factors regulating histone H2B ubiquitination and downstream events**

A) Recruitment and activity of H2B ubiquitinating and deubiquitinating enzymes at gene promoter and transcribed regions

B-C) Mechanisms underlying the crosstalk between H2Bub and H3 methylation (B) according to Lee et al. <sup>53</sup> and (C) according to Vitaliano-Prunier et al. <sup>54</sup>.

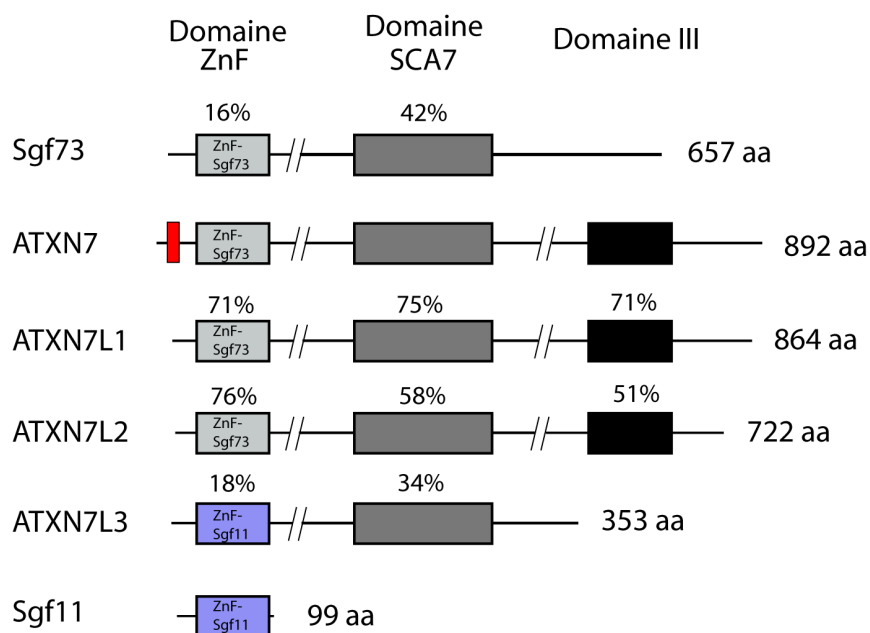
# **Résultats et Discussions**

## Partie 5 : Résultats

### I. Analyses structure fonction des domaines SCA7

#### Constat :

En 2004, il a été démontré que ATXN7 était l'orthologue de la protéine de levure Sgf73 et qu'elle était donc une sous-unité du complexe SAGA humain (Helmlinger et al, 2004b). Cette protéine, dont une forme mutée cause une maladie neurodégénérative appelée SCA7 (Ataxie Spinocérébelleuse de type 7), contient trois domaines fonctionnels séparés par des régions dont la séquence est peu conservée (Figure 31). Contrairement aux domaines N- et C-terminaux, la séquence du domaine central est très conservée au cours de l'évolution et ce domaine est retrouvé chez tous les orthologues de ATXN7 de la levure à l'homme, à l'exception de la drosophile. Ce domaine central a donc été appelé domaine SCA7 et il a été utilisé pour définir la famille des gènes homologues à *ATXN7* chez l'homme. Une analyse de type Blast a permis de mettre en évidence l'existence de deux protéines paralogues de ATXN7 (*ATXN7L1* et *ATXN7L2*), qui présentent une très



**Figure 31. La famille des protéines ATXN7.**

Représentation schématique de la famille des protéines ATXN7 chez l'homme et la levure *S. cerevisiae*. Pour chaque domaine conservé, le pourcentage d'identité de séquence par rapport à ATXN7 humaine est indiqué. Sgf73, ATXN7, ATXN7L1, ATXN7L2, ATXN7L3 possèdent toutes un domaine SCA7. Le domaine N-terminal de ATXN7L3 (qui contient un doigt de zinc) est par contre orthologue avec Sgf11. Le motif polyglutamine de ATXN7 est indiqué par un rectangle rouge.

forte conservation de séquence avec ATXN7 ainsi qu'une même organisation en domaines, ce qui suggère que ces deux protéines exercent la même fonction que ATXN7 (Figure 31) (Helmlinger et al, 2004b). Nos analyses récentes par spectrométrie de masse de purifications de SAGA montrent d'ailleurs que ATXN7L1 et ATXN7L2 sont des sous-unités du complexe. Nos résultats suggèrent également que l'on trouve une copie de l'un des trois paralogues (ATXN7, -L1 ou -L2) au sein d'un complexe SAGA indiquant que la fonction de ces trois protéines est redondante (Données non publiées du laboratoire de L. Tora). Un domaine SCA7 a également été identifié au sein d'une quatrième protéine qui a été appelée ATXN7L3 (Helmlinger et al, 2004b). De façon intéressante, une comparaison des séquences des quatre domaines SCA7 humains indique que les domaines SCA7 de ATXN7L1 et ATXN7L2 présentent un pourcentage d'identité très important avec celui de ATXN7 (respectivement 75% et 58%) alors que le domaine SCA7 de ATXN7L3 possède un pourcentage d'identité presque deux fois plus faible (34%) avec le domaine SCA7 d'ATXN7 (Helmlinger et al, 2004b). De plus, il apparaît que l'organisation en domaines de ATXN7L3 est différente de celle de ATXN7 (Figure 31) et des études additionnelles ont montré que le domaine N-terminal de ATXN7L3 était l'orthologue d'une autre sous-unité du DUBm de SAGA, la protéine Sgf11 (Zhao et al, 2008). Chez la levure, cette dernière protéine ne possède cependant pas de domaine SCA7 et celui-ci n'apparaît au cours de l'évolution qu'au sein des orthologues de Sgf11/ATXN7L3 chez certains insectes.

Nous avons également montré que chez l'homme, ATXN7 et ATXN7L3 sont deux sous-unités distinctes du DUBm de SAGA (Zhao et al, 2008), ce qui indique donc que dans la cellule, l'ensemble des domaines SCA7 sont retrouvés au sein de ce module de SAGA.

D'autre part, un alignement multiple de nombreux domaines SCA7 montre que trois cystéines et une histidine sont strictement conservées et il a été démontré que le domaine SCA7 d'ATXN7 lie un atome de zinc (Helmlinger et al, 2004b). Cependant, de façon remarquable, les résidus sensés coordonner l'atome de zinc présentent une configuration très particulière de type Cys-X<sub>9-10</sub>-Cys-X<sub>5</sub>-Cys-X<sub>2</sub>-His avec une très large insertion entre les deux premières cystéines du motif ce qui suggère que les domaines SCA7 sont des doigts de zinc atypiques et pose la question de leur organisation structurale.

Questions posées :

Nous avons donc voulu déterminer comment les doigts de zinc SCA7 de ATXN7 et ATXN7L3 sont repliés et comment les différences observées entre les séquences de ces deux domaines SCA7 se traduisent d'un point de vue structural. D'autre part, nous avons essayé de définir quelles étaient les fonctions exercées par chacun de ces deux domaines SCA7 au sein du DUBm de SAGA.

Résultats principaux :

Grâce à une collaboration avec le groupe de RMN de Bruno Kieffer à l'IGBMC, nous avons démontré que les domaines SCA7 de ATXN7 et ATXN7L3 présentent un nouveau repliement structural qui est spécifique à ces deux domaines SCA7. Ce repliement est défini par un doigt de zinc atypique, dans lequel la longue insertion entre les deux premiers résidus coordonnant le zinc n'est pas flexible mais forme un pseudo  $\beta$ -hairpin ce qui permet de positionner de façon optimale les résidus de coordination les uns par rapport aux autres. La seconde caractéristique de ce repliement est une petite hélice  $\alpha$ , présente dans les domaines SCA7 de ATXN7 et ATXN7L3, qui est maintenue à proximité du doigt de zinc par une interaction hydrophobe entre deux valines strictement conservées, une des valines étant située à proximité immédiate du doigt de zinc et l'autre étant intégrée à l'hélice  $\alpha$  (Figure 1 et 2 de l'article publié dans EMBO reports). Les deux domaines SCA7 partagent donc un cœur structural commun mais par contre, l'organisation des autres éléments de structure secondaire est très différente, en particulier la disposition relative de deux hélices  $\alpha$  caractéristique de chaque domaine. En effet, dans le domaine SCA7 de ATXN7, les deux hélices  $\alpha$  sont perpendiculaires alors que les deux hélices  $\alpha$  du domaine SCA7 de ATXN7L3 adoptent une conformation antiparallèle.

Ces différences structurales expliquent sans doute les divergences fonctionnelles que nous avons observées entre les deux types de domaines. En effet, nous avons démontré que le domaine SCA7 de ATXN7, mais pas celui de ATXN7L3, est capable d'interagir avec un nucléosome via le dimère H2A-H2B *in vitro*. Ceci a été démontré par des expériences de GST pull down ainsi que par des mesures biophysiques de RMN. Nous avons également démontré que la seconde hélice  $\alpha$  du domaine SCA7 de ATXN7, qui est absente chez ATXN7L3, est requise pour permettre l'interaction entre le domaine SCA7 et un nucléosome.

Pour préciser la fonction *in vivo* de cette interaction entre le domaine SCA7 de ATXN7 et un nucléosome, nous avons choisi de travailler chez la levure afin de pouvoir analyser l'effet de la délétion du domaine SCA7 de Sgf73, l'orthologue de ATXN7. Ce travail a été réalisé dans l'équipe de Marc Timmers à Utrecht. J'ai tout d'abord démontré que Sgf73 était parfaitement incorporée au sein de SAGA en l'absence de son domaine SCA7 ce qui indique que la fonction principale de ce domaine n'est pas d'interagir avec d'autres sous-unités de SAGA. D'autre part, nous avons analysé la capacité d'une souche délétée du domaine SCA7 à débiquitiner H2Bub (H2B ubiquitiné) et nous avons observé une augmentation modérée mais reproductible du niveau global de H2Bub dans cette souche mutante par rapport au niveau de H2Bub dans une souche sauvage. Ceci nous a permis de suggérer que l'interaction entre le domaine SCA7 de Sgf73/ATXN7 et un nucléosome pourrait participer à la régulation fine de l'activité de Ubp8/USP22, en positionnant de façon optimale le DUBm par rapport à son substrat (cf. Article publié dans la revue EMBO reports en 2010).

# The structural plasticity of SCA7 domains defines their differential nucleosome-binding properties

Jacques Bonnet<sup>1\*</sup>, Ying-Hui Wang<sup>2\*</sup>, Gianpiero Spedale<sup>3</sup>, R. Andrew Atkinson<sup>2</sup>, Christophe Romier<sup>2</sup>, Ali Hamiche<sup>1</sup>, W.W.M. Pim Pijnappel<sup>3</sup>, H.Th. Marc Timmers<sup>3</sup>, László Tora<sup>1</sup>, Didier Devys<sup>1+</sup> & Bruno Kieffer<sup>2++</sup>

<sup>1</sup>Department of Functional Genomics and <sup>2</sup>Department of Structural Biology and Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 964, Université de Strasbourg, Illkirch, France, and

<sup>3</sup>Department of Physiological Chemistry, University Medical Center Utrecht, Utrecht, The Netherlands

**SAGA (Spt-Ada-Gcn5 acetyltransferase), a coactivator complex involved in chromatin remodelling, harbours both histone acetylation and deubiquitination activities. ATXN7/Sgf73 and ATXN7L3, two subunits of the SAGA deubiquitination module, contain an SCA7 domain characterized by an atypical zinc-finger. We show that the yeast Sgf73-SCA7 domain is not required to recruit Sgf73 into SAGA. Instead, it binds to nucleosomes, a property that is conserved in the human ATXN7-SCA7 domain but is lost in the ATXN7L3 domain. The solution structures of the SCA7 domain of both ATXN7 and ATXN7L3 reveal a new, common zinc-finger motif at the heart of two distinct folds, providing a molecular basis for the observed functional differences.**

Keywords: SAGA; ATXN7; ubiquitin; zinc-finger; nucleosome  
EMBO reports (2010) 11, 612–618. doi:10.1038/embor.2010.98

## INTRODUCTION

Transcription initiation is regulated by the diverse functions of multiprotein coactivator complexes that include histone modification and nucleosome remodelling activities, as well as interactions with gene-specific activators and general transcription factors (Naar *et al.*, 2001; Martinez, 2002). SAGA (Spt-Ada-Gcn5 acetyltransferase) is one such coactivator complex that is highly conserved throughout evolution, with the histone acetyltransferase GCN5 being the catalytic subunit (Baker & Grant, 2007; Nagy & Tora, 2007). The structural organization of the SAGA complex revealed that its activities are localized in distinct functional modules (Wu *et al.*, 2004). This complex hosts a second enzymatic

activity, as it was shown to remove ubiquitin from mono-ubiquitinated histone H2B (Henry *et al.*, 2003; Daniel *et al.*, 2004). In agreement with this observation, the human ubiquitin protease USP22 (or its yeast orthologue Ubp8), together with ATXN7L3 (Sgf11), ENY2 (Sus1) and ATXN7 (Sgf73), forms a deubiquitination module (DUBm; Fig 1A; Powell *et al.*, 2004; Rodriguez-Navarro *et al.*, 2004; Ingvarsdottir *et al.*, 2005; Lee *et al.*, 2005; Zhang *et al.*, 2008; Zhao *et al.*, 2008). Analysis of ATXN7/Sgf73 protein sequences reveals the presence of two domains showing a strict conservation of cysteine and histidine residues (Fig 1B). The first amino-terminal domain contains a C2H2 zinc-finger (ZnF) and its deletion dissociates the DUBm from SAGA (Kohler *et al.*, 2008). This ZnF has an important structural role in the DUBm, mediating interactions with the three other components (Kohler *et al.*, 2010; Samara *et al.*, 2010). The second domain, hereafter referred to as the SCA7 domain (InterPro: IPR013243), shows an atypical sequence pattern with a Cys-X<sub>9-10</sub>-Cys-X<sub>5</sub>-Cys-X<sub>2</sub>-His motif, which is characterized by a long sequence insertion between the first two zinc-coordinating residues (Fig 1C; Helmlinger *et al.*, 2004). An SCA7 domain is also found in ATXN7L3, but not in its yeast orthologue Sgf11, suggesting that the ATXN7L3-SCA7 domain might have evolved to achieve a function distinct from that of ATXN7-SCA7 (Zhao *et al.*, 2008). Both the structure and the function of the two distinct SCA7 domains remain unknown.

To provide structural and functional insights into the SCA7 domains, we undertook a structure–function analysis of these atypical ZnFs. Our results unveil a new zinc-binding fold, and we link the striking structural differences observed between the two SCA7 domains to their ability to bind to nucleosomes.

## RESULTS AND DISCUSSION

### ATXN7 and ATXN7L3 contain distinct SCA7 domains

The SCA7 domain is found exclusively in members of the ATXN7 gene family, which includes two distinct subunits of SAGA complexes: ATXN7 and ATXN7L3 orthologues (Fig 1B). To examine the functional divergence of the SCA7 domains between ATXN7 and ATXN7L3 proteins, we performed multiple alignments of the corresponding sequences from different genomes. This analysis highlights the consensus signature for the

<sup>1</sup>Department of Functional Genomics and <sup>2</sup>Department of Structural Biology and Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 964, Université de Strasbourg, BP 10142-67404 Illkirch, France

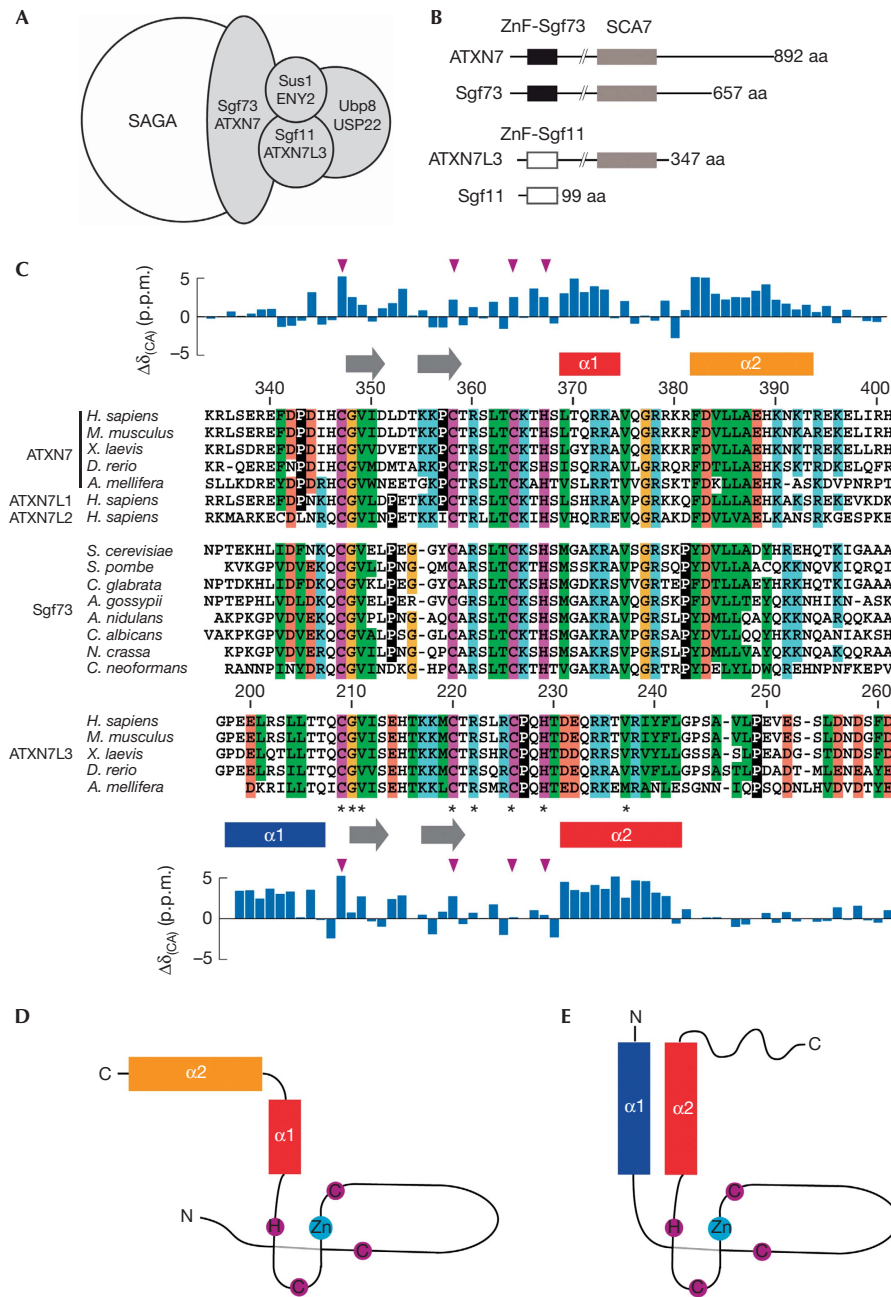
<sup>3</sup>Department of Physiological Chemistry, University Medical Center Utrecht, Universiteitsweg 100, Utrecht 3584 CG, The Netherlands

\*These authors contributed equally to this work

<sup>+</sup>Corresponding author. Tel: +33 388653441; Fax: +33 388653201; E-mail: devys@igbmc.fr

<sup>++</sup>Corresponding author. Tel: +33 368854722; Fax: +33 368854718; E-mail: kieffer@igbmc.fr

Received 26 October 2009; revised 1 June 2010; accepted 7 June 2010; published online 16 July 2010



**Fig 1** | Conservation of the SCA7 domain in ATXN7 and ATXN7L3 orthologues. (A) Schematic representation of the SAGA deubiquitination module. (B) Domain organization of ATXN7 homologues. The highly conserved SCA7 domain (grey box) is shared by all ATXN7 orthologues, as well as a subset of ATXN7L3 orthologues. Sequence conservation of the amino-terminal ZnF domain distinguishes ATXN7L3/Sgf11 orthologues (ZnF-Sgf11, white box) from ATXN7/Sgf73 orthologues (ZnF-Sgf73, black box). (C) Multiple sequence alignment of the SCA7 domain from selected species reveals the sequence similarities and differences that allow the SCA7 domains of ATXN7/Sgf73 orthologues to be distinguished from those of ATXN7L3 orthologues. Invariant residues are indicated by asterisks and zinc-chelating residues are coloured in magenta. Hydrophobic, positive, negative and glycine residues are coloured in green, blue, red and yellow, respectively. Prolines are shown in inverted contrast.  $\alpha$ -Helices are indicated as rectangles and extended segments as arrows, and the deviations of  $C\alpha$  chemical shifts from random coil values are represented above and below the ATXN7 and ATXN7L3 sequences. The positions of zinc-coordinating residues are indicated with magenta triangles. (D,E) Topologies of ATXN7-SCA7 (D) and ATXN7L3-SCA7 domains (E). SAGA, Spt-Ada-Gcn5 acetyltransferase; ZnF, zinc-finger.



**Table 1** | NMR and refinement statistics for protein structures

	ATXN7	ATXN7L3
<i>NMR distance and dihedral constraints</i>		
Distance constraints		
Total NOE	1589	970
Intra-residue	387	289
Inter-residue		
Sequential ( $ i-j  = 1$ )	474	383
Medium-range ( $ i-j  < 4$ )	376	167
Long-range ( $ i-j  > 5$ )	352	131
Protein–zinc restraints	10	10
Hydrogen bonds	0	0
Total dihedral angle restraints		
$\phi$	29	23
$\psi$	27	23
<i>Structure statistics</i>		
Violations (mean $\pm$ s.d.)		
Distance constraints (Å)	0.0664 $\pm$ 0.0034	0.0725 $\pm$ 0.0040
Dihedral angle constraints (°)	0.1706 $\pm$ 0.2998	0.1293 $\pm$ 0.1373
Max. dihedral angle violation (°)	4.935	2.441
Max. distance constraint violation (Å)	0.488	0.356
Deviations from idealized geometry		
Bond lengths (Å)	0.0177 $\pm$ 0.0009	0.0147 $\pm$ 0.0007
Bond angles (°)	2.0232 $\pm$ 0.0400	1.8581 $\pm$ 0.0471
Impropers (°)	1.9413 $\pm$ 0.0777	1.8466 $\pm$ 0.0865
Average pairwise r.m.s. deviation* (Å)		
Heavy	1.17 $\pm$ 0.16	1.44 $\pm$ 0.17
Backbone	0.60 $\pm$ 0.19	0.54 $\pm$ 0.12

NOE, nuclear Overhauser enhancement.

\*Pairwise r.m.s. deviations were calculated among 20 refined structures by using residues 342–394 for ATXN7–SCA7 and residues 199–211 and 220–244 for ATXN7L3–SCA7.

SCA7 domain (Fig 1C, asterisks), encompassing the putative zinc-coordinating residues, but also reveals the distinct features of the two proteins (Fig 1C). Marked differences are found mostly in the carboxy-terminal of the domain, suggesting that divergent evolution of the SCA7 ZnF domain occurred in order to achieve specific functions in the SAGA complex.

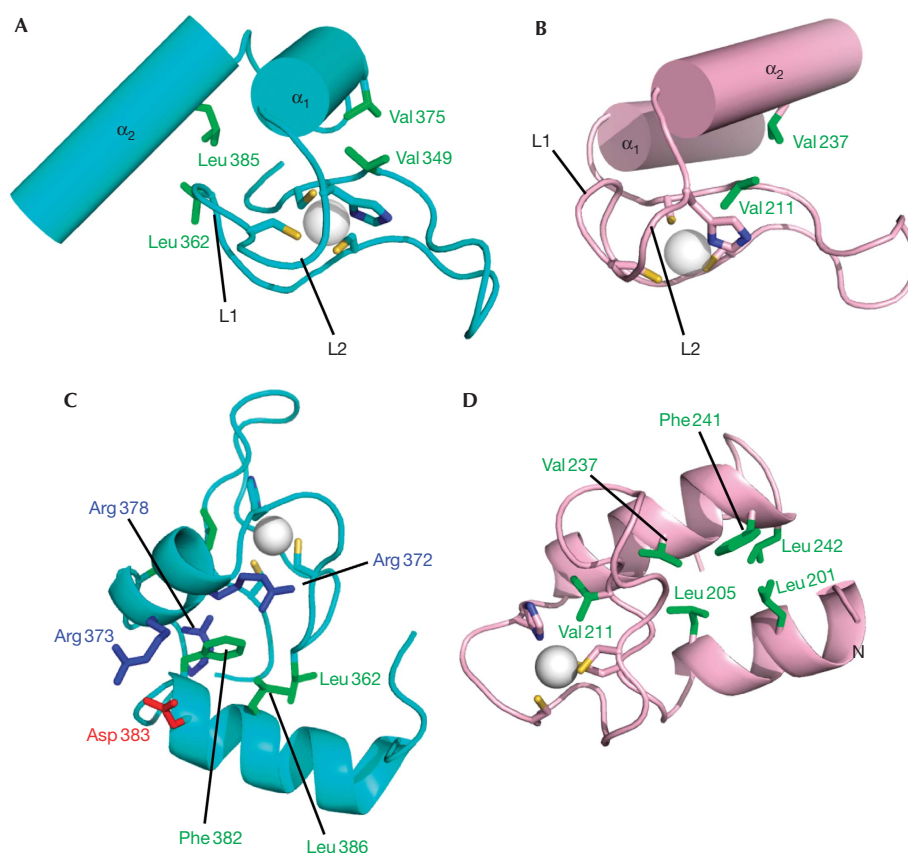
### SCA7 domains contain a new type of ZnF

On the basis of both sequence conservation analysis and biochemical screening, the SCA7 domains of ATXN7 and ATXN7L3, encompassing residues Lys 330–His 401 of ATXN7 (ATXN7–SCA7) and Gly 197–Gly 276 of ATXN7L3 (ATXN7L3–SCA7), were produced and analysed using NMR spectroscopy. Notably, analysis of  $C\alpha$  chemical shifts revealed the presence of disordered regions in both SCA7 domains, albeit not in the same region (Fig 1C). Whereas the first and last 10 residues of

ATXN7–SCA7 are not folded, the N-terminal region of ATXN7L3–SCA7 is well structured and the last 30 residues of this domain are not folded.

The structures of the two proteins were determined by using nuclear Overhauser enhancement spectroscopy (NOESY) spectra recorded at 950 MHz, allowing efficient automated analysis by ATNOS/CANDID software (Table 1 and supplementary Table S1 online). Both sets of calculations resulted in a well-defined fold as shown by the superimposition of the lowest energy models of ATXN7–SCA7 and ATXN7L3–SCA7 (supplementary Fig S1 online). The larger number of NOEs assigned for ATXN7–SCA7 reflects the greater number of folded residues that form an extended core, as compared with ATXN7L3–SCA7.

In both ATXN7–SCA7 and ATXN7L3–SCA7, the large sequence insertion between the first and second zinc-coordinating cysteines corresponds to a protruding extended hairpin structure (Fig 2A,B).



**Fig 2** | Solution structures of the SCA7 domains of human ATXN7 and ATXN7L3. (A,B) Ribbon representation of (A) ATXN7-SCA7 and (B) ATXN7L3-SCA7 domains showing the hydrophobic contacts that hook the two  $\alpha$ -helices onto the zinc-finger. (C,D) Conserved interactions define the relative orientations of the two  $\alpha$ -helices within (C) ATXN7-SCA7 and (D) ATXN7L3-SCA7 domains. The side chains of conserved hydrophobic, positive or negative residues are shown by using green, blue or red sticks, respectively. The figures were prepared using PyMOL (DeLano Scientific LLC).

The core of the zinc-binding site shows a conserved structure formed by two short adjacent loops (Fig 2A,B, L1 and L2) located at the bottom of the hairpin. This particular three-dimensional organization of zinc-coordinating residues has not been observed previously and represents a new fold, which seems to be specific to SCA7 domains as no homologous fold was found in the Protein Data Bank. The zinc-binding scaffold contains two conserved hydrophobic residues that hook onto the remaining parts of the domain. In particular, the valine residue of the SCA7-specific Cys-Gly-Val motif (Val349 in ATXN7, Val211 in ATXN7L3) interacts with the  $\alpha$ -helix located immediately downstream from the zinc-binding site and contributes to its orientation (Fig 2A,B). A second hook is provided by the Leu 362 located at the tip of the first loop of the zinc-coordination cage that anchors the second  $\alpha$ -helix in the ATXN7-SCA7 domain (Fig 2A).

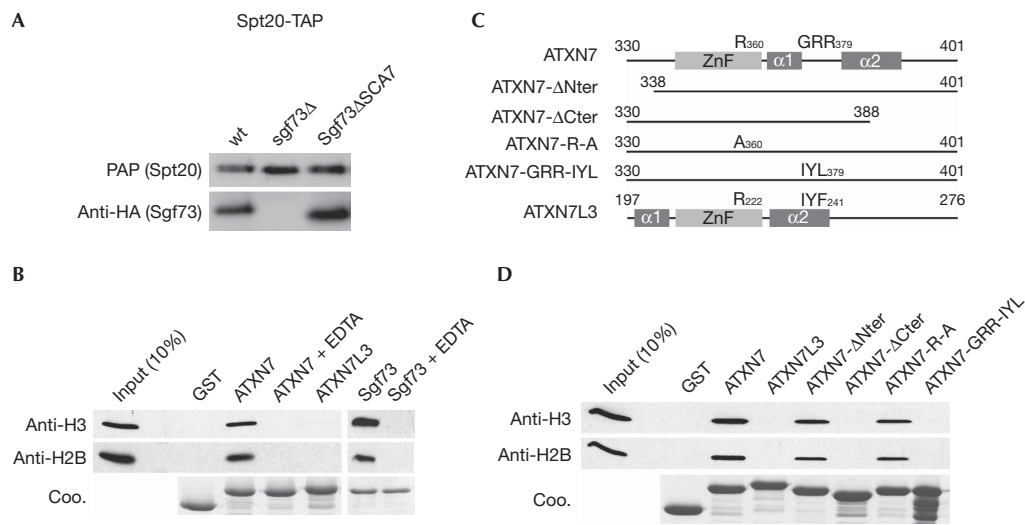
### SCA7 domains differ in their helical structures

Although the SCA7 domains of both ATXN7 and ATXN7L3 contain two  $\alpha$ -helices, these are not located at similar positions in the sequences. In ATXN7-SCA7, the two  $\alpha$ -helices (encompassing residues Leu 369–Ala 374 and Phe 382–Thr 393) are located downstream from the zinc-binding site and are separated by

a loop containing a large number of positively charged residues (Fig 1C,D). In ATXN7L3-SCA7, the two  $\alpha$ -helices (encompassing residues Pro 198–Thr 207 and Asp 231–Leu 242) lie to either side of the zinc-binding site (Fig 1C,E), leading to a different packing of the two helices (Fig 2C,D).

In ATXN7-SCA7, the two helices have an almost perpendicular orientation, the  $\alpha$ 2 helix being anchored to the zinc-binding site by the hydrophobic contact between Leu 362 and Leu 385 (Fig 2A). The closed conformation of the elbow formed by the two helices is due to the tight packing of several conserved residues, namely Phe 382 and Leu 386 in helix  $\alpha$ 2 and Arg 372 in helix  $\alpha$ 1 (Fig 2C). In addition, a conserved salt bridge is observed between Arg 373 in helix  $\alpha$ 1 and Asp 383 in helix  $\alpha$ 2. Notably, the side chain of Arg 378 is buried inside the elbow and might contribute to the closed conformation through a cation– $\pi$  interaction with the side chain of Phe 382.

In ATXN7L3-SCA7, the helices  $\alpha$ 1 and  $\alpha$ 2 adopt an anti-parallel orientation defined by hydrophobic interactions involving residues Leu 201 and Leu 205 in helix  $\alpha$ 1, and Val 237, Phe 241 and Leu 242 in helix  $\alpha$ 2 (Fig 2D). Interestingly, sequence analysis of the *Apis mellifera* ATXN7L3 orthologue suggests that the hydrophobic interaction between Leu 201 and Leu 242 is replaced



**Fig 3** | The SCA7 domains of human ATXN7 and yeast Sgf73 bind to nucleosomes. (A) Tandem affinity purifications (TAP) of SAGA using TAP-tagged Spt20 from the indicated strains were analysed by western blot using PAP or HA antibody. (B) SCA7 domains from ATXN7, ATXN7L3 or Sgf73 fused to GST or GST alone were immobilized on glutathione-Sepharose beads and incubated with mononucleosomes. The retained material was analysed by Coomassie blue staining (Coo.) and western blotting using H2B or H3 antibodies, as indicated. All interactions were lost when the fusion proteins were pre-incubated with EDTA solution. (C) Schematic representation of ATXN7-SCA7 mutant domains used in GST-pulldown experiments. (D) Mononucleosomes retained by immobilized SCA7 domains were analysed by western blotting using H3 and H2B antibodies. GST, glutathione S-transferase; HA, haemagglutinin; PAP, peroxidase anti-peroxidase; SAGA, Spt-Ada-Gcn5 acetyltransferase; ZnF, zinc-finger.

by an electrostatic interaction between a lysine and a glutamate residue (Fig 1C).

Analysis of the solution structures of ATXN7-SCA7 and ATXN7L3-SCA7 shows that sequence divergence between the two SCA7 domains translates into a major structural rearrangement of helical structures around a conserved zinc-coordination scaffold.

### The ATXN7-SCA7 domain binds to nucleosomes

We then asked whether these structural differences account for potential distinct functions of the SCA7 domains of ATXN7 and ATXN7L3. Although the function of the ATXN7L3-SCA7 domain has never been addressed, it has been suggested that SCA7 domains from both Sgf73 and ATXN7 are involved in the incorporation of these proteins into SAGA (Helmlinger *et al*, 2004; Kohler *et al*, 2008). To confirm this observation, we analysed a yeast strain in which the SCA7 domain of Sgf73 was deleted (Sgf73ΔSCA7). Tandem affinity purification (TAP) of TAP-tagged Spt20 from this mutant strain revealed normal incorporation of Sgf73ΔSCA7 (Fig 3A; supplementary Fig S2 online), indicating that Sgf73-SCA7 is dispensable for Sgf73-mediated incorporation of the DUBm into SAGA.

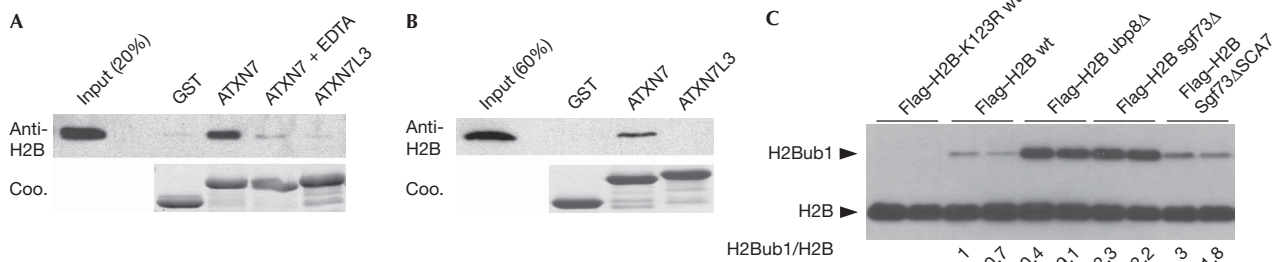
We then asked whether the ATXN7/Sgf73 SCA7 domain helps recruit nucleosomes to the DUBm of SAGA, as described for the SAGA histone acetyltransferase module that contains several chromatin-binding domains (Lee & Workman, 2007). In glutathione S-transferase (GST)-pulldown assays, nucleosomes were retained specifically by the GST-ATXN7-SCA7 fusion protein, but not by GST alone, nor by the fusion protein pretreated with EDTA solution (Fig 3B). Under the same experimental conditions, we observed the nucleosome-binding properties of

the SCA7 domain of Sgf73, whereas the ATXN7L3-SCA7 domain did not interact with nucleosomes (Fig 3B). Evidence of the nucleosome-binding property of the ATXN7-SCA7 domain was also obtained by NMR, as the addition of an increasing amount of mononucleosomes to a solution of <sup>15</sup>N-labelled ATXN7-SCA7 led to the progressive loss of free ATXN7-SCA7 resonances (supplementary Fig S3 online). The observation of a subset of correlation peaks, the intensity of which increases on addition of nucleosomes, suggests slow exchange between the bound and free states of ATXN7-SCA7. By assuming a diffusion-limited on-rate, an upper limit of 10 μM was obtained for the *K<sub>d</sub>*, suggesting a fairly strong interaction (supplementary Fig S3 online).

### Role of the C-terminal helix in nucleosome binding

The different organizations of the two helices found in the SCA7 domains of ATXN7 and ATXN7L3 suggested that the nucleosome-binding properties depend on sequence elements located outside the zinc-binding site. To confirm this, we generated and expressed mutant forms of the GST-ATXN7-SCA7 fusion protein, in which either the N-terminal or the C-terminal residues of the domain were deleted (ATXN7-ΔNter and ATXN7-ΔCter, respectively; Fig 3C).

GST-pulldown experiments revealed that the N-terminal deletion mutant interacted with nucleosomes as efficiently as the full-length domain. By contrast, deletion of the C-terminal residues of the domain completely abolished the interaction (Fig 3D). Superposition of the <sup>1</sup>H-<sup>15</sup>N heteronuclear single quantum coherence (HSQC) NMR spectrum of ATXN7-ΔCter on that of ATXN7-SCA7 revealed that the mutation disrupts α2 helix but leaves the rest of the structure unaffected, thus emphasizing the role of this helix in nucleosome binding



**Fig 4** | The yeast Sgf73–SCA7 domain regulates H2B deubiquitination. (A,B) GST-pulldown experiments using H2A–H2B dimers or tail-less histone octamers were analysed by Coomassie blue staining (Coo.) or western blotting using H2B antibodies. (C) Extracts of the indicated yeast strains carrying a Flag–HTB1 plasmid were separated on a 15% SDS–PAGE gel and subjected to western blot analysis using Flag antibodies. GST, glutathione *S*-transferase; SDS–PAGE, SDS–polyacrylamide gel electrophoresis.

(supplementary Fig S4 online). This result was reinforced by further mutagenesis experiments in which the Gly–Arg–Arg motif, found specifically in ATXN7 homologues (positions 377–379) in the elbow between the two helices (Fig 1C), was mutated to Ile–Tyr–Leu, reminiscent of the Ile–Tyr–Phe motif found in the ATXN7L3 domain. This mutation abolished nucleosome binding, whereas the mutation of the strictly conserved Arg360—located in the ZnF fold and exposed to solvent—to an alanine did not affect nucleosome binding, again suggesting that the core SCA7 ZnF is not involved directly in the interaction (Fig 3D). Together, these observations suggest that both the orientation and composition of the C-terminal helix in the SCA7 domain of ATXN7, which are strikingly different from those of ATXN7L3, have a role in nucleosome binding.

### Regulation of SAGA deubiquitination activity

We further demonstrated that the ATXN7–SCA7 domain interacted with H2A–H2B dimers (Fig 4A) and with octamers reconstituted from histones of which the N-terminal tails had been deleted (Fig 4B), but not with H3–H4 tetramers (supplementary Fig S5 online). Our *in vitro* binding experiments indicate that the ATXN7–SCA7 domain binds to the core or the C-terminal ends of the histone H2A and H2B dimer, a region located on the lateral face of the nucleosome that contains the ubiquitinated Lys 120 of H2B (Luger *et al*, 1997; Zhang, 2003). The ATXN7–SCA7 domain could therefore facilitate the recruitment of the SAGA DUBm to its substrate. To test this hypothesis, we analysed whether the deletion of the Sgf73–SCA7 domain affected the deubiquitination activity of SAGA. As reported previously, we observed a strong accumulation of monoubiquitinated H2B (H2Bub1) in *ubp8Δ* and *sgf73Δ* strains in which the SAGA deubiquitination activity is completely lost (Kohler *et al*, 2008; Fig 4C). We further analysed a mutant form of Sgf73 lacking the SCA7 domain (*Sgf73ΔSCA7*) but containing the N-terminal region that interacts with other subunits of the DUBm. In the *Sgf73ΔSCA7* strain, the mutant form of Sgf73 remained associated with SAGA (Fig 2A) and the global levels of H2Bub1 were reproducibly increased by 2.5-fold as compared with the wild-type strain (Fig 4C). Together, our results suggest that the nucleosome-binding properties of the ATXN7/Sgf73–SCA7 domain provide a further regulatory mechanism for USP22/Ubp8. Indeed, the first ZnF of Sgf73, together with other subunits of the DUBm, mediates the allosteric activation of Ubp8 (Bonnet *et al*,

2008; Kohler *et al*, 2008, 2010). In addition, we propose that the interaction of Sgf73–SCA7 with nucleosomes has a role in fine-tuning the deubiquitination activity of SAGA through optimal positioning of the DUBm relative to its substrate.

The structures of the SCA7 domains of ATXN7 and ATXN7L3 provide striking examples of the diversity and plasticity of ZnFs as protein interaction mediators. In contrast to previous observations, the nucleosome-binding ability that distinguishes the SCA7 domains of ATXN7 and ATXN7L3 is related to a significant change in the position of the helical structural elements around an otherwise conserved zinc-binding core. Indeed, whereas a unique prototypal structure can usually be associated with a given ZnF sequence signature, we found two distinct folds for the SCA7 sequence signature, providing the first example of a class of ZnF domains for which distinct interaction properties are associated with different folds.

### METHODS

**Plasmids.** Plasmids for the expression of GST–SCA7 fusion proteins (ATXN7: 330–401, ATXN7-ΔNter: 338–401 and ATXN7-ΔCter: 330–388) were described previously (Helmlinger *et al*, 2004). The DNA coding for SCA7 domains of Sgf73 (211–283) and of ATXN7L3 (197–276) were amplified by PCR and inserted into pGEX-4T1 (GE Healthcare). The *sgf73* genomic locus (ChrVII: 377258, 379763) was amplified by PCR from genomic DNA of a BY4742 strain and cloned into a pRS315 vector. This plasmid was modified to introduce the sequence coding for 3-haemagglutinin and a Myc Tag upstream from the Sgf73 stop codon. Mutants were generated by using the QuikChange Lightning Site-Directed Mutagenesis Kit (Stratagene).

**Expression and purification of SCA7 domains.** Expression of all proteins was achieved by using *Escherichia coli* BL21 [DE3] cells (Novagen) as described earlier (Nomine *et al*, 2005). The final purification step involved size exclusion chromatography using a Superdex G75 (HiLoad 16/60) gel filtration column equilibrated with phosphate buffer (50 mM phosphate, pH 7.0, 200 mM NaCl, 2 mM dithiothreitol) before concentration.

**NMR spectroscopy.** NMR samples contained 0.2 mM of <sup>13</sup>C,<sup>15</sup>N-labelled or unlabelled proteins. Most NMR experiments were recorded at 298 K on a Bruker DRX600 spectrometer. For structure calculations, homonuclear 2D NOESY spectra were recorded with a mixing time of 150 ms on a Bruker 950 MHz spectrometer.

The protonation state of the zinc-coordinating His367 of ATXN7 was set to  $N\epsilon_2$ , based on the frequencies of the  $^{15}N$  nuclei of the imidazole ring, which were measured using a long-range  $^1H$ - $^{15}N$  HSQC spectrum.

**Structure calculation.** 950 MHz 2D NOESY and 600 MHz 3D  $^{15}N$ - and  $^{13}C$ -edited NOESY spectra, together with backbone  $\Phi$  and  $\Psi$  angles derived from resonance assignments using the TALOS program (Cornilescu *et al*, 1999), were used as input data to the semi-automatic ATNOS/CANDID procedure (Herrmann *et al*, 2002). The final set of structures was refined in explicit water using the RECOORD protocol (Nederveen *et al*, 2005). The final ensemble of the 20 lowest energy structures contained no distance violations greater than 5 Å, nor dihedral angle violations greater than 5°. Three-dimensional structures of ATXN7-SCA7 and ATXN7L3-SCA7 have been deposited in the Protein Data Bank under accession numbers 2kkk and 2kkt, respectively.

**Generation of yeast strains.** The deletion of *sgf73* gene and the genomic integration of a TAP-Tag in the C-terminal of Spt20 were performed using standard methods involving the transformation of yeast with PCR fragments for homologous recombination. The constructs inserted in pRS315 and pRS413 were transformed, selected and grown on complete synthetic medium minus leucine and minus histidine, respectively. All strains were generated in a BY4742 background and are listed in supplementary Table S2 online.

**Analysis of the level of ubiquitinated H2B.** Two independent clones of each tested strain were grown in complete synthetic medium minus leucine and histidine to the midlog phase. A volume of 1 ml of each culture was resuspended in 100 µl of 200 mM NaOH for 5 min at room temperature (20 °C). Pellets of cells were then boiled for 5 min at 100 °C in 100 µl Laemmli buffer. In all, 10% of the whole-cell extract was analysed by western blot with the Flag antibody (M2, Sigma) to detect Flag-H2B. Quantification of the H2Bub1/H2B ratio was done by using ImageJ software.

**GST-pulldown assay and tandem affinity purifications.** GST-pulldown assay and tandem affinity purifications were performed according to standard procedures, with minor modifications (see the supplementary information online).

**Supplementary information** is available at *EMBO reports* online (<http://www.emboreports.org>).

#### ACKNOWLEDGEMENTS

We thank J. Koffler, G. Lang, E. Guittet, C. Ling, M. Stierle, M. Vitorino, X. Gao and M. Oulad-Abdelghani for their contribution to this work. Research in the authors' laboratories is supported by grants from the European Transcriptome, Regulome and Cellular Commitment Consortium, Institut National du Cancer, Agence Nationale de la Recherche, Structural Proteomics in Europe (SPINE-2), Ligue contre le cancer and NWO-Horizon. Access to high-field NMR spectrometers was granted by the CNRS through the Très Grands Equipements Résonance Magnétique Nucléaire programme.

#### CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

#### REFERENCES

Baker SP, Grant PA (2007) The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* **26**: 5329–5340  
Bonnet J, Romier C, Tora L, Devys D (2008) Zinc-finger UBPs: regulators of deubiquitylation. *Trends Biochem Sci* **33**: 369–375

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**: 289–302  
Daniel JA, Torok MS, Sun ZW, Schieltz D, Allis CD, Yates JR III, Grant PA (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* **279**: 1867–1871  
Helmlinger D *et al* (2004) Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet* **13**: 1257–1265  
Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NC, Kao CF, Pillus L, Shilatifard A, Osley MA, Berger SL (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* **17**: 2648–2663  
Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol NMR* **24**: 171–189  
Ingvarsdottir K, Krogan NJ, Emre NC, Wyce A, Thompson NJ, Emili A, Hughes TR, Greenblatt JF, Berger SL (2005) H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol Cell Biol* **25**: 1162–1172  
Kohler A, Schneider M, Cabal GG, Nehrbass U, Hurt E (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* **10**: 707–715  
Kohler A, Zimmerman E, Schneider M, Hurt E, Zheng N (2010) Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell* **141**: 606–617  
Lee KK, Workman JL (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**: 284–295  
Lee KK, Florens L, Swanson SK, Washburn MP, Workman JL (2005) The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol Cell Biol* **25**: 1173–1182  
Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260  
Martinez E (2002) Multi-protein complexes in eukaryotic gene transcription. *Plant Mol Biol* **50**: 925–947  
Naar AM, Lemon BD, Tjian R (2001) Transcriptional coactivator complexes. *Annu Rev Biochem* **70**: 475–501  
Nagy Z, Tora L (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**: 5341–5357  
Nederveen AJ *et al* (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* **59**: 662–672  
Nomine Y, Charbonnier S, Miguët L, Potier N, Van Dorsselaer A, Atkinson RA, Trave G, Kieffer B (2005)  $^1H$  and  $^{15}N$  resonance assignment, secondary structure and dynamic behaviour of the C-terminal domain of human papillomavirus oncoprotein E6. *J Biomol NMR* **31**: 129–141  
Powell DW, Weaver CM, Jennings JL, McAfee KJ, He Y, Weil PA, Link AJ (2004) Cluster analysis of mass spectrometry data reveals a novel component of SAGA. *Mol Cell Biol* **24**: 7249–7259  
Rodriguez-Navarro S, Fischer T, Luo MJ, Antunez O, Brettschneider S, Lechner J, Perez-Ortin JE, Reed R, Hurt E (2004) Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**: 75–86  
Samara NL, Datta AB, Berndsen CE, Zhang X, Yao T, Cohen RE, Wolberger C (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* **328**: 1025–1029  
Wu PY, Ruhlmann C, Winston F, Schultz P (2004) Molecular architecture of the *S. cerevisiae* SAGA complex. *Mol Cell* **15**: 199–208  
Zhang XY, Varthi M, Sykes SM, Phillips C, Warzecha C, Zhu W, Wyce A, Thorne AW, Berger SL, McMahon SB (2008) The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol Cell* **29**: 102–111  
Zhang Y (2003) Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev* **17**: 2733–2740  
Zhao Y *et al* (2008) A TIFC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell* **29**: 92–101

## Supplementary Methods

### GST-pull down assay

Glutathione sepharose beads (GE Healthcare) equilibrated in binding buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% NP40, 10  $\mu$ M ZnCl<sub>2</sub>, 1 mM DTT and protease inhibitors), were incubated with bacterial lysates containing GST or GST-SCA7 domains of hATXN7, hATXN7L3 or ySgf73. Mononucleosomes, recombinant H2A-H2B dimers and H3-H4 tetramers or recombinant tail-less histone octamers were prepared as described (Hamiche et al., 2001) and added to the GST-coated sepharose beads and incubated overnight at 4°C in binding buffer. Beads were washed with binding buffer containing 250 mM NaCl. 30% of the bound proteins were analysed by SDS-PAGE followed by Coomassie blue R250 coloration and 30% were analysed by immunoblot analysis. A monoclonal anti-H2B antibody (5HH2-2A8) was obtained by immunization of mice with a peptide corresponding to residues 111-125 (VSEGTKAVTKYTSSK) of H2B. An antibody against H3 (ab1791) was obtained from Abcam.

### Cell Extraction and Tandem Affinity Purifications

Cell extracts were prepared from *S. cerevisiae* strains TAP-tagged on Spt20 (WT,  $\Delta$ sgf73 and Sgf73 $\Delta$ SCA7, Supplementary Table 1). Cells were harvested from a 2 L suspension culture in CSM minus Leucine at an OD-600 of 2, disrupted by glass bead homogenization in extraction buffer (20 mM HEPES-NaOH, pH 8.0, 150 mM sodium chloride, 10% glycerol, 0.1% Tween 20, and Protease Inhibitor Cocktail (Sigma-Aldrich)) and cleared by ultracentrifugation. 40  $\mu$ L of biotinylated rabbit anti goat IgG (Bethyl Laboratories) were incubated with 400  $\mu$ L of Dynabeads® M-280 Streptavidin (Invitrogen). After extensive washes in extraction buffer, the resin was incubated for 2 h with 10 mL of cell extract and subsequently washed with 30 mL of extraction buffer and 10 mL of cleavage buffer (10 mM Tris-HCl, pH 8.0, 150 mM sodium chloride, 0.5 mM EDTA, pH 8.0, 0.1% Tween 20 and 1 mM DTT). The protein A

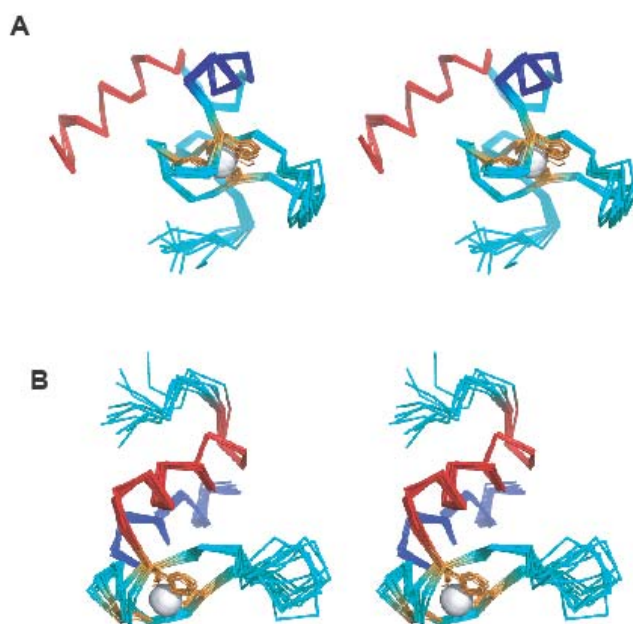
moiety of the tag was cleaved off with 10  $\mu$ L of TEV protease (Invitrogen) in 1 mL of cleavage buffer for 2 h. Binding to calmodulin agarose (Stratagene) for 1 h in binding buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM magnesium acetate, 1 mM imidazole, 2 mM calcium chloride, 10% glycerol, 0.1% Tween 20 and 10 mM mercaptoethanol) was followed by washing with 30 mL of binding buffer and elution in 400  $\mu$ L of elution buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM magnesium acetate, 1 mM imidazole, 2 mM EGTA, pH 8.0, 10% glycerol, 0.1% Tween 20 and 10 mM mercaptoethanol). The eluate was precipitated according to (Wessel and Flugge, 1984) and analysed on a NuPAGE® Novex 4-12% Bis-Tris Gel (Invitrogen), revealed by Coomassie blue staining or by western blotting using PAP (Sigma) or HA antibody (3F10, Roche).

## Supplementary References

- Hamiche, A., Kang, J.G., Dennis, C., Xiao, H. and Wu, C. (2001) Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. *Proc Natl Acad Sci U S A*, **98**, 14316-14321.
- Wessel, D. and Flugge, U.I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem*, **138**, 141-143.

**Supplementary Figure 1 : Solution structures of the SCA7 domains of human ATXN7 and ATXN7L3**

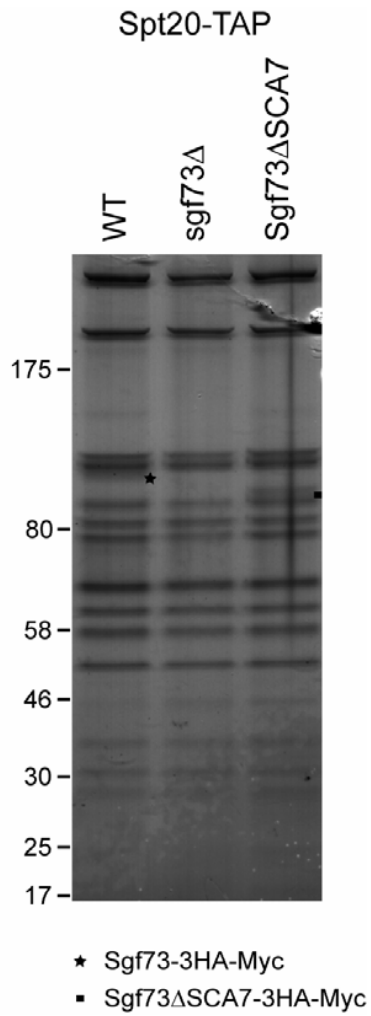
(**A,B**) Stereoview of the  $C_{\alpha}$  traces from the ten lowest energy structures of (**A**) ATXN7-SCA7 and (**B**) ATXNL3-SCA7 ZnFs. Only the folded parts of the proteins are represented. The heavy atoms of the zinc-coordinating residues are shown in orange and the position of the zinc ion is indicated by a gray sphere. Helices  $\alpha_1$  and  $\alpha_2$  are coloured in blue and red respectively.





**Supplementary Figure 2 : Sgf73-SCA7 is not required for Sgf73 incorporation into SAGA**

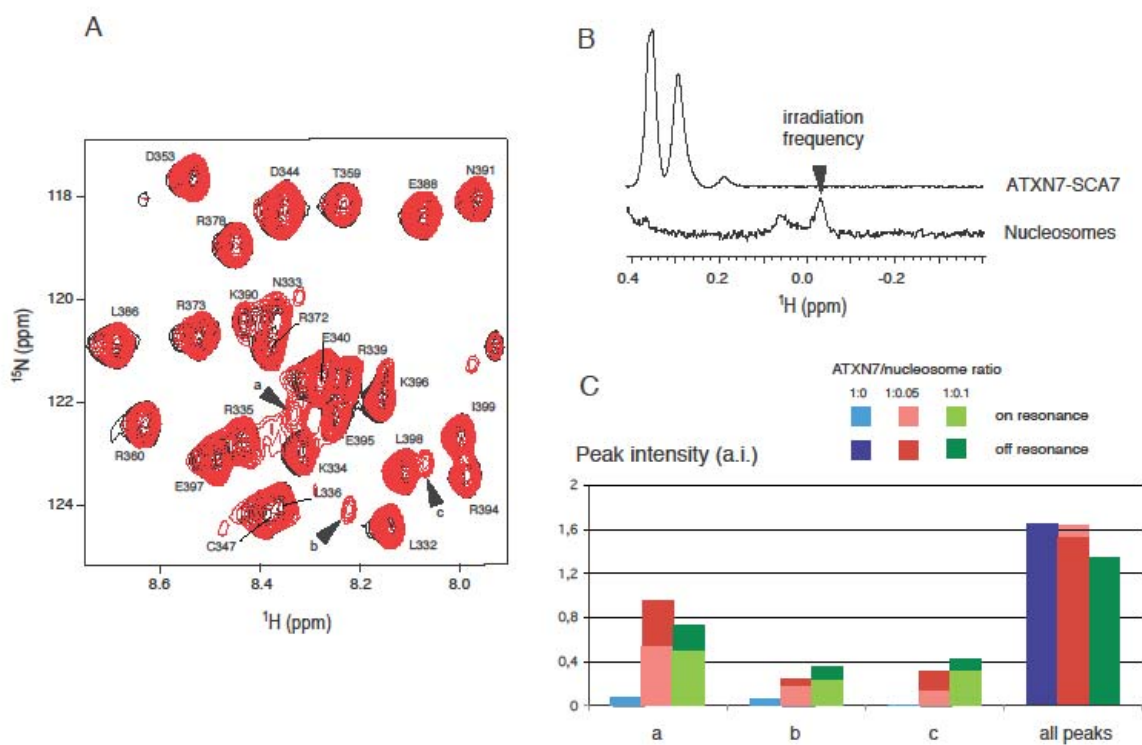
Tandem affinity purification of SAGA complexes using Tap-tagged Spt20 from the indicated strains were analyzed by Coomassie staining.



### Supplementary Figure 3 : Interaction of ATXN7-SCA7 domain with nucleosomes

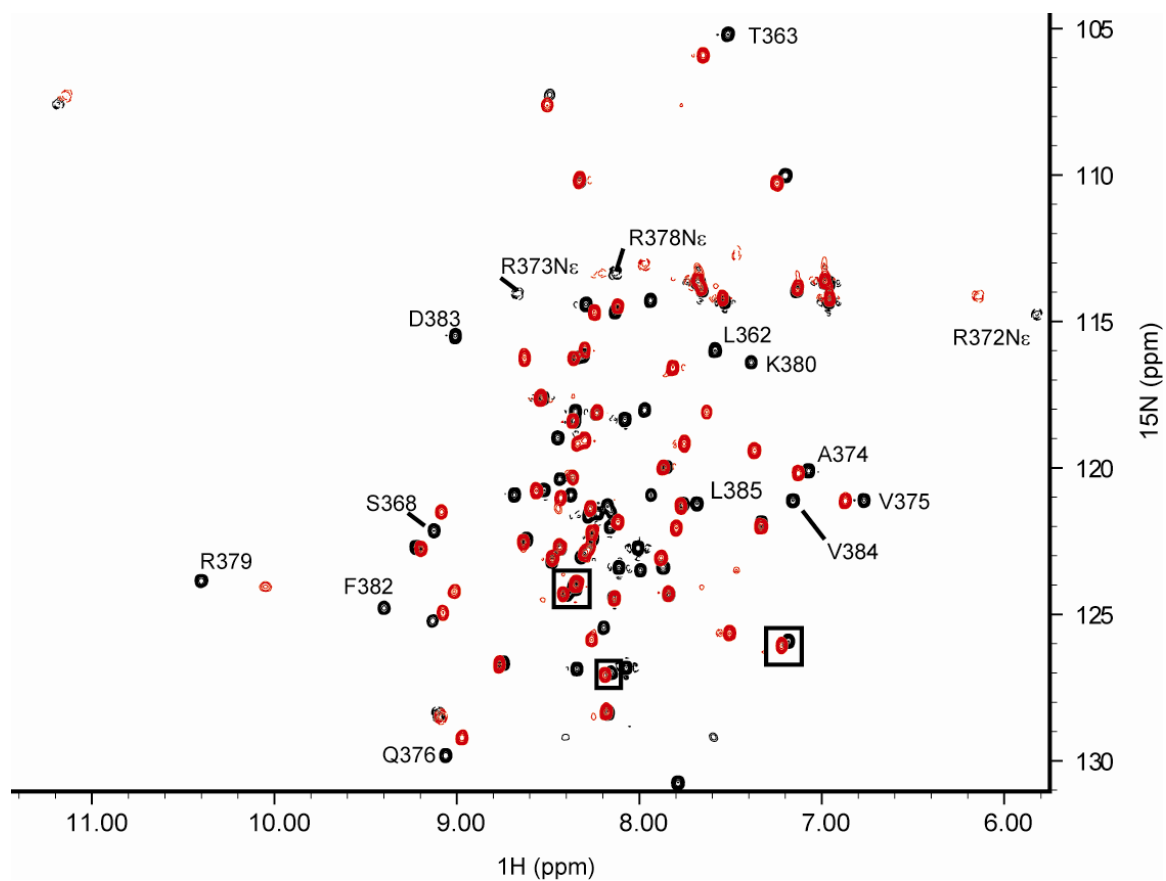
(A) The interaction between the SCA7 domain of ATXN7 and nucleosome particles was probed using NMR by mixing  $^{13}\text{C}$ ,  $^{15}\text{N}$  uniformly labelled ATXN7-SCA7 domain with different amounts of a preparation of purified nucleosomes. (B) The presence of an isolated methyl resonance at -0.015 ppm observed in the NMR spectrum of the nucleosome preparation was used to assess the relative concentration of nucleosome as compared to that of ATXN7 (B). The ratio of 1:30 (stock solution of nucleosome versus ATXN7) was found to be compatible to the 1:20 ratio found by Bradford assay. The absence of peak overlap between free ATXN7 and the nucleosome signal allowed specific saturation of the nucleosome particles using selective irradiation at -0.015 ppm during the relaxation time of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments. (C) Peak intensities were measured in  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of ATXN7-SCA7 domains recorded in absence and with two different substœchiometric amounts of nucleosome particles. In addition, peak intensities were recorded under a selective irradiation of the histone methyl resonance (light blue, red and green colours) or under off-resonance (-13.3 ppm) irradiation (dark blue, red and green colours). While the addition of increasing amounts of nucleosomes leads to an overall decrease of ATXN7 signals, a set of additional peaks, with increasing intensities could be observed in the region of the HSQC spectrum that corresponds to unfolded parts of the ATXN7 protein (see peaks labelled a,b and c in (A)). The intensity of these peaks is affected by nucleosome on-resonance irradiation, suggesting that they originate from a nucleosome bound form of ATXN7. Furthermore, analysis of frequencies of these secondary peaks suggests that they probably correspond to residues located at the edges of the SCA7 domain, where sufficient protein backbone flexibility is present in the complex to allow their observation, while frequencies are slightly affected by the nucleosome environment. The observation of a slow exchange regime between

free and nucleosome bound forms of ATXN7 suggests a tight binding involving all residues of the SCA7 domain, with an upper limit of the  $K_d$  of about 10  $\mu$ M.



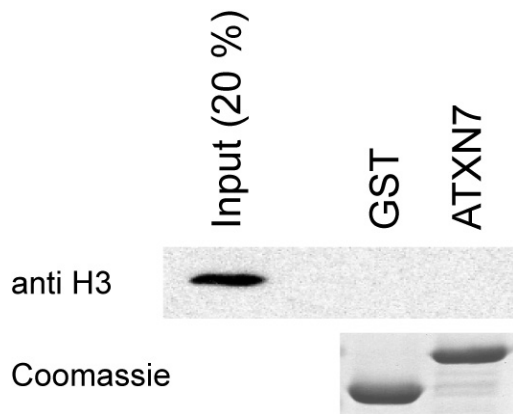
### Supplementary Figure 4 :

Superposition of HSQC spectra from ATXN7-SCA7 (black) and the mutant ATXN7- $\Delta$ Cter (red). Correlations with significant shifts upon mutation are labelled. Zinc-coordinating residues are boxed.



**Supplementary Figure 5 : ATXN7-SCA7 does not interact with H3H4 tetramers *in vitro***

GST-pull down experiments using H3-H4 tetramers were analyzed by Coomassie blue staining or western blotting using anti-H3 antibodies.



## Supplementary Table 1

Ramachandran statistics of ATXN7 and ATXN7L3 structures

	ATXN7	ATXN7L3
<b>Backbone dihedral angles statistics</b>		
Residues in most favorable regions (%)	72.4	76.0
Residues in additional favorable regions(%)	27.1	21.0
Residues in generously favorable regions (%)	0.6	2.0
Residues in disallowed regions (%)	0.0	1.1

## Supplementary Table 2

Name	Strain	Genotype	Source
	YGL066W	<i>Mat a; his3D1; leu2D0; lys2D0; ura3D0; sgf73Δ::Kan</i>	Euroscarf
	YMR223W	<i>Mat a; his3D1; leu2D0; lys2D0; ura3D0; ubp8Δ::Kan</i>	Euroscarf
Spt20-TAP sgf73Δ	YJB100	Isogenic to YGL066W except <i>SPT20-TAP::URA3; pRS315</i>	This study
Spt20-TAP WT	YJB101	Isogenic to YGL066W except <i>SPT20-TAP::URA3; pRS315-Sgf73-3HA-Myc</i>	This study
Spt20-TAP Sgf73ΔSCA7	YJB103	Isogenic to YGL066W except <i>SPT20-TAP::URA3; pRS315-Sgf73Δ 227-281-3HA-Myc</i>	This study
Flag-H2B sgf73Δ	YJB104	Isogenic to YGL066W except <i>pRS315; pRS413-FLAG-HTB1-HTA1</i>	This study
Flag-H2B WT	YJB105	Isogenic to YGL066W except <i>pRS315-Sgf73-3HA-Myc; pRS413-FLAG-HTB1-HTA1</i>	This study
Flag-H2B Sgf73ΔSCA7	YJB106	Isogenic to YGL066W except <i>pRS315-Sgf73Δ227-281-3HA-Myc; pRS413-FLAG-HTB1-HTA1</i>	This study
Flag-H2B-K123R WT	YJB107	Isogenic to YGL066W except <i>pRS315-Sgf73-3HA-Myc; pRS413-FLAG-htb1-K123R-HTA1</i>	This study
Flag-H2B ubp8Δ	YJB108	Isogenic to YMR223W except <i>pRS315; pRS413-FLAG-HTB1-HTA1</i>	This study

## II. Régulation de l'activité de USP22 par les sous-unités du DUBm

### Constat :

En 2008, nous avons identifié la déubiquitinase USP22 au sein du complexe SAGA chez l'homme et montré que celle-ci interagissait avec d'autres sous-unités de SAGA, à savoir ATXN7, ATXN7L3 et ENY2 pour former un module de déubiquitination (DUBm) (Zhao et al, 2008). Une des observations à la base de ce projet est que l'enzyme recombinante USP22 isolée n'est pas capable de déubiquitiner H2B *in vitro*, alors que dans les mêmes conditions, un complexe SAGA purifié à partir de cellules HeLa présente une activité de déubiquitination. Ceci suggère donc qu'il existe un mécanisme pour réguler l'activité enzymatique de USP22.

D'autre part, en position N-terminale par rapport à son domaine catalytique, USP22 possède un domaine à doigt de zinc, appelé ZnF-UBP et il est intéressant de noter que chez l'homme, une dizaine d'autres USP possèdent également un domaine ZnF-UBP. Une interaction entre plusieurs de ces domaines et l'extrémité C-terminale de l'ubiquitine libre ayant été observée, il a été suggéré que les domaines ZnF-UBP étaient des domaines de liaison de l'ubiquitine. De plus, une élégante étude structure fonction a démontré que l'interaction entre le domaine ZnF-UBP de USP5 et l'extrémité C-terminale de l'ubiquitine libre pouvait stimuler l'activité de cette enzyme (Reyes-Turcu et al, 2006).

### Question posée :

Nous avons donc voulu savoir si le domaine ZnF-UBP de USP22 participait au mécanisme d'activation de USP22, et si l'extrémité C-terminale de l'ubiquitine ainsi que les autres sous-unités du DUBm pouvaient jouer un rôle dans ce mécanisme de régulation enzymatique.

### Résultats principaux :

Dans une première étude, nous nous sommes intéressés au domaine ZnF-UBP de USP22. A partir de données de la littérature ainsi que d'un alignement multiple que j'ai généré et qui comprend l'ensemble des séquences des domaines ZnF-UBP codés par les génomes de l'homme et de la levure *S. cerevisiae*, nous avons observé que les résidus impliqués dans la liaison de l'ubiquitine par le ZnF-UBP de USP5 n'étaient pas conservés chez les domaines ZnF-UBP de USP22 et de son orthologue de levure Ubp8. Nous avons

donc émis l'hypothèse que la fonction du domaine ZnF-UBP de USP22 était de réaliser des interactions avec les autres sous-unités du DUBm ce qui devait permettre de stimuler l'activité enzymatique de USP22 (cf. Article publié dans la revue Trends in Biochemical Sciences en 2008).

Dans un deuxième temps, j'ai voulu démontrer expérimentalement que ce modèle était correct. Pour cela, nous avons généré différents baculovirus qui m'ont permis de purifier, à partir de cellules d'insectes, un DUBm recombinant contenant les protéines entières USP22 et ENY2 ainsi que les domaines N-terminaux de ATXN7 et de ATXN7L3. Nous avons observé que la formation d'un DUBm stable reposait sur un réseau complexe d'interactions entre les deux domaines de USP22 et les autres sous-unités du DUBm. D'autre part, les domaines N-terminaux de ATXN7 et de ATXN7L3 possèdent tous deux un doigt de zinc et nous avons montré que le doigt de zinc de ATXN7, mais pas celui de ATXN7L3, était impliqué dans la formation d'un module stable.

J'ai ensuite développé un test de déubiquitination semi-quantitatif, basé sur l'utilisation d'ubiquitine vinyl sulfone (UbVS), un substrat suicide pouvant interagir de façon covalente avec la forme active de nombreuses USP. Ce test m'a permis d'évaluer l'activité des différentes versions du DUBm que j'ai purifiées et nous avons démontré que la présence des quatre sous-unités du DUBm était requise pour obtenir l'activité optimale de USP22 dans des conditions données. Nous avons également observé que le doigt de zinc de ATXN7L3 participait de façon directe à la stimulation de l'activité de USP22. De plus, mes résultats indiquent que la présence de ce domaine est nécessaire pour que le DUBm puisse déubiquitiner H2Bub dans le contexte d'un nucléosome *in vitro*.

L'ensemble de ces données montre donc que chez l'homme, la formation d'un DUBm intègre permet d'activer USP22 et que le doigt de zinc du domaine N-terminal de ATXN7L3 joue un rôle direct dans ce mécanisme de régulation allostérique. Il apparaît également que ce doigt de zinc doit sans doute permettre au DUBm de reconnaître spécifiquement un nucléosome ubiquitiné (cf. Article publié dans la revue Molecular and Cellular Biology en 2011).

En parallèle de cette étude *in vitro*, Guillaume Lang a analysé la fonction du DUBm de SAGA *in vivo*. Pour cela, il a tiré avantage du fait que ATXN7L3 est indispensable à la formation du DUBm ainsi qu'à l'activation enzymatique de USP22. Nous avons ainsi pu



supprimer l'activité DUB de SAGA en générant une lignée cellulaire (appelée lignée shATXN7L3) qui exprime de façon stable un petit ARN en épingle à cheveux qui permet la dégradation de l'ARNm endogène de ATXN7L3 par ARN interférence. Dans ces cellules, Guillaume a démontré que le niveau global de H2Bub augmentait près de 5 fois par rapport à la situation sauvage et que le niveau global de H2Aub semblait aussi augmenter, mais de façon beaucoup plus modérée que dans le cas de H2B. Ceci démontre donc que SAGA est la principale déubiquitinase de H2B *in vivo* et que sa fonction ne peut pas être compensée par d'autres déubiquitinases.

Dans un second temps, nous avons étudié un groupe de gènes qui sont régulés par un stress du réticulum endoplasmique et qui sont connus pour être dépendants de SAGA (Nagy et al, 2009). Guillaume Lang a ainsi pu démontrer que l'activité DUB de SAGA était requise pour permettre une induction optimale de ces gènes. Pour déterminer ensuite où le DUBm déubiquitine H2B *in vivo*, il a comparé le niveau de H2Bub sur les promoteurs et la région transcrite de ces gènes dans des cellules contrôles ainsi que des cellules shATXN7L3 en réalisant des expériences d'immunoprécipitation de la chromatine (ChIP) avec un anticorps reconnaissant de façon très spécifique H2Bub (Minsky et al, 2008). Ces résultats indiquent que le niveau de H2Bub augmente de façon très importante sur la région transcrite des gènes analysés et de façon beaucoup moins marquée au niveau des promoteurs de ces gènes. D'autre part, des expériences de ChIP réalisées avec des anticorps dirigés contre des sous-unités de SAGA ne permettent de détecter la présence du complexe SAGA qu'au niveau des promoteurs, après l'induction des gènes analysés.

Ces résultats préliminaires suggèrent donc que pour des gènes SAGA dépendants, ce complexe déubiquitine l'histone H2B sur l'ensemble de la séquence du gène alors que par ChIP, nous ne détectons SAGA qu'au niveau du promoteur correspondant. Cette observation semble indiquer que le complexe SAGA est probablement recruté de façon stable sur le promoteur (où il est détecté par ChIP) mais qu'il agit sur l'ensemble de la séquence transcrite de ce gène (où il ne serait pas détectable par ChIP en raison d'un temps de résidence très court) (cf. Article publié dans la revue *Molecular and Cellular Biology* en 2011).

# Zinc-finger UBPs: regulators of deubiquitylation

Jacques Bonnet<sup>1</sup>, Christophe Romier<sup>2</sup>, László Tora<sup>1</sup> and Didier Devys<sup>1</sup>

<sup>1</sup> Department of Functional Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 596, Université Louis Pasteur de Strasbourg, BP 10142-67404 ILLKIRCH Cedex, CU de Strasbourg, France

<sup>2</sup> Department of Structural Biology and Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 596, Université Louis Pasteur de Strasbourg, BP 10142-67404 ILLKIRCH Cedex, CU de Strasbourg, France

**Deubiquitylating enzymes have key regulatory roles in multiple cellular processes by mediating ubiquitin removal and processing. The ubiquitin-specific processing proteases (USPs) represent the largest subclass of deubiquitylases. Recently, several USPs that recognize the monoubiquitylated histones H2A and/or H2B have been identified. Among these enzymes, three USPs contain a zinc-finger ubiquitin-specific protease (ZnF-UBP) domain, indicating that this domain plays a crucial part in regulating their activity. To address the putative function of this domain, we systematically analysed and aligned yeast and human ZnF-UBP-containing proteins. By complementing our analysis with structural and functional data, we present a classification of the different ZnF-UBP-containing proteins and a model for their regulation.**

## Different classes of deubiquitylating enzymes

Post-translational modification of intracellular proteins by ubiquitin is a key regulator of many cellular processes [1,2]. Ubiquitin is a highly conserved 76-amino acid polypeptide that is covalently attached to target proteins via an isopeptide bond between the C-terminal glycine of ubiquitin and the amino group of a lysine in substrate proteins. This occurs through a three-step process involving ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin ligase (E3) enzymes. Proteins can be monoubiquitylated (attachment of a single ubiquitin molecule), multiubiquitylated (attachment of single ubiquitin molecules to different lysine residues of the protein) or polyubiquitylated through alternative linkage leading to the regulation of different functions. Besides polyubiquitylation mediated by linkage through Lys48 and its role in protein degradation, other types of ubiquitylation are involved in many different cellular functions (reviewed in Refs [2–5]). For example, the ubiquitylation of histones H2A and H2B occurs mainly in a monoubiquitylated form and is not linked to degradation but, rather, regulates transcription initiation and elongation, silencing and DNA repair [6,7].

Another level of regulation is provided by deubiquitylating enzymes (DUBs) that mediate the removal and processing of ubiquitin. There are now ~100 human DUBs that have been identified by *in silico* efforts and activity-

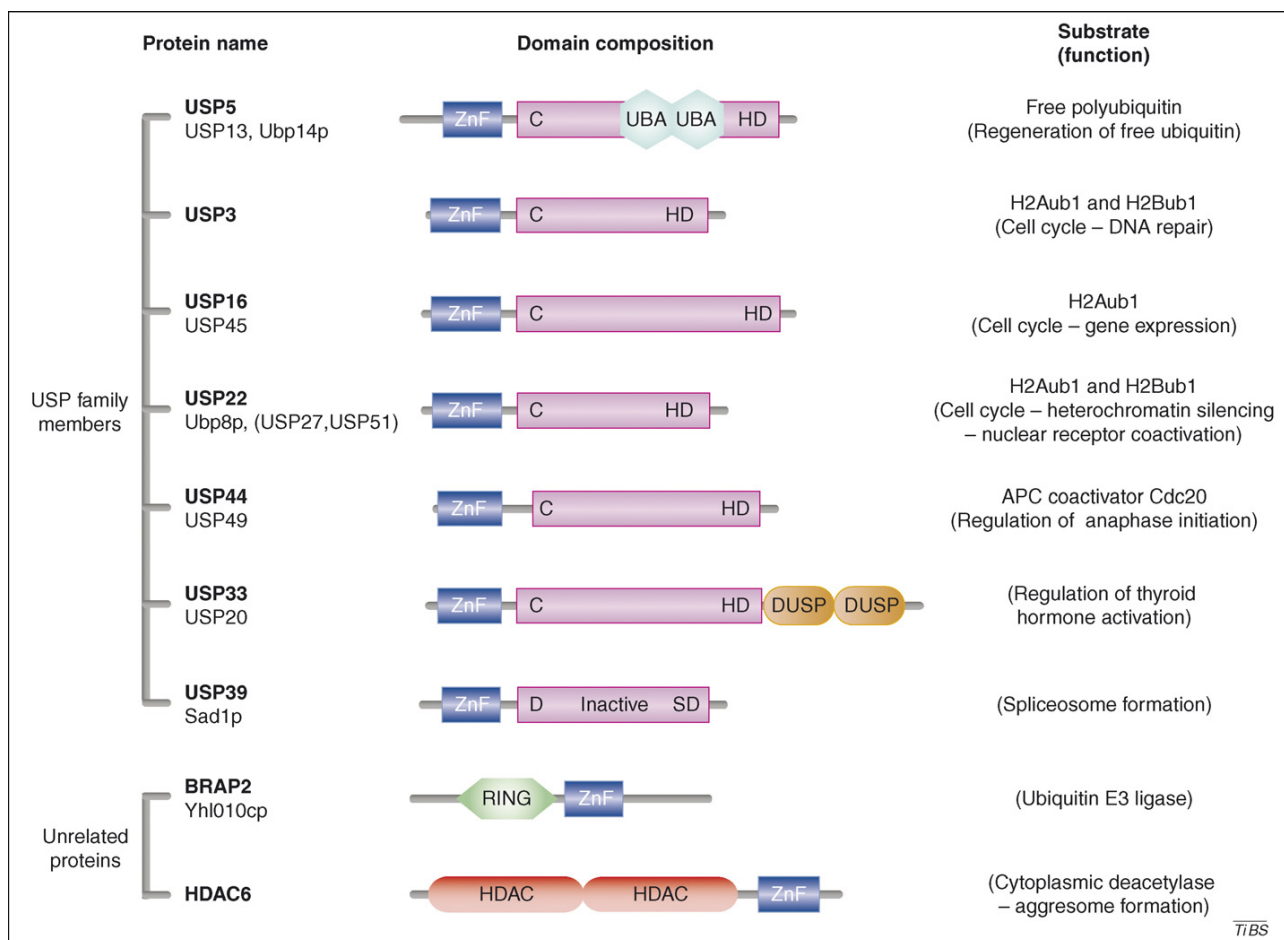
based profiling, which can be divided into five subclasses: (i) ubiquitin C-terminal hydrolases, (ii) ubiquitin-specific processing proteases (USPs), (iii) Machado-Joseph disease protein domain proteases, (iv) ovarian tumour proteases and (v) JAMM (JAB1/MPN/Mov34 metalloenzyme) motif proteases (reviewed in Ref. [8]). The largest subclass is the USPs, which contains >50 members and belongs to the cysteine protease family. All USPs have a catalytic domain characterized by two highly conserved histidine and cysteine boxes that contain the catalytic residues. However, the size of the protease domain varies from 300 to 800 amino acids between members, owing to the insertion of unrelated sequences [8]. The catalytic domain is surrounded by one or more additional domains, some of which contribute to target recognition, such as ubiquitin associated (UBA) or DUSP (domain in USPs) domains (Figure 1). The ZnF-UBP domain [also known as deacetylase and ubiquitin-specific protease (DAUP), polyubiquitin-associated zinc-finger (PAZ) or binder of ubiquitin zinc-finger (BUZ)] is another additional domain of ~100 residues found in several USPs, which might regulate their protease activity (reviewed in Refs [4,9]).

To better understand the function of ZnF-UBP domains, we performed searches for all budding yeast and human proteins containing this domain and analysed multiple alignments of their ZnF-UBP domains. By combining these data with known structural and functional data, we distinguished two classes of ZnF-UBPs that can or cannot bind ubiquitin. In addition, we propose a model in which USP catalytic activity is regulated by these domains through different mechanisms, depending on their ability to bind to the free C terminus of unanchored ubiquitin.

## The ZnF-UBP domain structure

Investigation of the budding yeast and human genomes has identified two yeast (Ubp14p and Ubp8p) and ten human USPs (USP5, USP13, USP3, USP16, USP45, USP22, USP44, USP49, USP33 and USP20) containing a ZnF-UBP domain (Figure 1). This domain is also found in two predicted protein sequences (USP27 and USP51) that are highly similar to USP22, but encoded by intronless genes. The functionality of these putative genes is therefore questionable (Figure 1). Interestingly, among USP family members the ZnF-UBP domain is always localized N-terminally to the catalytic domain. The conserved ZnF-UBP domain is also found in USP39, a protein related to

Corresponding authors: Tora, L. (laszlo@igbmc.u-strasbg.fr); Devys, D. (devys@igbmc.u-strasbg.fr).



**Figure 1.** Human and budding yeast ZnF-UBP-containing proteins. Proteins were classified according to their ZnF-UBP sequence similarities. A multiple alignment (Figure 2c) was used to generate a phylogenetic tree with the PHYLO\_WIN software [36]. Where the bootstrap value was significant (>75%), the sequences were grouped together. Within each class, we have arbitrarily chosen a representative sequence and indicated their domain composition. In addition to the ZnF-UBP domain (blue, ZnF), all USP family members contain a protease domain (purple) characterized by a catalytic Cys-His-Asp triad (CHD). In USP39, an inactive protease, the catalytic triad is mutated as indicated. UBA (light blue) and DUSP (orange) domains are additional domains found in some USPs. RING (green) and HDAC (red) domains are associated with ZnF-UBP in unrelated proteins. Substrates of the representative proteins are indicated at the right-hand side. The associated functions are indicated below in brackets. SwissProt accession numbers are as follows: USP5/IsoT, P45974 [14,18–21]; USP13/IsoT3, Q92995; Ubp14p, P38237; USP3, Q9Y6I4 [22]; USP16/Ubp-M, Q9Y5T5 [15,28,29]; USP45, Q70EL2; USP22, Q9UPT9 [30,31]; Ubp8p, P50102 [24,33]; USP27, A6NNY8; USP51, Q70EK9; USP44, Q9H0E7 [37]; USP49, Q70CQ1; USP33, Q8TEY7 [16,38]; USP20, Q9Y2K6; USP39/SNUT2, Q53GS9 [10]; Sad1p, P43589; BRAP2/IMP, Q7Z569 [11]; Yhl010cp, P38748; and HDAC6, Q9UBN7 [12,13,15,17,34,35].

USPs but devoid of protease activity and involved in spliceosome maturation [10]. Finally, this domain is also present in otherwise unrelated proteins, including the E3 ubiquitin ligase BRAP2 [BRCA1-associated protein 2, also known as IMP (impedes mitogenic signal propagation)] [11] and the cytoplasmic deacetylase (HDAC6) [12,13] (Figure 1).

During the past two years, the structures of three ZnF-UBPs from different USPs (USP5, USP16 and USP33) have been reported [14–16]. All domains are organized around a central five-stranded twisted  $\beta$ -sheet and a nearby  $\alpha$ -helix (Figure 2). Although this central region is highly conserved in all structures, a larger variability is observed in the other elements of secondary structure that surround it, possibly owing to different structural and/or functional specificities (Figure 2b).

The different ZnF-UBP domains coordinate, or are predicted to coordinate, one, two or three zinc ions [14,17]. For example, USP5 contains a single zinc ion, whereas in USP16, USP33 and HDAC6, three zinc ions are coordi-

nated by 12 cysteine or histidine residues that are arranged in an atypical cross-braced ring finger and encapsulated within a third zinc finger in the primary structure [15] (Figure 2c). In HDAC6, substitutions of each residue involved in zinc binding impair ubiquitin binding of this domain (see later) [17], underlying the importance of these zinc ions for ZnF-UBP structural stability. By contrast, the ZnF-UBPs in BRAP2 and its yeast orthologue Yhl010cp are predicted to contain two zinc ions. It is unclear why the number of bound zinc ions varies between domains, even in cases of known functional homologues. For instance, USP5, the human homologue of Ubp14p, can functionally replace its yeast counterpart in complementation assays but, surprisingly, the two proteins do not bind the same number of zinc ions (Ubp14 binds three, whereas USP5 binds only one) [18].

#### USP5 regulation by ubiquitin–ZnF-UBP binding

Initial biochemical studies demonstrated that murine HDAC6 ZnF-UBP binds ubiquitin [12,13]. In addition,

Seigneurin-Berny *et al.* [13] showed that USP5 (also called isopeptidase T), an enzyme that enables polyubiquitin-chain degradation following their release from conjugated proteins, is retained by ubiquitin-agarose [19]. The recent X-ray structure of the USP5 ZnF-UBP, free and in complex with ubiquitin, provides a clear mechanism for ubiquitin recognition by this domain [14]. The free ubiquitin C terminus, notably the C-terminal Arg-Gly-Gly sequence, penetrates a pocket formed by the conserved five-stranded  $\beta$ -sheet and a loop extending from it (Figure 2a). Major residues involved in ubiquitin C terminus recognition are Trp209, Arg221, Tyr223, Tyr259, Tyr261 and Asp264 (the numbering corresponds to the USP5 sequence); all but Tyr223 are located within, or in the vicinity of, the central  $\beta$ -sheet (Figure 2a). It should be noted that, unlike many domains that bind ubiquitin with a low affinity, the ZnF-UBP-ubiquitin affinity is high ( $K_D \sim 3 \mu\text{M}$  and 60 nM, for USP5 and HDAC6 ZnF-UBPs, respectively) [14,17].

The structural and functional studies of USP5 provided clues toward understanding its catalytic regulation. USP5 disassembles polyubiquitin chains that contain an intact and unblocked Arg-Gly-Gly sequence at the C terminus of the proximal subunit (defined as the end that contains the free C terminus) [19]. The USP5 catalytic activity relies on two ubiquitin-binding domains with different functions: the catalytic site and a distinct activation domain. Indeed, USP5 catalytic activity against the artificial substrate ubiquitin-amino-methylcoumarin increased after the addition of free ubiquitin which binds the activation domain [20,21]. The activation domain was later identified as the ZnF-UBP [14]. Indeed, Reyes-Turcu *et al.* [14] have demonstrated that binding of ubiquitin through its C-terminal Arg-Gly-Gly motif to the ZnF-UBP domain is sufficient to enhance USP5 catalytic activity *in vitro*. Moreover, substitutions of cysteine and histidine residues that coordinate the zinc ion (Figure 2c) interfere with ubiquitin-mediated USP5 activation. Similarly, alanine substitutions for most of the residues interacting with ubiquitin decrease (Asp264) or completely abolish (Trp209, Arg221 and Tyr261) USP5 activation by free ubiquitin [14]. Altogether, these results indicate a model in which the USP5 ZnF-UBP binds the free C terminus of proximal ubiquitin in polyubiquitin chains, a binding event that is required for full USP5 activation (Figure 3a).

### The binding of ZnF-UBPs to the unanchored ubiquitin C terminus

In agreement with the observations made for USP5, nuclear magnetic resonance studies have revealed similar interactions between the USP16 ZnF-UBP domain and the unanchored ubiquitin C-terminal Arg-Gly-Gly motif [15]. Different biochemical studies have reported an interaction between ubiquitin and ZnF-UBP domains from HDAC6, USP3 and Yhl010cp [14,17,22]. Together with the conservation between these proteins and USP5 of all residues involved in ubiquitin recognition, this result indicates that these ZnF-UBPs interact with ubiquitin through binding of the C-terminal Arg-Gly-Gly motif (Figure 2c). On the basis of these results, it could have been tempting to speculate that ubiquitin-binding is common to all ZnF-UBP domains.

However, recent biochemical and structural studies have revealed that the USP33 ZnF-UBP domain cannot bind ubiquitin, although the binding pocket is still formed and most of the ubiquitin-binding residues defined in USP5 are conserved in USP33 [16]. However, in USP33 Asp264 is replaced by a serine and, notably, Arg221 is replaced by a glutamate. The negative charge of this latter residue leads to the disappearance of an otherwise crucial interaction between free ubiquitin and the USP33 ZnF-UBP domain, and probably accounts for the loss of binding between these two molecules.

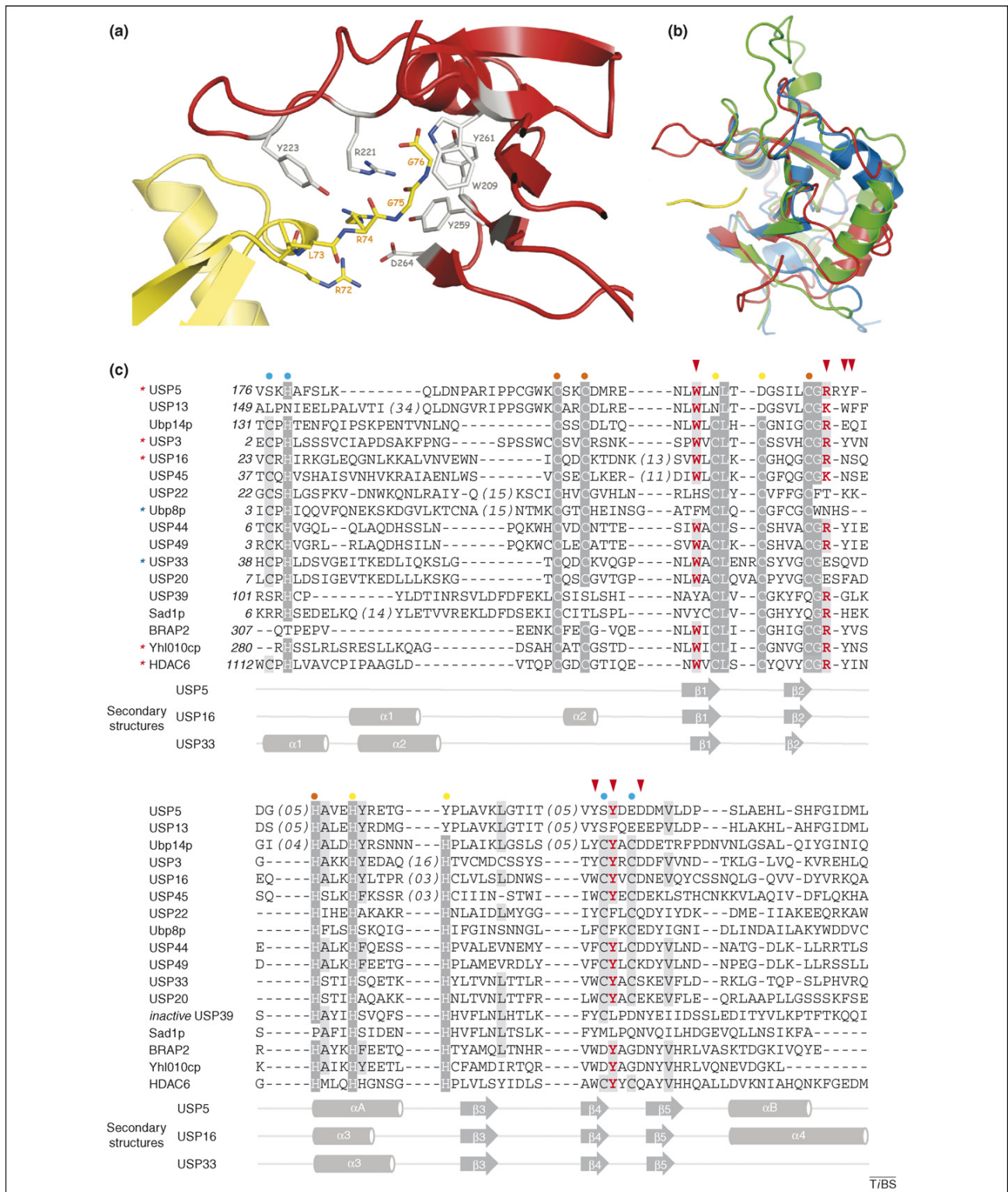
Consequently, to predict whether a ZnF-UBP can interact with ubiquitin, we have aligned all the known human and yeast ZnF-UBPs. Multiple alignments using ClustalX [23] have revealed that several uncharacterized human ZnF-UBPs (in USP44, USP49, USP45 and BRAP2) either share all the crucial residues with USP5, USP16, USP3, Yhl010cp and HDAC6 or harbour residues with similar properties for ubiquitin binding, as defined from structural studies. Thus, we predict that the USP44, USP49, USP45 and BRAP2 ZnF-UBPs interact with ubiquitin. By contrast, the crucial Arg221 residue present in USP5 but missing in USP33 (see earlier) [16] is also missing in three additional ZnF-UBPs, Ubp8p, USP22 and USP20, indicating that a subclass of these domains cannot bind ubiquitin. Moreover, other residues that come into contact with ubiquitin are poorly conserved within Ubp8p and USP22. In agreement, Ubp8p cannot be retained by immobilized ubiquitin [24], lending further credence to the idea that the Ubp8p and USP22 ZnF-UBP domains do not interact with free ubiquitin.

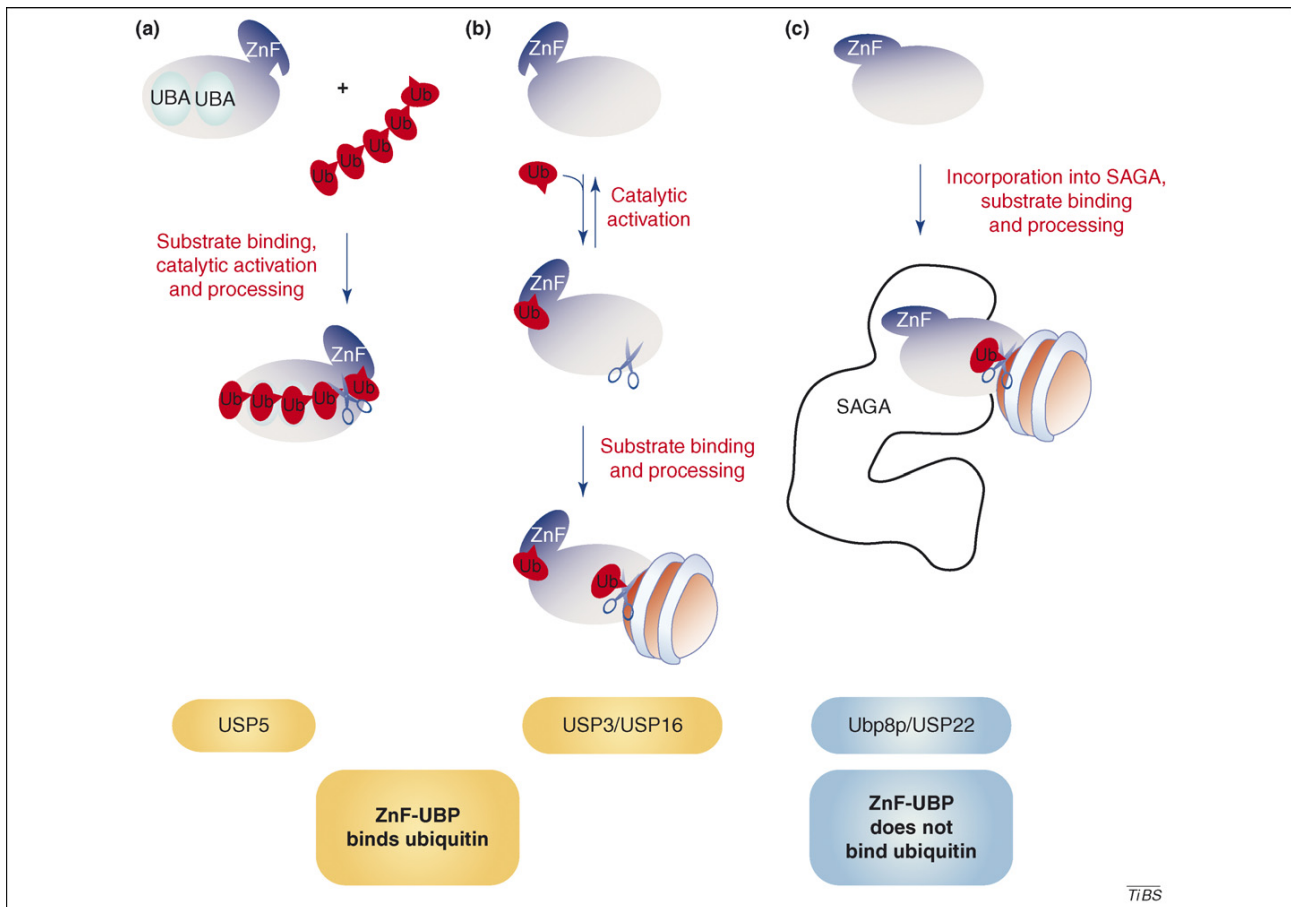
The ZnF-UBPs from USP13 and USP39 harbour another substitution among the residues involved in ubiquitin binding: Tyr261 is replaced with phenylalanine and leucine, respectively (Figure 2). The lack of the hydroxyl group brought by this Tyr261 probably affects ubiquitin binding. However, it remains to be determined experimentally whether these ZnF-UBP domains can bind ubiquitin.

It is thus possible, for the subgroup of ZnF-UBPs that bind ubiquitin, that enhancement of USP catalytic activity following free ubiquitin binding would be a general regulatory mechanism as demonstrated for USP5. Indeed, such a mechanism could apply to USP3 and USP16, two proteins that were recently shown to be involved in histone deubiquitylation.

### Histone deubiquitylation by ZnF-UBP-containing USPs

USP3, recently identified as a deubiquitylase enzyme for monoubiquitylated histones H2A (H2Aub1) and H2B (H2Bub1), is involved in S-phase progression and DNA repair [22]. Interestingly, the substitution of a histidine residue within the ZnF-UBP that coordinates zinc abrogated USP3 deubiquitylation activity to the same extent as a substitution of the cysteine in the catalytic triad. Although immunoprecipitation experiments indicate that the USP3 ZnF-UBP would be involved in substrate recognition, the fact that both USP3 and USP5 are retained by ubiquitin-agarose strongly indicates that USP3, like USP5, binds the ubiquitin free C-terminal Arg-Gly-Gly motif [22]. Thus, as has been demonstrated by different structural studies [25–27], it seems that the USP3 catalytic





**Figure 3.** Classification and proposed regulation of ZnF-UBP-containing proteins. Three different modes of USP regulation by the ZnF-UBP domain are schematized. Representative proteins are classified according to the ability of their ZnF-UBP domain to bind the C terminus of free ubiquitin. (a) USP5 participates in free polyubiquitin processing. The ZnF-UBP domain (ZnF; dark blue) binds the C terminus of free polyubiquitin chains (red) and therefore participates in substrate recognition together with the catalytic domain (light blue) and, probably, the two additional UBA domains. Binding of ubiquitin to ZnF-UBP triggers USP5 catalytic activation (represented by scissors). (b) The USP3 and USP16 ZnF-UBPs can also bind free ubiquitin; however, both enzymes process monoubiquitylated H2A. We propose that these ZnF-UBPs function as sensors for free ubiquitin. We speculate that ubiquitin binding regulates the catalytic activity of ZnF-UBP, as demonstrated for USP5. (c) By contrast, the Ubp8p and USP22 ZnF-UBPs do not interact with ubiquitin but, rather, mediate interactions between SAGA complex subunits. Incorporation of Ubp8p or USP22 into SAGA is required for its full histone deubiquitylation activity.

domain recognizes the ubiquitylated target and that the ZnF-UBP domain binds free ubiquitin, performing the same activating function as observed for USP5 (Figure 3).

USP16 (also termed Ubp-M) was also identified as a histone deubiquitylase; however, its activity is specific for H2Aub1 [28]. Through this function, USP16 regulates cell-cycle progression and gene expression. As indicated by sequence and structural similarities between the USP5, USP3 and USP16 ZnF-UBPs, we propose that these domains have a comparable positive regulatory activity within these proteins. Therefore, we suggest that the two H2Aub1 deubiquitylases, USP3 and USP16, are both regulated by their respective ZnF-UBPs, which could act as free ubiquitin pool sensors. However, the functional specificities of these enzymes might be directed by additional regulatory mechanisms; indeed, USP16 activity is affected by oligomerization and/or phosphorylation [28,29].

A third human ZnF-UBP-containing USP, USP22, was recently identified as a deubiquitylase for histones H2Aub1 and H2Bub1 [30,31]. USP22 is involved in cell-cycle progression, heterochromatin silencing and nuclear receptor coactivation. In contrast to the previously described ZnF-UBPs, no interaction could be detected between ubiquitin and the Ubp8p ZnF-UBP (Ubp8p is the yeast USP22 orthologue) [24]. Although it has not yet been experimentally demonstrated, protein sequence analysis strongly indicates that the USP22 ZnF-UBP does not bind free ubiquitin (see earlier and Figure 2c). Furthermore, the ZnF-UBP domains of either Ubp8p or USP22, which are components of the Spt-Ada-Gcn5-acetyltransferase (SAGA) complexes, are required for interactions with other subunits of the corresponding SAGA complexes [24,31,32]. Interestingly, neither free recombinant USP22 nor Ubp8p can efficiently deubiquitylate H2Bub1 or H2Aub1 (in the case of USP22) *in vitro*, indicating that

binding in other sequences are highlighted in red. ZnF-UBP domains that can or cannot bind ubiquitin, as demonstrated experimentally, are indicated by a red or a blue star, respectively. Secondary structures of ZnF-UBP from USP5, USP16 and USP33 are drawn under the alignment. Conserved histidine and cysteine residues are indicated with blue, red or yellow circles according to the coordinated zinc ion. Sequence accession numbers are as in Figure 1.

an interaction(s) between the ZnF-UBP and other SAGA subunits is required for the full activity of USP22 and Ubp8p [30,31,33] (Figure 3c).

#### Functional classification of ZnF-UBP domains

We propose that the ZnF-UBP domain is one of several means by which to regulate USP activity. In addition, our classification of ZnF-UBPs that can or cannot bind free ubiquitin has enabled us to identify two main types of regulation by these domains. First, we propose that, for the ZnF-UBP domains that can bind ubiquitin, the interaction with the ubiquitin C-terminal Arg-Gly-Gly motif enhances the catalytic activity of the corresponding USPs. This activation has been clearly demonstrated for USP5, a protein involved in free polyubiquitin processing. In this situation, the substrate contains a free C-terminal Arg-Gly-Gly motif that is recognized by the ZnF-UBP domain. Therefore, recognition of free polyubiquitin chains by USP5 is mediated by both its catalytic and ZnF-UBP domains. Thus, USP5 activation is mediated by binding of the ubiquitin Arg-Gly-Gly motif to its ZnF-UBP (Figure 3a). Other ZnF-UBPs that can bind free ubiquitin participate in the processing of different substrates that do not contain the free Arg-Gly-Gly motif. For these enzymes we propose, based on structural and functional data, a model in which the ZnF-UBP domains act as free ubiquitin sensors and bind the ubiquitin C-terminal Arg-Gly-Gly motif, leading to enzymatic activation (Figure 3b). This model could apply to two recently identified histone DUBs, USP3 and USP16, because their DUB activity could be regulated by the pool of free nuclear ubiquitin, the levels of which undergo variations throughout the cell cycle.

It has been proposed that HDAC6, a cytoplasmic deacetylase, could be regulated by the binding of free C-terminal ubiquitin to its ZnF-UBP domain [15]. In addition, many studies have reported that HDAC6 binds polyubiquitylated proteins in a ZnF-UBP-dependent manner, thereby transporting misfolded ubiquitylated proteins to aggresomes ([34] and references therein). Therefore, HDAC6 ZnF-UBP undergoes two types of ubiquitin-binding depending on the accessibility of a free Arg-Gly-Gly end. According to this model, the strong binding of free ubiquitin to HDAC6 ZnF-UBP would impede interaction of HDAC6 with polyubiquitylated proteins, thereby regulating the turnover of misfolded proteins [15,34,35]. BRAP2, a RING (really interesting new gene) E3 ligase, might also be regulated by free ubiquitin levels [11,14]. Indeed, free ubiquitin sensing could be a more general role of each ubiquitin-binding USP.

Surprisingly, another novel class of ZnF-UBP domains cannot be regulated by binding free ubiquitin. Indeed, the Ubp8p and USP22 ZnF-UBPs cannot bind ubiquitin, but are necessary for the incorporation of these enzymes in the multi-subunit SAGA complexes [24,31,32] (Figure 3c). Interestingly, SAGA complex histone- deubiquitylation activity is mediated by Ubp8p or USP22 in the corresponding yeast or human complexes, whereas both recombinant enzymes are inactive in isolation [31,33]. Therefore, the Ubp8p and USP22 ZnF-UBP domains that mediate interactions with other SAGA complex subunits could regulate the activity of these enzymes (Figure 3c). Such regulation

could be explained by conformational changes induced by these interactions or by the dependence upon other subunits for the recruitment of SAGA complexes to nucleosomes. The precise mechanisms leading to Ubp8p and USP22 regulation by the ZnF-UBP domain in the SAGA complexes remain to be elucidated.

#### Concluding remarks

Increasing evidence indicates that the fine tuning of ubiquitylation is controlled by both the ubiquitin conjugation and deconjugation systems. This model shows a tight regulation of DUBs, which use a variety of mechanisms to regulate their activities. ZnF-UBP domains are found in a subset of USPs and provide a paradigm for the regulation of DUB activity by specific domains. Recent biochemical and structural data differentiate two types of ZnF-UBPs according to their ubiquitin-binding ability. The first subclass of ZnF-UBP domains do not bind ubiquitin directly but, rather, regulate the activity of the corresponding enzyme by promoting interaction with other proteins (Figure 3c). By contrast, the second subclass of ZnF-UBPs bind the free ubiquitin C terminus; this binding leads to USP catalytic activation as demonstrated for USP5 (Figure 3a,b). In addition, we propose that the latter subclass of ZnF-UBPs act as sensors of the free ubiquitin pool.

Despite the recent progress toward understanding the role of ZnF-UBP in USP activation, several questions remain unanswered. More functional and structural studies are required to identify the catalytic domain conformational changes that are mediated by the ZnF-UBP upon either ubiquitin binding or interaction with other partners. Although the DUB activity of USP5 is increased by binding of ubiquitin to USP5 ZnF-UBP, it remains to be determined whether the binding of ubiquitin to other ZnF-UBP domains would activate or inhibit the activity of the corresponding enzymes. Lastly, because the ubiquitin C-terminal Arg-Gly-Gly motif is conserved in other ubiquitin-like proteins (e.g. NEDD8 and ISG15) we propose that such proteins would bind to ZnF-UBP domains, thereby regulating USP activity.

#### Acknowledgements

Research in the authors' laboratory is supported by grants from the Fondation pour la Recherche Médicale (FRM) and the Agence Nationale de la Recherche (ANR-05-MRAR-029-01) to D.D. and from ANR (05-BLAN-0396-01) and European Community (HPRN-CT 00504228) to L.T.

#### References

- 1 Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* 67, 425–479
- 2 Pickart, C.M. and Eddins, M.J. (2004) Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* 1695, 55–72
- 3 Hicke, L. (2001) Protein regulation by monoubiquitin. *Nat. Rev. Mol. Cell Biol.* 2, 195–201
- 4 Hicke, L. *et al.* (2005) Ubiquitin-binding domains. *Nat. Rev. Mol. Cell Biol.* 6, 610–621
- 5 Kerscher, O. *et al.* (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu. Rev. Cell Dev. Biol.* 22, 159–180
- 6 Weake, V.M. and Workman, J.L. (2008) Histone ubiquitination: triggering gene activity. *Mol. Cell* 29, 653–663
- 7 Zhang, Y. (2003) Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev.* 17, 2733–2740
- 8 Nijman, S.M. *et al.* (2005) A genomic and functional inventory of deubiquitinating enzymes. *Cell* 123, 773–786

- 9 Hurley, J.H. *et al.* (2006) Ubiquitin-binding domains. *Biochem. J.* 399, 361–372
- 10 Makarova, O.V. *et al.* (2001) The 65 and 110 kDa SR-related proteins of the U4/U6.U5 tri-snRNP are essential for the assembly of mature spliceosomes. *EMBO J.* 20, 2553–2563
- 11 Matheny, S.A. *et al.* (2004) Ras regulates assembly of mitogenic signalling complexes through the effector protein IMP. *Nature* 427, 256–260
- 12 Hook, S.S. *et al.* (2002) Histone deacetylase 6 binds polyubiquitin through its zinc finger (PAZ domain) and copurifies with deubiquitinating enzymes. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13425–13430
- 13 Seigneurin-Berny, D. *et al.* (2001) Identification of components of the murine histone deacetylase 6 complex: link between acetylation and ubiquitination signaling pathways. *Mol. Cell. Biol.* 21, 8035–8044
- 14 Reyes-Turcu, F.E. *et al.* (2006) The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* 124, 1197–1208
- 15 Pai, M.T. *et al.* (2007) Solution structure of the Ubp-M BUZ domain, a highly specific protein module that recognizes the C-terminal tail of free ubiquitin. *J. Mol. Biol.* 370, 290–302
- 16 Allen, M.D. and Bycroft, M. (2007) The solution structure of the ZnF UBP domain of USP33/VDU1. *Protein Sci.* 16, 2072–2075
- 17 Boyault, C. *et al.* (2006) HDAC6-p97/VCP controlled polyubiquitin chain turnover. *EMBO J.* 25, 3357–3366
- 18 Amerik, A. *et al.* (1997) *In vivo* disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J.* 16, 4826–4838
- 19 Wilkinson, K.D. *et al.* (1995) Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. *Biochemistry* 34, 14535–14546
- 20 Dang, L.C. *et al.* (1998) Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amido-4-methylcoumarin by deubiquitinating enzymes. *Biochemistry* 37, 1868–1879
- 21 Stein, R.L. *et al.* (1995) Kinetic studies of isopeptidase T: modulation of peptidase activity by ubiquitin. *Biochemistry* 34, 12616–12623
- 22 Nicassio, F. *et al.* (2007) Human USP3 is a chromatin modifier required for S phase progression and genome stability. *Curr. Biol.* 17, 1972–1977
- 23 Thompson, J.D. *et al.* (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882
- 24 Ingvarsdottir, K. *et al.* (2005) H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol. Cell. Biol.* 25, 1162–1172
- 25 Hu, M. *et al.* (2002) Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* 111, 1041–1054
- 26 Hu, M. *et al.* (2005) Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J.* 24, 3747–3756
- 27 Renatus, M. *et al.* (2006) Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. *Structure* 14, 1293–1302
- 28 Joo, H.Y. *et al.* (2007) Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* 449, 1068–1072
- 29 Cai, S.Y. *et al.* (1999) A mutant deubiquitinating enzyme (Ubp-M) associates with mitotic chromosomes and blocks cell division. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2828–2833
- 30 Zhang, X.Y. *et al.* (2008) The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol. Cell* 29, 102–111
- 31 Zhao, Y. *et al.* (2008) A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol. Cell* 29, 92–101
- 32 Kohler, A. *et al.* (2006) The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Mol. Biol. Cell* 17, 4228–4236
- 33 Lee, K.K. *et al.* (2005) The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol. Cell. Biol.* 25, 1173–1182
- 34 Boyault, C. *et al.* (2007) HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates. *Genes Dev.* 21, 2172–2181
- 35 Kawaguchi, Y. *et al.* (2003) The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. *Cell* 115, 727–738
- 36 Galtier, N. *et al.* (1996) SEAVIEW and PHYLO\_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* 12, 543–548
- 37 Stegmeier, F. *et al.* (2007) Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature* 446, 876–881
- 38 Curcio-Morelli, C. *et al.* (2003) Deubiquitination of type 2 iodothyronine deiodinase by von Hippel-Lindau protein-interacting deubiquitinating enzymes regulates thyroid hormone activation. *J. Clin. Invest.* 112, 189–196



# The Tightly Controlled Deubiquitination Activity of the Human SAGA Complex Differentially Modifies Distinct Gene Regulatory Elements<sup>∇</sup>

Guillaume Lang,<sup>#</sup> Jacques Bonnet,<sup>#</sup> David Umlauf, Krishanpal Karmodiya, Jennifer Koffler, Matthieu Stierle, Didier Devys,<sup>\*</sup> and László Tora<sup>\*</sup>

*Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 964, Université de Strasbourg, Illkirch, France*

Received 17 February 2011/Returned for modification 29 March 2011/Accepted 30 June 2011

**The multisubunit SAGA coactivator complex facilitates access of general transcription factors to DNA through histone acetylation mediated by GCN5. USP22 (ubiquitin-specific protease 22) was recently described as a subunit of the human SAGA complex that removes ubiquitin from monoubiquitinated histone H2B and H2A *in vitro*. Here we demonstrate an allosteric regulation of USP22 through multiple interactions with different domains of other subunits of the SAGA deubiquitination module (ATXN7, ATXN7L3, and ENY2). Downregulation of ATXN7L3 by short hairpin RNA (shRNA) specifically inactivated the SAGA deubiquitination activity, leading to a strong increase of global H2B ubiquitination and a moderate increase of H2A ubiquitination. Thus, SAGA is the major H2Bub deubiquitinase in human cells, and this activity cannot be fully compensated by other deubiquitinases. Here we show that the deubiquitination activity of SAGA is required for full activation of SAGA-dependent inducible genes. Interestingly, the reduction of the SAGA deubiquitination activity and the parallel increase in H2B ubiquitination at inducible target genes before activation do not induce aberrant gene expression. Our data together indicate that different dynamic equilibria of H2B ubiquitination/deubiquitination are established at different gene regulatory elements and that H2B ubiquitination changes are necessary but not sufficient to trigger parallel activation of gene expression.**

Transcription initiation by RNA polymerase II (Pol II) is a highly regulated process that requires the coordinated action of numerous factors. Coactivators are multisubunit complexes that are recruited to promoters by gene-specific activators to facilitate transcription initiation either by direct interaction with general transcription factors and Pol II or indirectly through modifications of the chromatin structure (reviewed in references 18 and 19).

SAGA (*Spt-Ada-Gcn5-Acetyltransferase*) is a coactivator complex that is functionally and structurally conserved from yeast to mammals and contains two enzymatic activities (14, 32). Histone acetylation is mediated by the catalytic activity of the GCN5 subunit and is thought to facilitate the recruitment of the general transcription machinery (19). Yeast SAGA contains the ubiquitin (Ub)-specific protease (USP) yUbp8, which was shown to remove ubiquitin from monoubiquitinated histone H2B (2, 6). yUbp8 strongly interacts with ySgf73, ySgf11, and ySus1 to form a functional deubiquitination module (DUBm) that is anchored within SAGA through ySgf73 (7, 10, 11, 17, 25). More recently, a similar enzymatic activity carried out by hUSP22 has been described in the human SAGA (hSAGA) complex deubiquitinating monoubiquitinated histones H2A (H2Aub) and H2B (H2Bub) *in vitro* (35–37). Similarly to the yeast complex, hUSP22 was shown to be a component of the DUBm that contains hATXN7, hATXN7L3, and hENY2, the orthologues of ySgf73, ySgf11, and ySus1, respec-

tively. Two ATXN7 paralogues, ATXN7L1 and ATXN7L2, are also components of the hSAGA complex, and due to their strong sequence similarity with ATXN7 it is expected that these proteins have redundant functions with ATXN7 (4, 15). Interestingly, an expansion of a polyglutamine motif found in the N-terminal region of ATXN7 but missing from the other paralogues is responsible for an inherited neurodegenerative disorder called spinocerebellar ataxia type 7 (SCA7) (reviewed in reference 5).

yUbp8 was shown to be inactive unless it is associated with the other proteins of the yeast DUBm. Recent determination of the yeast DUBm structure revealed how this module assembles and activates yUbp8 (12, 29). Similarly to the yeast protein, hUSP22 alone has almost no activity, and we previously hypothesized that hUSP22 has to be incorporated into SAGA to be active (37). hUSP22 and yUbp8 belong to a subclass of USPs that contain a zinc finger (ZnF) domain called ZnF-UBP in their N-terminal region (13, 24). This domain was first characterized as an interactor of free ubiquitin and has been involved in the catalytic activation of USP5, an enzyme which degrades free polyubiquitin chains (27). However, it has been shown that the Ubp8 ZnF-UBP is not able to bind free ubiquitin (7) and sequence conservation analysis revealed that this ubiquitin-binding property is likely absent from USP22 (1). In line with these studies, a mechanism for USP22 catalytic activation that involves the ZnF-UBP domain and its interactions with the other subunits of the DUBm has been proposed (1).

In mammals, monoubiquitination of H2B is catalyzed by the E2 conjugating enzymes HR6A/HR6B and the E3 ligase RNF20/RNF40 (8, 9, 38). The H2B ubiquitination machinery is recruited at the promoter through interactions with activators and is then activated through interactions with additional

<sup>\*</sup> Corresponding author. Mailing address: IGBMC, 1 rue Laurent Fries, BP10142, 67404 Illkirch Cedex, France. Phone: 33 388 65 34 44. Fax: 33 388 65 32 01. E-mail for László Tora: laszlo@igbmc.fr. E-mail for Didier Devys: devys@igbmc.fr.

<sup>#</sup> These authors equally contributed to this manuscript.

<sup>∇</sup> Published ahead of print on 11 July 2011.

factors such as the PAF complex (34; reviewed in reference 33). However, the precise function of this mark in transcription regulation remains elusive. In mammalian cells, H2Bub was found genome-wide to associate preferentially with transcribed regions of highly expressed genes, suggesting a positive role in transcription (20). In human cells, the SAGA deubiquitination activity is required for full transcriptional activity mediated by nuclear receptors (37). In yeast, increased levels of H2Bub were detected on the GAL1 core promoter and throughout the transcribed region upon transcriptional activation, with both ubiquitination and deubiquitination being required for optimal transcription (6). Hence, H2B monoubiquitination seems to be a highly dynamic process, although its precise functions in transcription regulation remain unclear.

In this study, we show that interactions of USP22 with other proteins of the DUBm are required for its incorporation in the human SAGA complex and for its activation. We also demonstrate that *in vivo*, the SAGA deubiquitination activity is directed against H2Bub and to a lesser extent H2Aub. This activity is required to fully activate SAGA-regulated inducible genes, and our results suggest that this occurs via distinct ubiquitination/deubiquitination cycles at the promoters and the transcribed regions of these genes.

#### MATERIALS AND METHODS

**Plasmids.** Plasmids that were used to generate the baculoviruses expressing Fl-USP22, Fl-ATXN7L3, and ENY2 were previously described (37). A full-length USP22 cDNA was PCR amplified using a 5' primer encoding a hemagglutinin (HA) epitope and cloned into pVL1393 baculovirus transfer vector (BD Biosciences). cDNAs corresponding to ATXN7L3 fragments from amino acid 3 to amino acid 151 [ATXN7L3(3-151)], ATXN7L3(3-76), and ATXN7L3(3-55) were PCR amplified and cloned into a derivative of the pBacPAK8 baculovirus transfer vector (Clontech) containing a sequence encoding a Flag epitope upstream of the cloning site. cDNA corresponding to ATXN7(75-172) and ATXN7(75-142) was amplified by PCR and cloned into a derivative of the pFBDOCS (a kind gift from Imre Berger) containing a sequence encoding the HA and Flag epitopes upstream of the cloning site. Production of recombinant viruses, Sf9 cell infection, and preparation of whole-cell extracts were performed as described previously (37).

**Antibodies.** The anti-ATXN7L3 (2997) and anti-SPT20 (3006) polyclonal antibodies were obtained by immunization of rabbits with protein fragments corresponding to residues 197 to 276 of human ATXN7L3 and 330 to 531 of mouse SPT20. Antibodies against the following proteins have been described earlier: GCN5 (5GC 2A6), SGF29 (2461), ATXN7L3 (2325), USP22 (2391), ENY2 (2585), and H2B (LG11-2) (15, 22, 37, 21). Other antibodies used in this study are as follows: FLAG (M2 [Sigma]), HA (12CA5), AR (PG21 [Upstate]), RNA Pol II (7G5), H2Bub (NR03 [Medimabs]), H2A (07-146 [Upstate]), H2Aub (E6C5 [Millipore]), and Ub (FK2 [Biomol International]).

**Cell culture.** HeLa cells were grown in Dulbecco's modified medium supplemented with 1 g/liter glucose, 5% fetal calf serum, and gentamicin. Stress treatments were carried out using 100 nM thapsigargin (Sigma) or dimethyl sulfoxide (DMSO) for 3 h or 4 h. LNCaP cells were grown in RPMI medium supplemented with 10% fetal calf serum, 10 mM HEPES, 1 mM sodium pyruvate, and gentamicin. Before induction, cells were grown in charcoal-treated medium and then treated for 210 min with  $10^{-5}$  M R1881 or with ethanol (EtOH).

**shRNA cell line constructions.** HEK293T were cotransfected with pLVTH (lenticular vector containing the short hairpin RNA [shRNA] sequence), pCMVdR8.2, and pMD2G-VSVG plasmids with JetPEI (polyplus transfection). Forty-eight hours later, medium containing lentiviruses was collected and used to infect HeLa or LNCaP cells. Positive cells were sorted by fluorescence-activated cell sorting (FACS). Sequences were as follows: for shATXN7L3, 5'-GGAAGCTTAGCAATTCGGATCCTTCAAGAGAGGATCCGAATTGCTAAGTCC-3'; and for shControl, 5'-CTTACGCTGAGTACTTCGATTCGAAGAGATCGAAGTACTCAGCGTAAAG-3'.

**Complex purifications.** Whole-cell extracts from Sf9 cells were immunoprecipitated using anti-HA antibodies coupled to protein A-agarose beads (EZview Red anti-HA affinity gel [Sigma]) for 3 h at 4°C. Antibody-protein A-bound

protein complexes were washed two times with IP buffer (25 mM Tris-HCl, pH 7.9; 10% glycerol; 0.1% NP-40; 1 mM dithiothreitol [DTT]; 5 mM MgCl<sub>2</sub>) containing 0.5 M KCl and twice with IP buffer containing 100 mM KCl. The immunoprecipitated material was then eluted by peptide competition overnight at 4°C in 100 mM KCl IP buffer. Eluted complexes were analyzed by Coomassie blue staining or Western blotting using the indicated primary antibodies and chemiluminescence detection was performed according to manufacturer's instructions (GE-Healthcare). Whole-cell extracts from HeLa stable cell lines were immunoprecipitated using monoclonal anti-GCN5 (5GC 2A6) and protein G-Sepharose (GE-Healthcare).

***In vitro* deubiquitination assay.** Histones were prepared from HeLa cells by lysing the cells in 10 mM HEPES, pH 7.9; 1.5 mM MgCl<sub>2</sub>; 10 mM KCl; 0.5 mM DTT; 1.5 mM phenylmethylsulfonyl fluoride (PMSF); 10 mM *N*-ethylmaleimide supplemented with 0.2 M HCl. Samples were incubated on ice for 30 min and centrifuged, and the supernatant was dialyzed first against 0.1 M acetic acid and then against water. Nucleosomes were purified from HeLa cells as described previously (30). Free histones or mononucleosomes were incubated with SAGA complexes immunopurified using an anti-USP22 antibody or purified recombinant USP22 in 100 mM Tris-HCl, pH 8.0; 5% glycerol; 1 mM EDTA; 3 mM DTT for 2 h at 37°C and analyzed by Western blotting using specific antibodies. For *in vitro* assays using ubiquitin vinyl sulfone (Ub-VS [BostonBiochem]), the purified complexes were incubated with 5 μM Ub-VS in a reaction buffer (100 mM Tris-HCl, pH 8.0; 5% glycerol; 100 mM KCl; 3 mM DTT) for 15 min at room temperature (RT). Alternatively, purified complexes were added on mononucleosomes in the presence or absence of 5 μM Ub-VS in a reaction buffer (20 mM Tris-HCl, pH 8.0; 100 mM KCl; 4 mM EDTA; 8 mM MgCl<sub>2</sub>; 3 mM DTT) for 2 h at 37°C. The reactions were stopped by the addition of Laemmli blue and then the reaction products were analyzed by Western blotting.

**RNA isolation, reverse transcription, quantitative PCR (qPCR).** Total RNA was isolated using the TRIzol reagent (Invitrogen). Reverse transcription was performed using SuperScript II (Invitrogen) and random hexamers according to the manufacturer's instructions. cDNAs were quantified by real-time PCR using SYBR green PCR master mix (Qiagen or Roche) on a LightCycler 480 instrument (Roche). Primers used are shown at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011).

**Chromatin immunoprecipitation (ChIP).** Cells were cross-linked with 1% formaldehyde at RT for 10 min, resuspended in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8.0]), and protease inhibitor cocktail [Roche]), and sonicated using a Bioruptor apparatus (Diagenode) until an average DNA fragment size of 200 to 500 bp was achieved. Supernatants were diluted in dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, and 20 mM Tris-HCl, pH 8.0) followed by pre-clearing with sheared salmon sperm DNA, bovine serum albumin (BSA), and protein G-Sepharose (Sigma) (for monoclonal mouse antibodies) or protein A-Sepharose (for polyclonal rabbit antibodies). The pre-cleared chromatin samples were shaken overnight at 4°C with the antibody, and then beads were added for 4 h to the samples to pull down specific protein-DNA complexes. Antibodies used are as follows: RNA Pol II (7G5), H2Bub (NR03 [Medimabs]), H2B (LG11-2), AR (PG21 [Upstate]), anti-ATXN7L3 (2997), and anti-SPT20 (3006). The following washes were carried out at 4°C: twice with TSEI (0.1% SDS; 1% Triton X-100; 2 mM EDTA; 150 mM NaCl; 20 mM Tris-HCl, pH 8.0; and protease inhibitor cocktail), twice with buffer TSEII (0.1% SDS; 1% Triton X-100; 2 mM EDTA; 500 mM NaCl; 20 mM Tris-HCl, pH 8.0; and protease inhibitor cocktail), twice with buffer III (250 mM LiCl; 1% NP-40; 1% Na-deoxycholate; 1 mM EDTA; 10 mM Tris, pH 8.0; and protease inhibitor cocktail), and finally twice with Tris-EDTA buffer (10 mM Tris, pH 8.0; 1 mM EDTA). Bound fractions of the chromatin were eluted with 200 μl of elution buffer (50 mM Tris, pH 8.0; 1 mM EDTA; 1% SDS) at 65°C and treated with RNase A (5 μg/ml), cross-link was reversed in 200 mM NaCl at 65°C overnight, and immunopurified material was incubated with proteinase K. DNA was phenol-chloroform extracted, precipitated by ethanol, and then used for qPCR analysis using a Roche LightCycler 480 with SYBR green master mix (Qiagen or Roche). Primers used in the ChIP experiments are available at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011).

#### RESULTS

**USP22 enzymatic activity requires its incorporation within the human SAGA complex.** Preliminary observations suggested that the free recombinant USP22 has a very weak deubiquitination activity on recombinant histones (36, 37). To test whether free USP22 would have a stronger activity on nucleosomal templates, we directly compared the deubiquitination

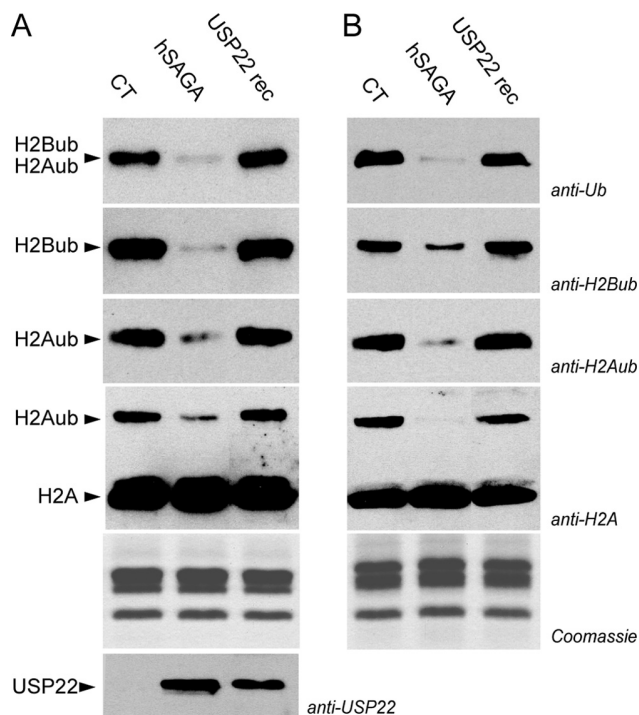


FIG. 1. H2Aub or H2Bub are deubiquitinated by the purified human SAGA complex but not by the free USP22 enzyme. Mononucleosomes (A) or free histones (B) extracted from HeLa cells and enriched in H2Aub and H2Bub were incubated with no enzyme (CT) or with purified SAGA complex (hSAGA) or recombinant USP22 (USP22 rec). Specific purification of hSAGA complex was verified by mass spectrometry analysis (see Table S1 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). The deubiquitination reaction was analyzed by Western blotting using antibodies directed against ubiquitin, H2Bub, H2Aub, H2A, and USP22 as indicated. The Coomassie blue staining shows that similar amounts of mononucleosomes or free histones were used.

activity of the recombinant USP22 with that of the endogenous hSAGA complex after normalization for the amount of USP22 (Fig. 1A, bottom panel) on either free histones (Fig. 1A) or nucleosomes (Fig. 1B) purified from HeLa cells. Although the purified SAGA complex efficiently removed the ubiquitin moiety from H2Aub and H2Bub on histone and nucleosomal substrates, the recombinant USP22 had no detectable activity on the same substrates (Fig. 1). These observations demonstrate that SAGA deubiquitinates H2Aub and H2Bub also in a nucleosomal context and that USP22 has to interact with other SAGA subunits to be fully active.

**The human SAGA subunits, ATXN7, ATXN7L3, ENY2, and USP22, assemble in a stable deubiquitination module.** To analyze how USP22 gets activated, we first analyzed how the DUBm is assembled. To this end, we coexpressed in Sf9 cells various combinations of the different subunits of the DUBm and immunopurified the corresponding subcomplexes (Fig. 2). In order to study the catalytic regulation of USP22, we analyzed the assembly of the DUBm using a full-length USP22 construct. Furthermore, interaction experiments using truncated versions of USP22 showed that neither the ZnF-UBP domain nor the catalytic domain alone could assemble into a stable complex with ATXN7L3 and ENY2 (see Fig. S1 at

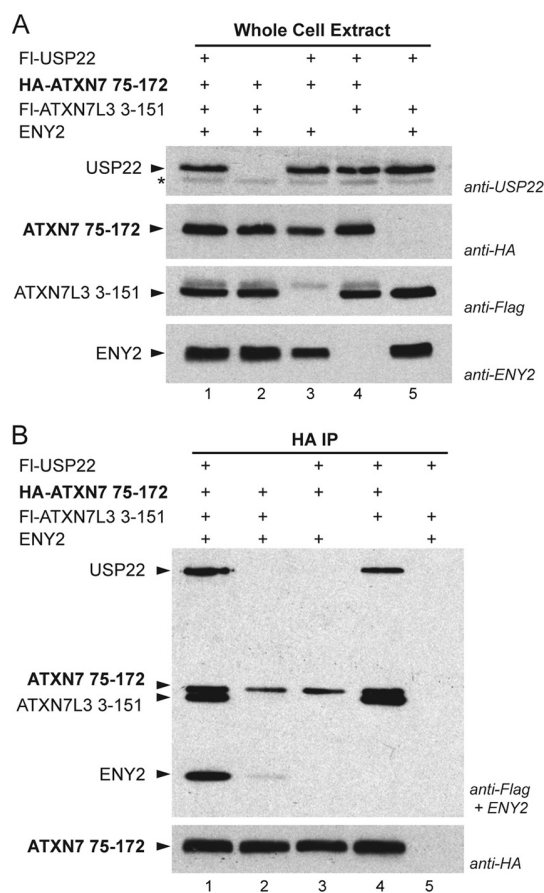


FIG. 2. The DUB module is stabilized by a network of interactions between the different subunits. (A) Sf9 cells were infected with different combinations of baculoviruses expressing Flag-USP22, HA-ATXN7(75–172), Flag-ATXN7L3(3–151), and ENY2 as indicated. Whole-cell extracts were analyzed by Western blotting and revealed with the indicated antibodies. (B) Immunoprecipitations using an anti-HA antibody followed by peptide elution were analyzed by Western blotting revealed with the indicated antibodies. ATXN7(75–172) (in bold) is harboring the HA epitope. The asterisk indicates a nonspecific band.

[http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)) (37). We also used a full-length ENY2 construct as this small protein of 101 amino acids is expected from yeast studies to bind a long  $\alpha$  helix in the N-terminal region of ATXN7L3 (3). Similarly, we expressed an N-terminal fragment of ATXN7L3 [ATXN7L3(3–151)] that contains the sequence that is homologous to Sgf11, the yeast orthologue of ATXN7L3.

A stable recombinant DUBm could be purified when this ATXN7L3 fragment and the full-length ENY2 and USP22 were coexpressed together with an N-terminal fragment of ATXN7 [ATXN7(75–172)] (Fig. 2B, lane 1). ENY2 appeared dispensable for the DUBm assembly as a complex composed of ATXN7, ATXN7L3, and USP22 could be purified (Fig. 2B, lane 4), even though it appeared slightly less stable than the complex formed with the four proteins. Similarly a stable USP22-ATXN7L3-ENY2 complex was purified in the absence of ATXN7 (Fig. 3B). By contrast, no stable complex could be formed in the absence of either USP22 or ATXN7L3 (Fig. 2B, lanes 2 and 3; also see Fig. S2 at <http://igbmc.fr/Lang>

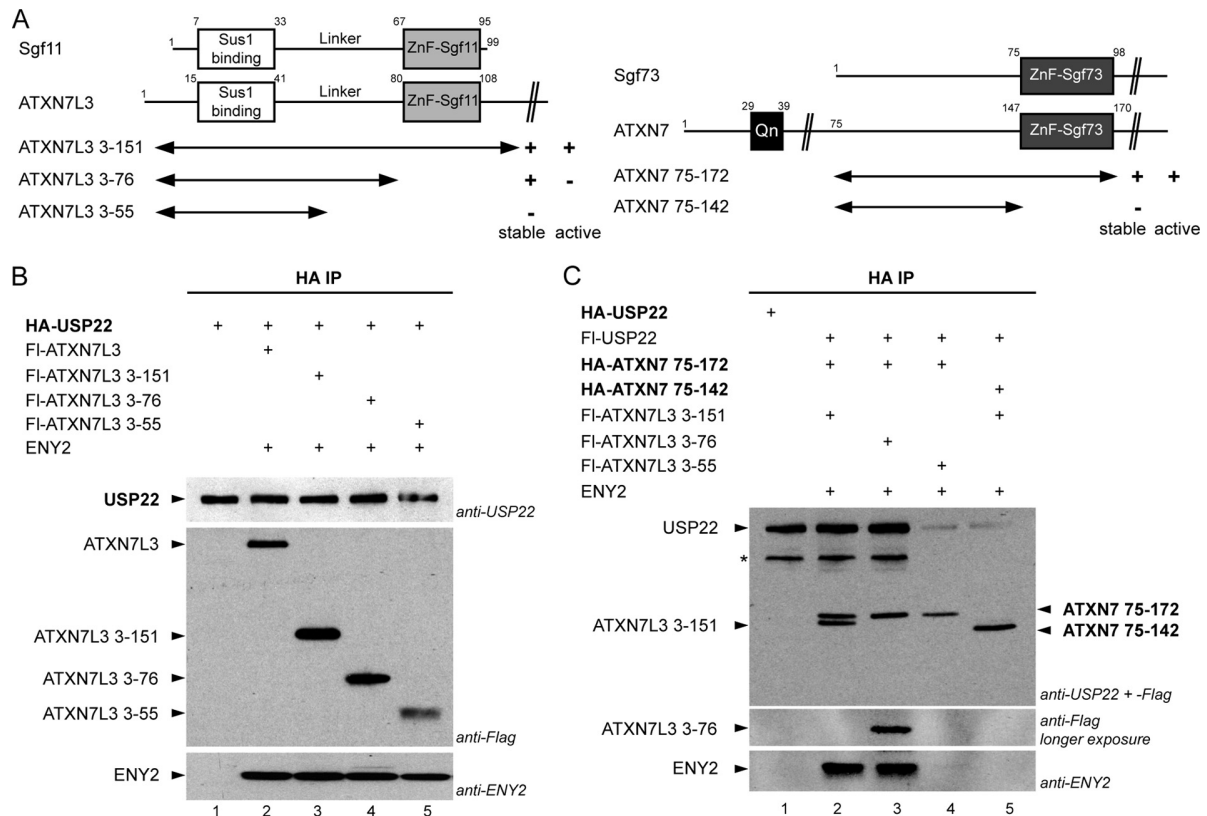


FIG. 3. Identification of the domains required for the formation of a stable DUB module. (A) Schematic representation of the domains of ATXN7L3 and ATXN7 involved in the formation of the DUBm. The functional regions of the human proteins are indicated as compared to their yeast orthologues Sgf11 and Sgf73, respectively (also see Fig. S3 and S4 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). The different truncated mutants of ATXN7L3 and ATXN7 that were used in immunopurification experiments are shown and their ability to form a stable/active DUBm composed of the four proteins is summarized. (B and C) Sf9 cells were infected with several combinations of baculoviruses expressing different versions of USP22, ATXN7, ATXN7L3, and ENY2 as indicated. Immunoprecipitations using an anti-HA antibody followed by peptide elution were analyzed by Western blotting revealed with the indicated antibodies. Coomassie blue staining of the purified complexes are shown in Fig. S5 at the URL given above. The subunits harboring an HA epitope (that were used to pull down the complexes) are indicated in bold.

\_mcb2011). Taken together, these results show that the DUBm is assembled through multiple interactions, rather than pairwise interactions, occurring between the four different subunits of the module.

**Determination of the minimal domains required to assemble the deubiquitination module of the human SAGA complex.** We next asked which domains of the subunits forming the DUBm are involved in the interaction networks that drive the assembly of this module. Both ATXN7 and ATXN7L3 incorporated within the DUBm through their N-terminal regions, which are composed of their respective ZnF domains (hereafter called ZnF-Sgf73 and ZnF-Sgf11) and poorly conserved N-terminal extensions (Fig. 3A; also see Fig. S3 and S4 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). Based on this common domain organization of the two proteins, we hypothesized that the ZnF of ATXN7 and ATXN7L3 plays a pivotal role in the assembly and/or the enzymatic regulation of the DUBm.

To test this hypothesis and to investigate the minimal domains of ATXN7L3 that are required to incorporate into the DUBm, we expressed different truncation mutants of ATXN7L3. A mutant containing the whole N-terminal region [ATXN7L3(3–151)] and two mutants deleted for the

ZnF domain [ATXN7L3(3–76) and ATXN7L3(3–55)] interacted with USP22 and ENY2 as well as the full-length ATXN7L3 (Fig. 3B, lanes 2 to 5; also see Fig. S2 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). In contrast with the smallest mutant [ATXN7L3(3–55)], the ATXN7L3(3–76) mutant, which does not contain the ZnF, could assemble a stable ATXN7-containing DUBm (Fig. 3C, lanes 2 to 4). These results indicate that the ZnF of ATXN7L3 is dispensable for its incorporation into the DUBm. In addition, the “linker” region of ATXN7L3 (amino acids 41 to 80), which is located N terminally from its ZnF-Sgf11 domain, is required for a stable interaction with ATXN7. Finally, the most N-terminal region of ATXN7L3 (Sus1 binding region) (Fig. 3A) bound ENY2 (data not shown), in agreement with the yeast structural data (3).

Similarly, we constructed deletion mutants of ATXN7 to analyze which region mediates its interaction with other subunits of the DUBm. We excluded the most N-terminal low-complexity sequences of ATXN7, which include the polyglutamine motif and are missing from its yeast orthologue, Sgf73, or from human paralogues, ATXN7L1 and ATXN7L2 (Fig. 3A; also see Fig. S4 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). A mu-

tant containing the ZnF-Sgf73 and the poorly conserved upstream sequence [ATXN7(75–172)] stably integrated within the DUBm, whereas another mutant lacking the ZnF-Sgf73 [ATXN7(75–142)] did not interact with ATXN7L3, USP22, or ENY2 (Fig. 3C, lanes 2 and 5). The ZnF of ATXN7 is therefore required for the assembly of the DUBm, in contrast with that of ATXN7L3, which is dispensable. This analysis allowed us to better understand the domains that are necessary to reconstitute a minimal recombinant DUBm, composed of the full-length USP22 and ENY2, the N-terminal region of ATXN7 (amino acids 75 to 172), and the most N-terminal region of ATXN7L3 (amino acids 3 to 76).

**Regulation of USP22 catalytic activity.** To understand which of the domains of the DUBm are required for the activation of USP22, we analyzed the enzymatic activities of the above-reconstructed USP22-containing complexes. To this end, we assessed the catalytic activity of USP22 in the different recombinant complexes using ubiquitin vinyl sulfone (Ub-VS), an irreversible and specific inhibitor of deubiquitinating enzymes. Ub-VS binds exclusively to an active deubiquitinase but not to an inactive one. Such binding creates a covalent bond between ubiquitin and the active site of the deubiquitinase, which can be revealed by Western blotting as a 7-kDa shift of the active enzyme corresponding to the size of ubiquitin.

When the free recombinant USP22 was incubated with Ub-VS, no shift could be detected, indicating that the enzyme alone is inactive under the assay conditions used (Fig. 4A). We then treated the purified ATXN7L3(3–151)-ENY2-USP22 and an ATXN7(75–172)-ATXN7L3(3–151)-ENY2-USP22 recombinant complexes with Ub-VS and analyzed the activity of USP22. USP22 was only partially shifted in the complex containing three subunits but was almost completely shifted in the complex containing the four subunits (Fig. 4A). These results together suggest that USP22 is partially activated by its interactions with ATXN7L3 and ENY2 but that its full activity is also dependent on additional interactions with ATXN7. Therefore, the full-length ENY2 and the N-terminal fragments of both ATXN7(75–172) and ATXN7L3(3–151) are sufficient to incorporate USP22 in a stable and fully active DUBm.

We showed that the ZnF from ATXN7, but not that from ATXN7L3, is required for the assembly of the DUBm. We therefore asked whether the ATXN7L3 ZnF plays a role in the regulation of USP22 enzymatic activity. Using two different fragments of ATXN7L3 that contain or do not contain the ZnF-Sgf11 [ATXN7L3(3–151) or ATXN7L3(3–76), respectively], we purified stable recombinant complexes as above (Fig. 3C) and tested them in the enzymatic assay using Ub-VS. This analysis revealed that the ATXN7L3 ZnF-Sgf11 has an additional stimulatory effect on USP22 catalytic activity (Fig. 4B).

We further analyzed USP22 activity in the different complexes in another *in vitro* deubiquitination assay using monoubiquitinated histone H2B, one of the natural substrates of USP22 (see below). Mononucleosomes purified from HeLa cells were incubated with different USP22-containing complexes and the levels of H2Bub were analyzed by Western blotting. Similarly to the results obtained by using Ub-VS, the free USP22 alone had no detectable activity on H2Bub in this assay. In contrast, a recombinant DUBm containing the ATXN7L3(3–151) mutant efficiently deubiquitinated H2Bub. This activity was lost when Ub-VS was added to the reaction

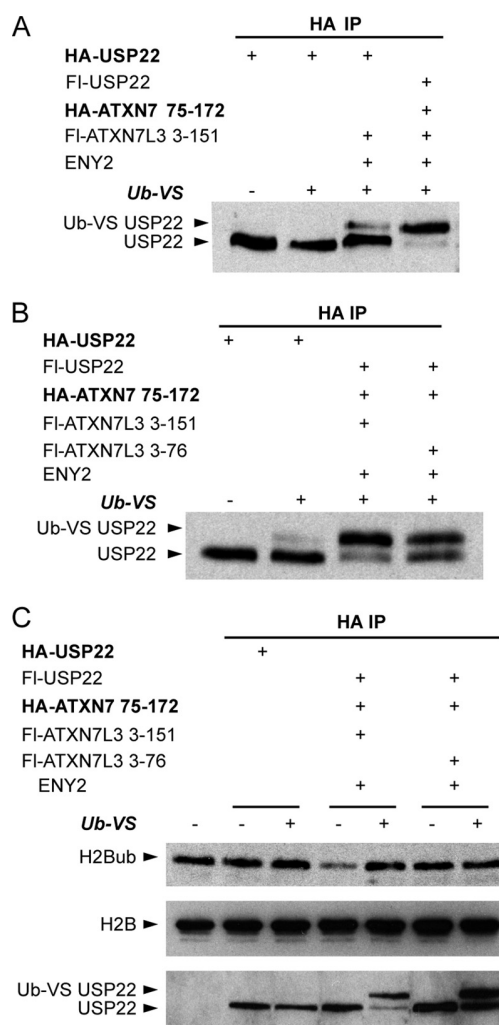


FIG. 4. USP22 catalytic activity is regulated by the other subunits of the DUB module. (A) The DUB module containing the first domain of ATXN7 is active. (B) The ZnF-Sgf11 of ATXN7L3 is required for the full activity of USP22. Equal amounts of the different purified complexes described for Fig. 2 and 3 were incubated with or without 5  $\mu$ M the suicide substrate ubiquitin vinyl sulfone (Ub-VS). This compound can form a covalent bond only with the active form of USP22, which is revealed by a 7-kDa shift of the enzyme when analyzed by Western blotting. (C) The above-described complexes were incubated with purified mononucleosomes with or without Ub-VS. H2B monoubiquitination was assessed by Western blotting using an anti-H2Bub antibody and USP22 reactivity with Ub-VS was measured as described for panel B.

mixture, further indicating that Ub-VS reacted with the catalytic cysteine of USP22 (Fig. 4C). Interestingly, a complex containing an ATXN7L3 mutant that lacks the ZnF-Sgf11 [ATXN7L3(3–76)] did not deubiquitinate H2Bub in this assay. However, the addition of Ub-VS in this reaction induced a partial shift of USP22 from this latter complex (Fig. 4C). Altogether, these results suggest that the ATXN7L3 ZnF-Sgf11 plays a role in the allosteric regulation of USP22 but also in the binding of the SAGA DUBm to nucleosomes. This hypothesis is strongly supported by structural analysis of the yeast complex showing that the ZnF-Sgf11 of Sgf11 contains a patch of strictly

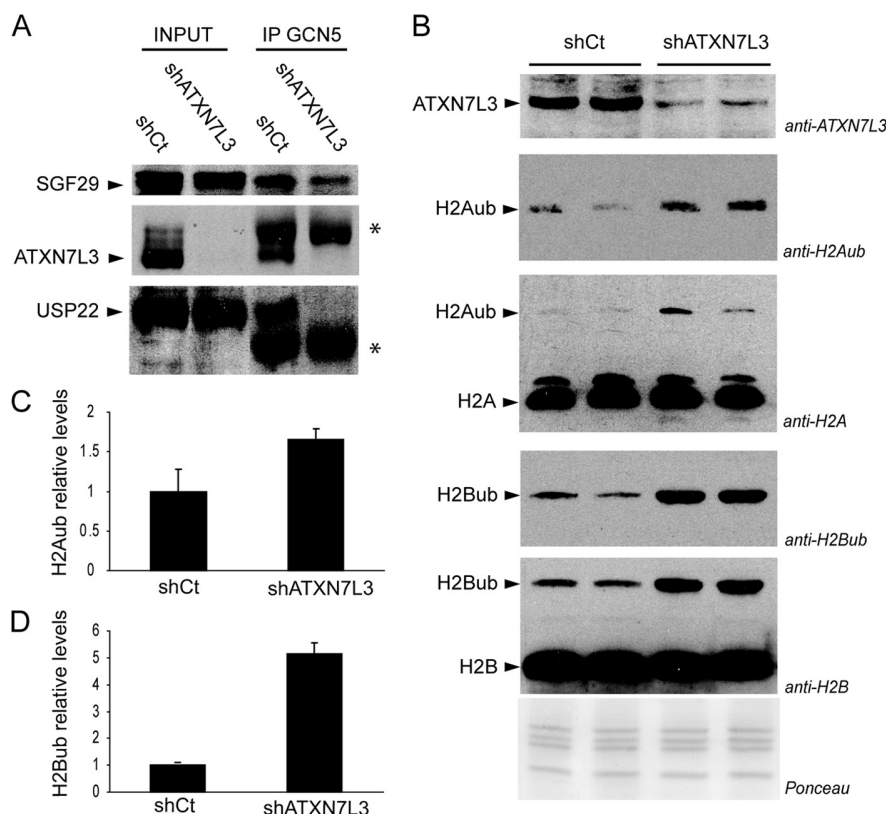


FIG. 5. Monoubiquitinated H2A and H2B are *in vivo* substrates of hSAGA deubiquitination activity. (A) SAGA complexes were immunopurified from HeLa cells stably expressing a shRNA against ATXN7L3 (shATXN7L3) or a control shRNA (shCt) by using anti-GCN5 antibodies and analyzed by Western blotting with the indicated antibodies. IgG heavy chains are indicated (\*). (B) Total histones were purified by acidic extractions from these cell lines. Levels of ubiquitinated histones were analyzed with specific antibodies as indicated. (C and D) H2Aub and H2Bub relative levels were quantified and normalized to H2A and H2B levels, respectively.

conserved positively charged residues, which are exposed to solvent and may be involved in interactions with nucleosomal DNA to target the DUBm to its substrate (12, 29) (see Fig. S3 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). Thus, these results show an allosteric regulation of USP22 through multiple interactions with different domains of the other subunits of the DUBm including domains that are not necessary for a stable assembly of this module.

***In vivo* deubiquitination activity of the SAGA complex.** Although the human SAGA complex was shown to deubiquitinate H2Aub and H2Bub *in vitro* (35–37), it has not been demonstrated that these substrates are recognized by SAGA *in vivo*. To specifically inactivate the SAGA deubiquitination activity, we generated a HeLa cell line expressing a specific shRNA against ATXN7L3, an hSAGA subunit that is required for both the assembly of the DUBm and the activation of USP22. Immunopurification of GCN5-associated proteins from this cell line, in which the ATXN7L3 levels were strongly reduced, revealed that not only ATXN7L3 but also USP22 was lost from the SAGA complex (Fig. 5A). This result further indicates that ATXN7L3 is crucial for the integrity and thus the activity of the SAGA DUBm, in good agreement with our *in vitro* interaction studies (Fig. 2 and 3).

Next, we asked whether the global levels of monoubiquitinated histones are affected in this cell line, in which the SAGA

deubiquitination activity is impaired. We observed a 1.5-fold increase of H2Aub levels in ATXN7L3 knockdown cells, indicating that human SAGA can participate in the deubiquitination of histone H2A *in vivo* (Fig. 5B and C). Moreover, the global levels of H2Bub were dramatically increased after ATXN7L3 knockdown, indicating that SAGA is a major H2Bub deubiquitinase and that this activity can not be fully compensated for by other deubiquitinases (Fig. 5B and D).

**Induction of SAGA-dependent gene expression strongly increased H2B ubiquitination at the transcribed regions but not at the promoters.** To further understand the regulation of H2A and H2B ubiquitination in human cells (22, 37), we analyzed the dynamics of these modifications at SAGA-regulated genes in two different cell types. We used LNCaP cells, in which we studied the androgen receptor (AR)-dependent stimulation of the *PSA* gene in the presence or absence of an AR ligand, the dihydrotestosterone analogue R1881 (Fig. 6A and B). We also studied two endoplasmic reticulum (ER) stress genes (*HERPUD* and *ERP70*) in HeLa cells before and following induction of ER stress with thapsigargin (Fig. 7A to C). To measure *in vivo* histone ubiquitination, we used chromatin immunoprecipitation (ChIP) experiments with antibodies recognizing either H2Aub or H2Bub in the two cellular systems. Unfortunately, the commercially available anti-H2Aub anti-



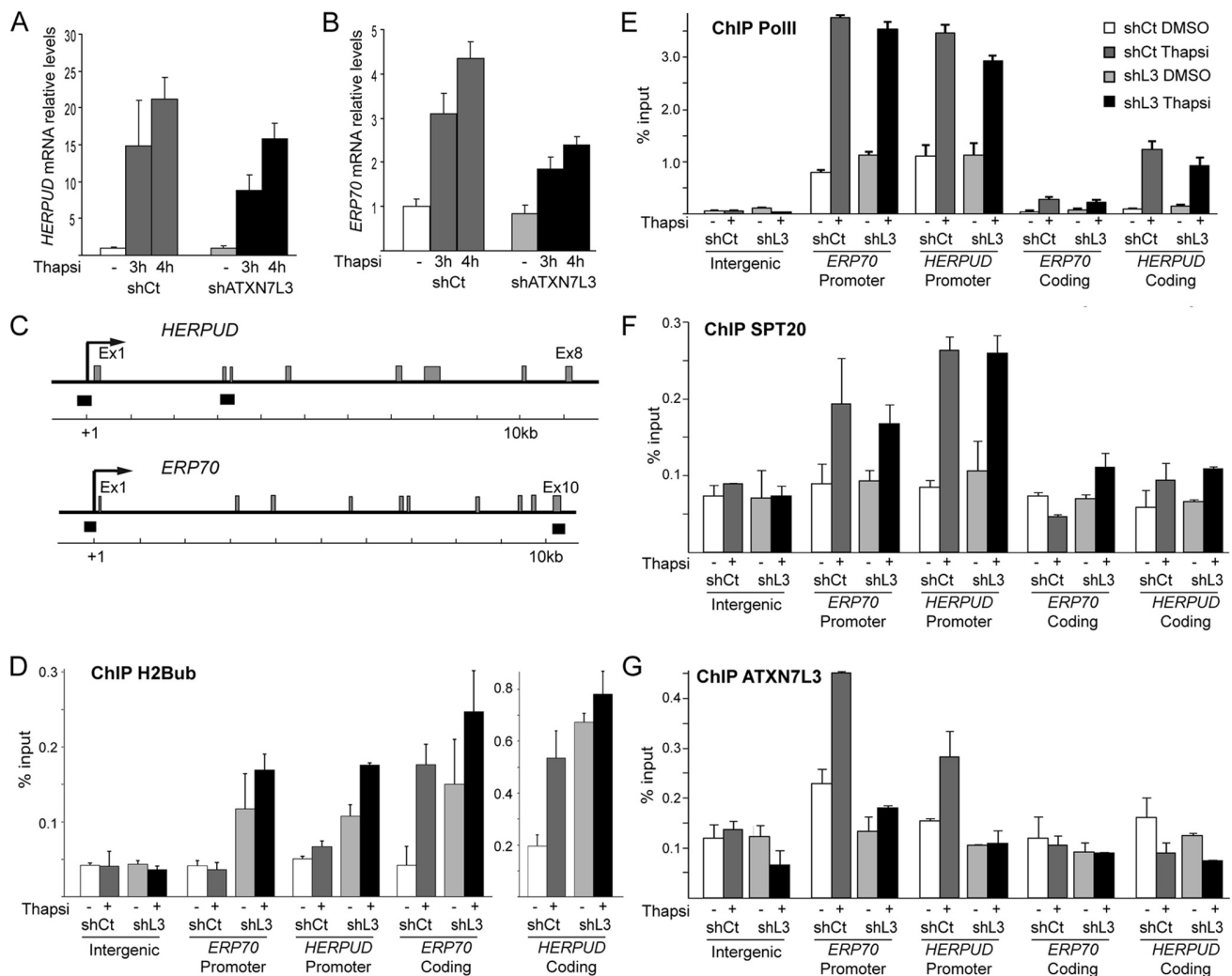


FIG. 7. Dynamics of H2B ubiquitination at the promoter and the transcribed region of ER stress genes. (A) HeLa cells stably expressing a shRNA against ATXN7L3 or a control shRNA were treated with thapsigargin for 3 or 4 h or with DMSO as a control. mRNA levels of ER stress genes (*HERPUD* and *ERP70*) were analyzed by qPCR of reverse transcribed RNA and normalized to *CycloB* mRNA levels and are presented as mean values  $\pm$  SD of three independent experiments. (B) Schematic representation of the genomic locus of the *HERPUD* and *ERP70* genes. Black boxes represent primer sets used for ChIP analysis. (D to G) Control and ATXN7L3 shRNA HeLa cell lines were incubated with thapsigargin for 4 h or with vehicle only (DMSO). Quantitative ChIP assays were performed for H2Bub (D), RNA Pol II (E), SPT20 (F), and ATXN7L3 (G). Precipitated DNA was quantified by real-time qPCR using primers on *ERP70* and *HERPUD* promoters and coding regions and on an intergenic control region. The values (mean  $\pm$  SD of triplicates) are expressed as percentages of the amount of immunoprecipitated DNA normalized to the respective input DNA signal (% input). These results are representative of three independent experiments. The same color code is used in all panels.

revealed by a SPT20 ChIP (Fig. 7F and G). These results indicate that the deubiquitination activity associated with the hSAGA complex is required for full activation of different SAGA-dependent genes.

**SAGA deubiquitinates H2Bub at the promoter and transcribed regions of SAGA-regulated genes.** To understand how the control of H2B ubiquitination by SAGA contributes to the regulation of gene expression, we compared H2Bub levels in ATXN7L3 knockdown cells with those of control cells at the different gene regulatory regions. Surprisingly, our ChIP analyses showed that H2B ubiquitination at the three tested promoters was increased by 2.5- to 3.5-fold in ATXN7L3 knockdown cells before induction when compared to the noninduced control cells (Fig. 6C and 7D, compare shCt - to shL3 - in

promoter panels). The increase of H2Bub levels at promoters when the SAGA deubiquitination activity is impaired reveals the activity of H2B-ubiquitinating enzymes at the promoter region of these genes, even in the absence of induction. Interestingly, this strong increase of H2Bub under noninduced conditions in the ATXN7L3 knockdown cells had no effect on transcription, as we did not measure transcriptional changes from the three studied promoters before induction (Fig. 6A and 7A and B). At these promoters, we observed a further 1.5- to 2-fold increase of H2Bub after induction in the shATXN7L3 cell line as compared to the noninduced conditions. This result demonstrates an increased activity of H2B-ubiquitinating enzymes at promoters following induction of gene expression.

Next, we analyzed the transcribed regions of these genes and



observed that before induction, the H2Bub levels strongly increased (by 3- to 4-fold) in the ATXN7L3 knockdown cells as compared to control cells (Fig. 6C and 7D). This result indicates that SAGA deubiquitinates H2Bub not only at the promoter but also in the transcribed region of these SAGA-dependent genes. In contrast, an increased recruitment of SAGA was observed only at the promoter but not in the gene body of *HERPUD* and *ERP70* upon induction (Fig. 7F and G). Interestingly, the H2Bub increase in the coding regions before induction did not induce aberrant transcription from the studied genes in these cells (Fig. 6A and 7A and B). After induction, H2Bub levels were further increased (by about 1.5-fold) in the ATXN7L3 knockdown cells when compared to control cells (Fig. 6C and Fig. 7D, compare shCt + with shL3 + in coding panels). Concomitantly, transcription activation was reduced by about 25 to 45% (Fig. 6A and 7A and B). However, the recruitment of RNA Pol II at the promoter and the coding regions of *HERPUD* and *ERP70* was not modified after ATXN7L3 knockdown in either induced or noninduced conditions (Fig. 7E).

It is noteworthy that, at each individual region examined in the two cellular systems, the total H2B levels were not significantly changed after ATXN7L3 knockdown (see Fig. S7 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)). However, at the *PSA* enhancer and promoter region, an approximately 2-fold reduction of H2B occupancy could be detected after induction in both control and ATXN7L3 knockdown cells (see Fig. S7 at the URL given above), which likely reflects nucleosome eviction upon transcriptional activation.

In summary, our data indicate that different dynamic equilibria of H2B ubiquitination are established at different gene regulatory elements. Moreover, our results suggest that at the studied promoter and transcribed regions, changes of H2B ubiquitination alone are not sufficient to trigger parallel changes in gene expression.

## DISCUSSION

**The allosteric regulation of human USP22.** Recent structural and functional studies of the yeast DUBm demonstrated an allosteric regulation of Ubp8 through multiple interactions with other subunits of the DUBm (12, 29). In good agreement, we show here that the enzymatic activity of human USP22 is similarly regulated by its interactions with ATXN7, ATXN7L3, and ENY2. We show that this regulation is mediated not only by domains that are required for the assembly of DUBm but also by the ZnF of ATXN7L3, which is not involved in the assembly of the DUBm. Our results thus suggest an allosteric regulation of human USP22 in which the ZnF domains of ATXN7 and ATXN7L3 play a crucial role. Therefore, the mechanisms of the regulation of SAGA deubiquitination activity appear remarkably conserved through evolution.

Similarly to the yeast complex, the assembly of the human DUBm is mediated through multiple interactions between the four different subunits. These interactions mainly involve the ZnF-UBP of USP22 and the N-terminal regions of ATXN7 and ATXN7L3. However, subtle differences between the yeast and the human complex can be detected. Our reconstruction experiments showed that, in contrast to yeast Ubp8, (12) the ZnF-UBP from USP22 is not sufficient to build up a stable

DUBm (see Fig. S1 at [http://igbmc.fr/Lang\\_mcb2011](http://igbmc.fr/Lang_mcb2011)), suggesting that the protease domain of USP22 is required to stabilize the DUBm. Functional studies of the yeast DUBm revealed that Sgf73 was absolutely required for Ubp8 activation, as an Ubp8-Sgf11-Sus1 complex was inactive (11), whereas in the case of the human DUBm the USP22-ATXN7L3-ENY2 complex was partially active (Fig. 4A). In addition, the N-terminal ZnF-Sgf73 domain of human ATXN7 is required for the assembly of the DUBm (Fig. 3), although the homologue domain in yeast Sgf73 is dispensable (12). Therefore, it appears that structural elements such as the ZnF domain of Sgf73 or the protease domain of Ubp8, which are mainly involved in the enzymatic regulation of the yeast DUBm, have evolved to participate in the stability of the human DUBm in addition to playing their role in the allosteric regulation of the enzyme. This suggests a coevolution of both the stability of the DUBm and the enzymatic regulation of the deubiquitinase, highlighting the biological relevance of the regulation of the deubiquitination activity of SAGA.

The allosteric regulation of USP22 could play an important role in controlling the specific recognition of the substrates by the enzyme. USP22 within the SAGA complex is recruited at specific loci on the genome, where it will act on its specific substrates, H2Aub and H2Bub. Importantly, the mechanism of allosteric regulation described here would prevent the free USP22 enzyme that is not incorporated into the SAGA complex from acting on nonspecific substrates.

**The majority of SAGA DUB activity on H2Bub is not redundant with other USPs.** A recent study suggested that USP22 might be found in complexes other than SAGA (31), and ENY2 is expected to be part of an mRNA export complex (16, 28). Therefore, the downregulation of USP22 or ENY2 might alter functions other than the SAGA deubiquitination activity. Moreover, human cells contain two additional ATXN7 paralogues (ATXN7L1 and ATXN7L2), which are highly similar with ATXN7 (4) and which could likely compensate for the loss of ATXN7. In contrast, the downregulation of ATXN7L3, which has not been described in any protein complex other than SAGA and whose function is not redundant with that of any other human protein, should exclusively impair the SAGA deubiquitination activity.

Therefore, by comparing H2Bub levels between control cells and cells in which ATXN7L3 is downregulated, we could directly measure where the SAGA deubiquitination activity is acting on different genomic positions and indirectly where the SAGA complex is recruited. The analysis of different inducible SAGA-dependent genes using anti-H2Bub ChIP experiments revealed several unexpected features of the SAGA deubiquitination activity. First, SAGA appears to have a global deubiquitination activity as revealed by the important increase of global H2Bub levels in ATXN7L3 knockdown cells. In agreement, our anti-H2Bub ChIP experiments suggest that SAGA could maintain a low level of H2B ubiquitination at different gene regulatory elements genome-wide. Indeed, even in non-induced conditions, the inactivation of SAGA deubiquitination activity increased the H2Bub levels at the enhancer, promoter, and transcribed regions of the three tested genes. This demonstrates that SAGA regulates H2Bub at different gene regions irrespective of their transcriptional status. However, our ChIP experiments failed to detect stable SAGA recruitment at

these regions in the noninduced conditions, suggesting a very transient and dynamic association of SAGA on these genes. After gene induction, SAGA recruitment was detected only at the tested promoters but not in the transcribed regions. In contrast, the deubiquitination activity of SAGA, as tested by H2Bub ChIP in the shATXN7L3 versus wild-type cells, revealed SAGA activity at both promoter and transcribed regions. These results suggest that after activation, SAGA association with chromatin in the gene body is transient, whereas it is less dynamic at the promoter. Such stable recruitment of SAGA at the promoter could be explained by interactions with promoter-associated activators and histone posttranslational modifications recognized by the different reader domains of SAGA, which in turn would enhance the residency time of SAGA. Finally, the striking difference in H2Bub levels between wild-type and ATXN7L3 knockdown cells indicates that most of the SAGA deubiquitination activity is not redundant with that of other deubiquitinases. Although different H2B deubiquitinases have been identified in yeast and in *Drosophila*, USP3 was the only other human enzyme suggested to deubiquitinate H2Bub (23). However, the precise role of USP3 has still to be determined.

**The dynamics of H2B ubiquitination and deubiquitination at promoters and transcribed regions are differentially regulated.** Our ChIP analyses on different SAGA-dependent genes showed an increased H2B ubiquitination at the corresponding promoters in ATXN7L3 knockdown cells as compared to control cells, under noninduced conditions. This observation suggests the interesting possibility that the deubiquitination activity of SAGA is already present together with the activities that ubiquitinate H2B at the inducible gene promoters before activation, leading to equilibrium between these two antagonistic activities. Moreover, the lack of changes in transcription under noninduced conditions in the ATXN7L3-downregulated cell line versus wild-type cells suggests that at the studied inducible promoters, the deregulation of the deubiquitination activity and the consequent upregulation of H2B ubiquitination have no direct consequence on transcription.

Our results also show that the SAGA deubiquitination activity can very rapidly reverse the action of H2B-ubiquitinating enzymes at the promoter of SAGA-dependent genes after induction. In yeast cells, it has been shown that the machinery of H2B ubiquitination is recruited to promoters by activators where coactivator complexes such as SAGA are also expected to be recruited through interactions with specific activators. However, the time course recruitment of each complex is unknown, and it is possible to speculate that both complexes are recruited simultaneously by an activator at the promoter of SAGA-dependent genes. In this case, a very fast deubiquitination may be necessary to start efficiently a new cycle of transcription and to maintain a high level of H2B-ubiquitinating enzymes around the promoter. Therefore, a constant cycling of H2B ubiquitination and deubiquitination could be required to maintain the respective enzymes and their associated factors at the promoter of SAGA-dependent genes.

In the transcribed region of these SAGA-dependent genes, the H2Bub levels were significantly increased after induction, suggesting that SAGA is less stably recruited at these regions than at the promoters. Alternatively, much higher levels of H2Bub in the transcribed region could exceed the ability of

SAGA to deubiquitinate H2B. In addition, it has been suggested that H2Bub facilitates the action of FACT, thereby facilitating transcription elongation (26). Our results do not support this model, as in ATXN7L3 knockdown cells in the noninduced conditions, a dramatic increase of H2Bub levels in the transcribed region of SAGA-dependent genes did not correlate with increased transcription of these genes.

Our study highlights the tight control of USP22 catalytic activity through an allosteric mechanism based on complex interactions with other subunits of the SAGA DUBm. Furthermore, our data suggest that SAGA has a global deubiquitination function that appears required to maintain low levels of H2Bub at the promoter and in the transcribed region of genes in noninduced conditions. Our study demonstrates that, after induction, SAGA deubiquitinates H2Bub at both the promoters and the transcribed regions of SAGA-regulated genes but that this complex has very different dynamics on these distinct regions.

#### ACKNOWLEDGMENTS

We thank Arnaud Krebs and Anne Helmrich for comments on the manuscript. We also thank Pang-Hung Hsu for mass spectrometry analysis, Christophe Romier and Imre Berger for gifts of materials, and Elisabeth Scheer and Isabelle Kolb-Cheynel for technical help.

J.B. received an FRM fellowship and G.L. was supported by a fellowship from the Ligue contre le cancer and from CSC. This work was supported by grants from the ANR (ANR-09-MNPS-023, ANR-09-PRI-0031, ANR-10-INTB-1201) to D.D. and by grants from ANR (ANR-09-BLAN-0266), the EU (EUTRACC, LSHG-CT-2007-037445 and EPIDIACAN), and the INCA (2008 Ubian) to L.T.

#### REFERENCES

- Bonnet, J., C. Romier, L. Tora, and D. Devys. 2008. Zinc-finger UBPs: regulators of deubiquitylation. *Trends Biochem. Sci.* **33**:369–375.
- Daniel, J. A., et al. 2004. Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J. Biol. Chem.* **279**:1867–1871.
- Ellisdon, A. M., D. Jani, A. Kohler, E. Hurt, and M. Stewart. 2010. Structural basis for the interaction between yeast Spt-Ada-Gcn5 acetyltransferase (SAGA) complex components Sgf11 and Sus1. *J. Biol. Chem.* **285**:3850–3856.
- Helminger, D., et al. 2004. Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum. Mol. Genet.* **13**:1257–1265.
- Helminger, D., L. Tora, and D. Devys. 2006. Transcriptional alterations and chromatin remodeling in polyglutamine diseases. *Trends Genet.* **22**:562–570.
- Henry, K. W., et al. 2003. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev.* **17**:2648–2663.
- Ingvarsdottir, K., et al. 2005. H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol. Cell Biol.* **25**:1162–1172.
- Kim, J., S. B. Hake, and R. G. Roeder. 2005. The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interaction. *Mol. Cell Biol.* **25**:759–770.
- Kim, J., and R. G. Roeder. 2009. Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. *J. Biol. Chem.* **284**:20582–20592.
- Kohler, A., et al. 2006. The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitylation through its interaction with Ubp8 and Sgf11. *Mol. Biol. Cell* **17**:4228–4236.
- Kohler, A., M. Schneider, G. G. Cabal, U. Nehrass, and E. Hurt. 2008. Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat. Cell Biol.* **10**:707–715.
- Kohler, A., E. Zimmerman, M. Schneider, E. Hurt, and N. Zheng. 2010. Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell* **141**:606–617.
- Komander, D., M. J. Clague, and S. Urbe. 2009. Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**:550–563.
- Koutelou, E., C. L. Hirsch, and S. Y. Dent. 2010. Multiple faces of the SAGA complex. *Curr. Opin. Cell Biol.* **22**:374–382.
- Krebs, A. R., et al. 2010. ATAC and Mediator coactivators form a stable complex and regulate a set of non-coding RNA genes. *EMBO Rep.* **11**:541–547.

16. Kurshakova, M. M., et al. 2007. SAGA and a novel *Drosophila* export complex anchor efficient transcription and mRNA export to NPC. *EMBO J.* **26**:4956–4965.
17. Lee, K. K., S. K. Swanson, L. Florens, M. P. Washburn, and J. L. Workman. 2009. Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. *Epigenetics Chromatin* **2**:2.
18. Li, B., M. Carey, and J. L. Workman. 2007. The role of chromatin during transcription. *Cell* **128**:707–719.
19. Martinez, E. 2002. Multi-protein complexes in eukaryotic gene transcription. *Plant Mol. Biol.* **50**:925–947.
20. Minsky, N., et al. 2008. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat. Cell Biol.* **10**:483–488.
21. Monestier, M., P. Decker, J. P. Briand, J. L. Gabriel, and S. Muller. 2000. Molecular and structural properties of three autoimmune IgG monoclonal antibodies to histone H2B. *J. Biol. Chem.* **275**:13558–13563.
22. Nagy, Z., et al. 2009. The human SPT20-containing SAGA complex plays a direct role in the regulation of endoplasmic reticulum stress-induced genes. *Mol. Cell. Biol.* **29**:1649–1660.
23. Nicassio, F., et al. 2007. Human USP3 is a chromatin modifier required for S phase progression and genome stability. *Curr. Biol.* **17**:1972–1977.
24. Nijman, S. M., et al. 2005. A genomic and functional inventory of deubiquitinating enzymes. *Cell* **123**:773–786.
25. Pascual-Garcia, P., et al. 2008. Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2. *Genes Dev.* **22**:2811–2822.
26. Pavri, R., et al. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**:703–717.
27. Reyes-Turcu, F. E., et al. 2006. The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* **124**:1197–1208.
28. Rodriguez-Navarro, S., et al. 2004. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**:75–86.
29. Samara, N. L., et al. 2010. Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* **328**:1025–1029.
30. Schnitzler, G. R. 2001. Isolation of histones and nucleosome cores from mammalian cells. *Curr. Protoc. Mol. Biol.* **21**:21.5.
31. Sowa, M. E., E. J. Bennett, S. P. Gygi, and J. W. Harper. 2009. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**:389–403.
32. Timmers, H. T., and L. Tora. 2005. SAGA unveiled. *Trends Biochem. Sci.* **30**:7–10.
33. Weake, V. M., and J. L. Workman. 2008. Histone ubiquitination: triggering gene activity. *Mol. Cell* **29**:653–663.
34. Xiao, T., et al. 2005. Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol. Cell. Biol.* **25**:637–651.
35. Zhang, X. Y., H. K. Pfeiffer, A. W. Thorne, and S. B. McMahon. 2008. USP22, an hSAGA subunit and potential cancer stem cell marker, reverses the polycomb-catalyzed ubiquitylation of histone H2A. *Cell Cycle* **7**:1522–1524.
36. Zhang, X. Y., et al. 2008. The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol. Cell* **29**:102–111.
37. Zhao, Y., et al. 2008. A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol. Cell* **29**:92–101.
38. Zhu, B., et al. 2005. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol. Cell* **20**:601–611.

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Oligonucleotides used for the qPCR experiments:

ATXN7L3 RT f: 5'-GTCTTTGTCTGGCCTGGATAAC-3'  
ATXN7L3 RT r: 5'-GGTGTACCTCAAAGCAGAATCC-3';

HERPUD RT f: 5'-GGTTGGGGTCTTCAGTTTCAGG-3'  
HERPUD RT r: 5'-CTACTCCTCCCTGAGCAGATTC-3'

ERP70 RT f: 5'-TGACGGGTCCCTTGTGTTCT-3'  
ERP70 RT r: 5'-AGAGTGGGGAGGATGTCAATG-3'

CyclophilinB RT f: 5'-CTTCCCCGATGAGAACTTCCAAACT-3'  
CyclophilinB RT r: 5'-CACCTCCATGCCCTCTAGAACTTT-3'

PSA RT f: 5'-CACCTGCTCGGGTGATTCTG-3'  
PSA RT r: 5'-CCACTTCCGGTAATGCACCA-3'

GAPDH RT f: 5'-GAAGGTGAAGGTCGGACTC-3'  
GAPDH RT r: 5'-GAAGATGGTGATGGGATTTC-3'

PSA ARE III f: 5'-GGGGTTTGTGCCACTGGTGAG-3'  
PSA ARE III r: 5'-GGGAGGCAATTCTCCATGGTT-3'

PSA ARE I/II f: 5'-GCCAAGACATCTATTTTCAGGAGC-3'  
PSA ARE I/II r: 5'-CCCACACCCAGAGCTGTGGAAGG-3'

PSA exon4 f: 5'-GTGTGTGGACCTCCATGTTATT-3'  
PSA exon4 r: 5'-CCACTCACCTTTCCCCTCAAG-3'

HERPUD Promoter f: 5'-CTGCAACGACAGTTCACGTCTCT-3'  
HERPUD Promoter r: 5'-GGTTGCATCAGCCCGTGCCC-3'

ERP70 Promoter f: 5'-CTCCGAGCCCTAACGTGAGGTGCC-3'  
ERP70 Promoter r: 5'-CCCAGGCTCCGCCTCTCCTGC-3'

HERPUD exon2-3 f: 5'-CTGGGAAGCTGTTGTTGGAT-3'  
HERPUD exon2-3 r: 5'-AAAACATGCCGTTTTTCCTG-3'

ERP70 exon10 f: 5'-GGCCTTCAAAGCTCTTCCTT-3'  
ERP70 exon10 r: 5'-AGCCTGGCCAAGAAGTACAA-3'

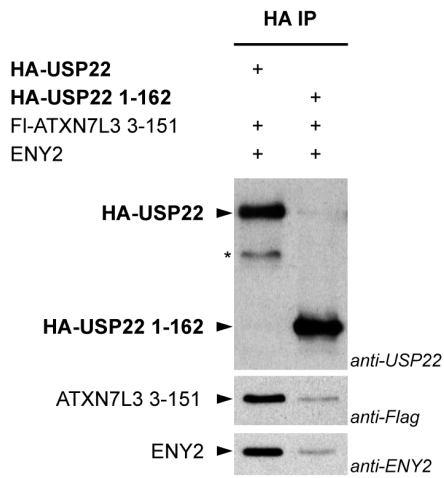
Non coding region f: 5'-TGGAACCTTCTGGAAGACACTGGAA-3'  
Non coding region r: 5'-TACACCACTCAAGGGAAACTGGAA-3'

**SUPPLEMENTAL TABLE 1 : SAGA subunits identified in a mass spectrometry analysis of USP22-associated proteins**

<b>Usual name</b>	<b>Swiss-Prot Entry</b>	<b>Unique peptides</b>	<b>Protein coverage (%)</b>
TRRAP	gi 116242829 sp Q9Y4A5.3	381	34
ATXN7	gi 5921169 sp O15265.1	19	18
GCN5	gi 209572743 sp Q92830.3	73	26.1
ATXN7L2	gi 74745198 sp Q5T6C5.1	16	14.1
ATXN7L1	gi 160413379 sp Q9ULK2.2	7	5.9
SPT20	gi 71151886 sp Q8NEM7.1	29	19.6
TAF6L	gi 46577572 sp Q9Y6J9.1	74	37
TAF5L	gi 46577305 sp O75529.1	43	28.9
SPT7L	gi 23396859 sp O94864.1	14	16.9
USP22*	gi 78103328 sp Q9UPT9.2	13	12
ADA3	gi 46397743 sp O75528.1	39	39.8
ADA2b	gi 166225686 sp Q86TJ2.2	20	20.2
ADA1	gi 74760738 sp Q96BN2.1	37	43.9
ATXN7L3	gi 121948758 sp Q14CW9.1	5	10.8
SPT3	gi 20978763 sp O75486.2	11	22.3
SGF29	gi 74731608 sp Q96ES7.1	29	42
TAF9	gi 2498981 sp Q16594.1	16	16.7
TAF10	gi 3024688 sp Q12962.1	12	34.4
TAF12	gi 3024708 sp Q16514.1	21	39.8

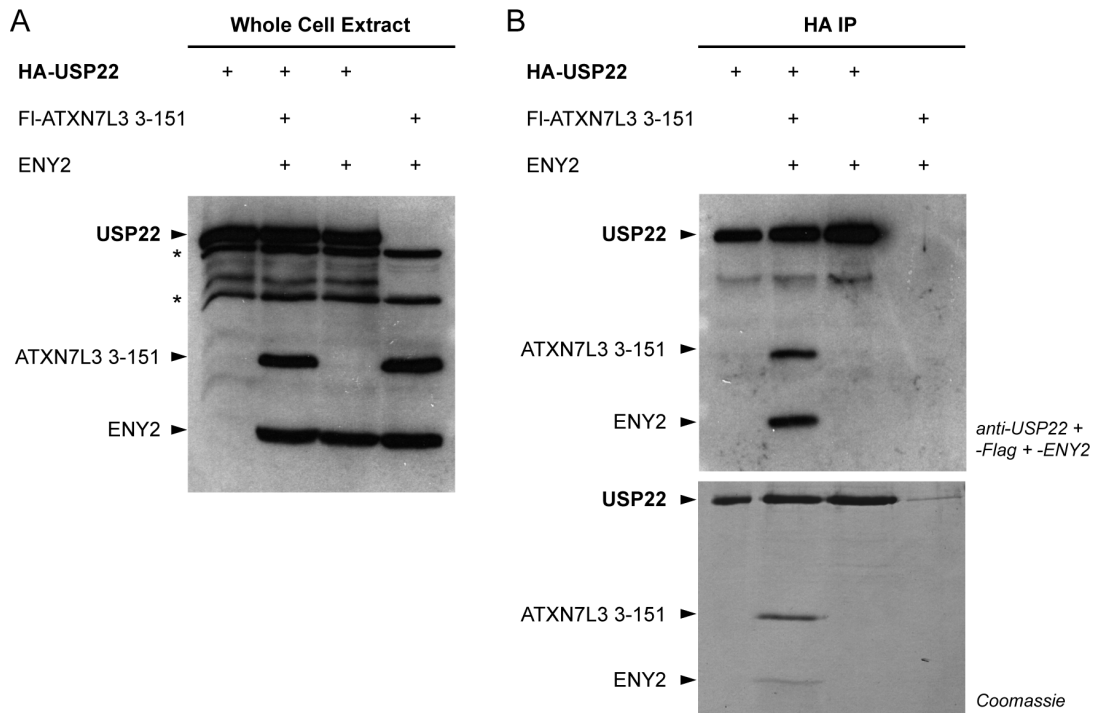
\* This analysis retrieved all known SAGA subunits except ENY2 but did not identify any other deubiquitinase with more than two unique peptides.

## SUPPLEMENTAL FIGURES



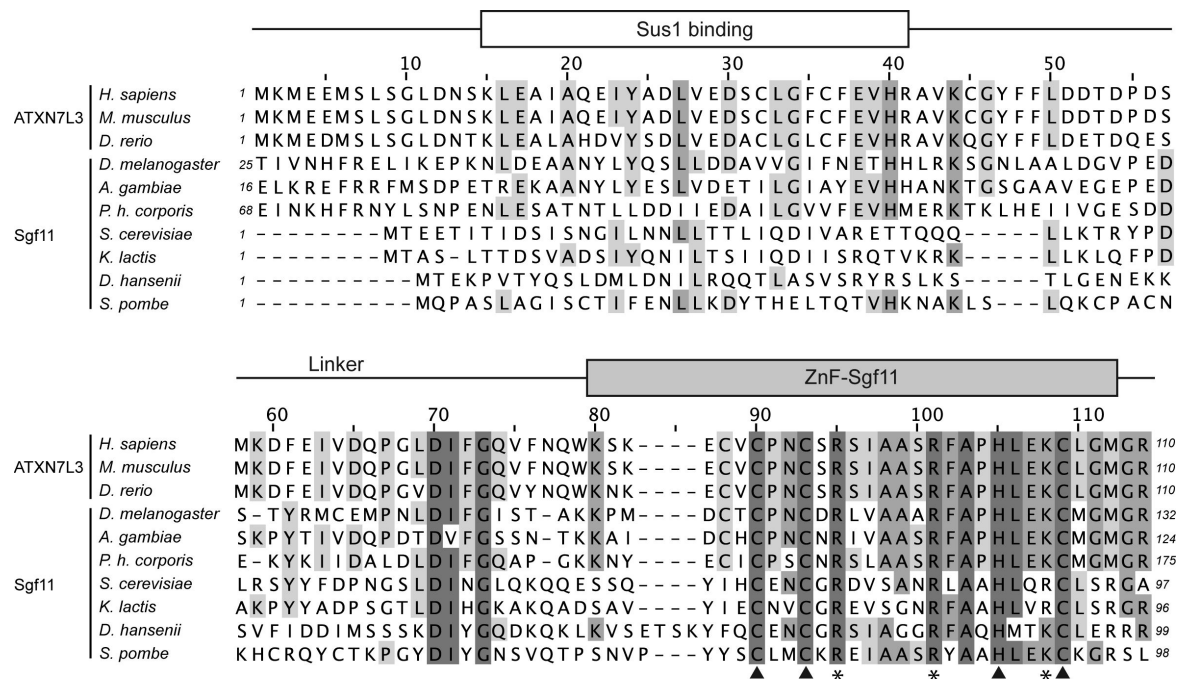
**Figure S1.**

The ZnF-UBP of USP22 did not assemble a stable complex with ATXN7L3 and ENY2. ENY2 and the N-terminal part of ATXN7L3 (FI-ATXN7L3 3-151) were co-expressed with either the full-length USP22 (HA-USP22) or the ZnF-UBP domain of USP22 (HA-USP22 1-162). Proteins associated with these HA-tagged USP22 constructs were analyzed by western blotting using the indicated antibodies after immunopurification using anti-HA antibodies (HA-IP).



### Figure S2.

Specific interaction of ATXN7L3 with USP22 and ENY2 (A) Sf9 cells were infected with different combinations of baculoviruses expressing HA-USP22, Flag-ATXN7L3 3-151 and ENY2 as indicated. Whole cell extracts were analyzed by western blotting. (B) USP22-associated proteins were analyzed after anti-HA immunoprecipitation followed by peptide elution. Western blotting were revealed using a anti-USP22, anti-Flag and anti-ENY2 antibodies. Asterisks (\*) indicate non specific bands.

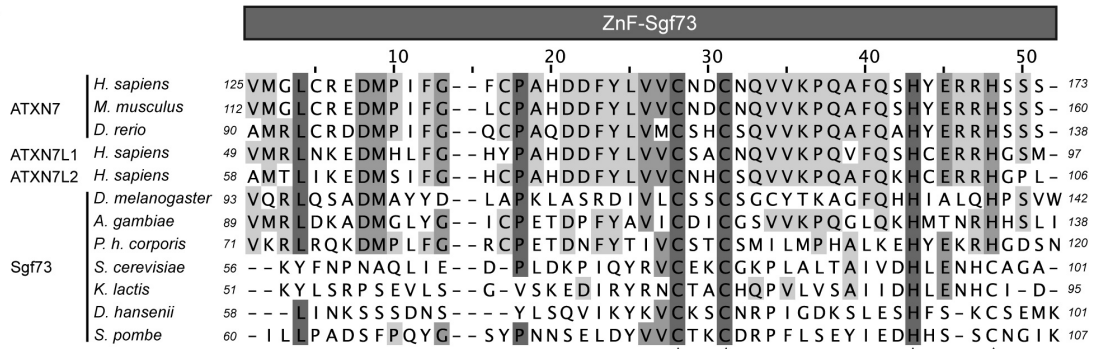


### Figure S3.

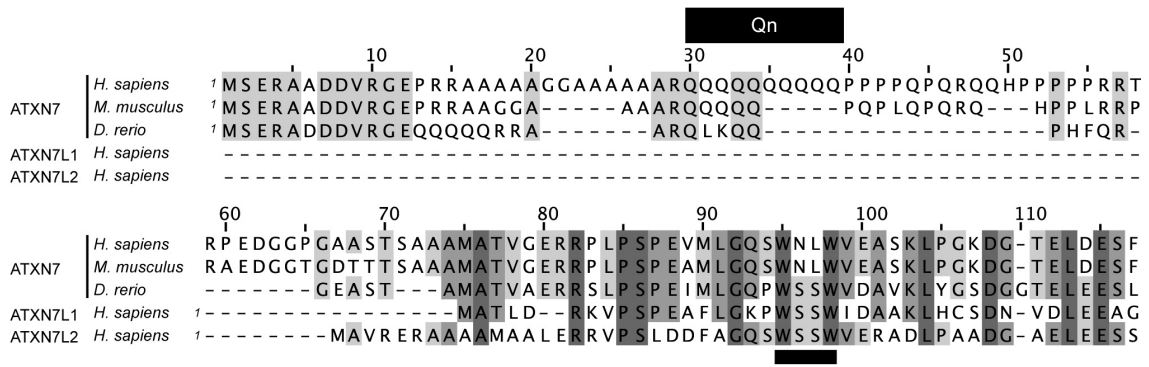
Sequence alignment of hATXN7L3 orthologs from representative species.

Only the homology region with Sgf11 was taken into account. Three main regions of sequence conservation (Sus1 binding, Linker and ZnF-Sgf11) corresponding to functional and structural regions of ySgf11 are indicated. Conserved residues with 90-100%, 70-80%, 50-60% identity are represented with dark grey, grey and light grey shading, respectively. Black triangles label the residues of the ZnF-Sgf11 that coordinate the zinc ion. Asterisks indicate the conserved basic patch within the ZnF-Sgf11 that may be involved in interactions with the phosphate backbone of nucleosomal DNA.

A



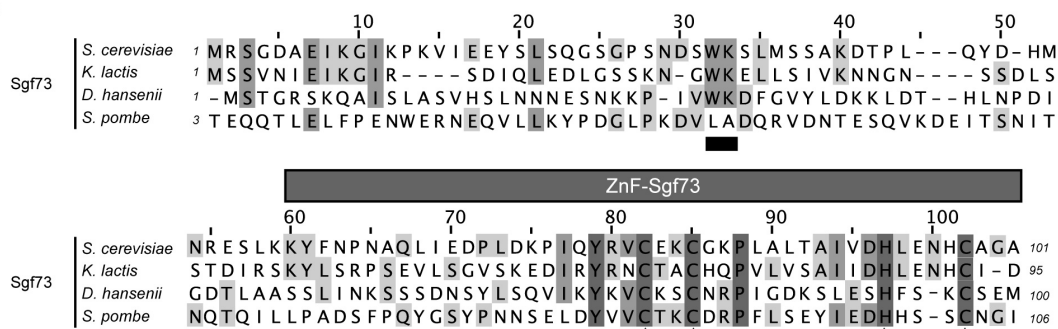
B



C



D

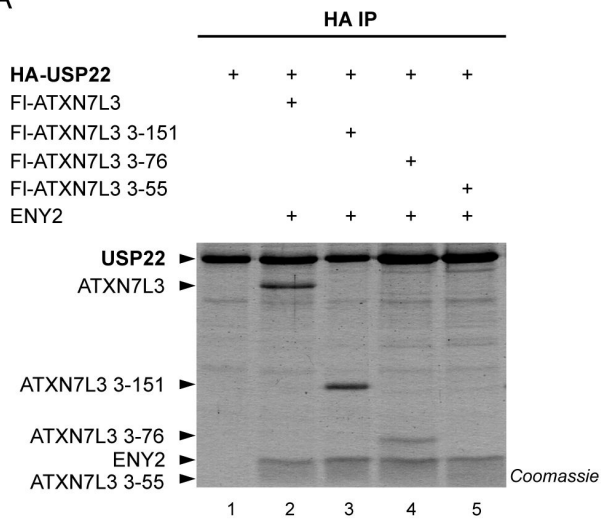




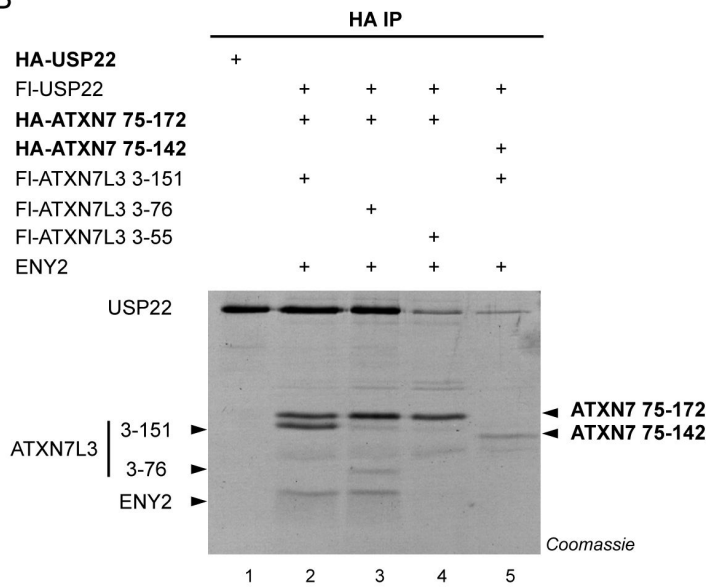
**Figure S4.**

Sequence analysis of the ATXN7 region that is involved in the formation of the DUB module. This domain is composed by a long and poorly conserved N-terminal region and by the ZnF-Sgf73 (see Figure 3). **(A)** Sequence alignment of the ZnF-Sgf73 of hATXN7 orthologs from representative species. **(B,C,D)** Sequence alignments of the entire domain for three subgroups of hATXN7 orthologs. Due to the lack of significant similarity between the N-terminal regions of these different proteins, all sequences could not be aligned together and were therefore divided in three subgroups (**B** vertebrate species, **C** insect species and **D** yeast species). Conserved residues with 90-100%, 70-80%, 50-60% identity are represented with dark grey, grey and light grey shading respectively. The polyQ stretch (Qn in a black box in panel **B**) is specific to the vertebrate orthologs of hATXN7 and is not found in the closely related vertebrate paralogs ATXN7L1 and ATXN7L2. In ySgf73, a WK motif was shown to be important for the assembly of the DUB module (black rectangle in **D**). A similar motif can be observed in the insect sequences (black rectangle in **C**) and may be involved in an interaction with Nonstop, the insect ortholog of USP22. A reminiscent motif, WxxW is also present in all the vertebrate sequences (black rectangle in **B**). Black triangles label the residues of the ZnF-Sgf73 that coordinate the zinc ion.

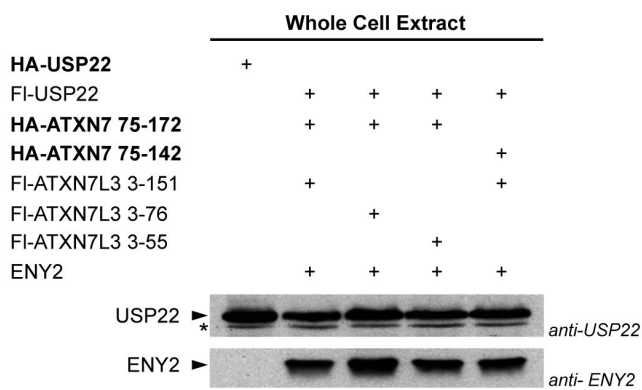
**A**



**B**

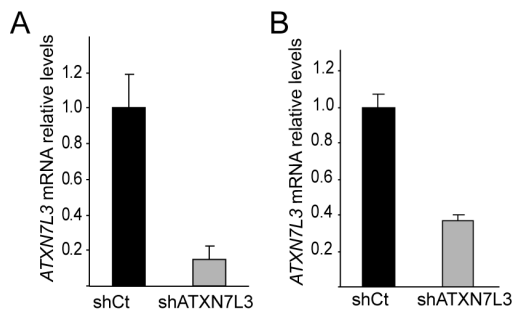


**C**



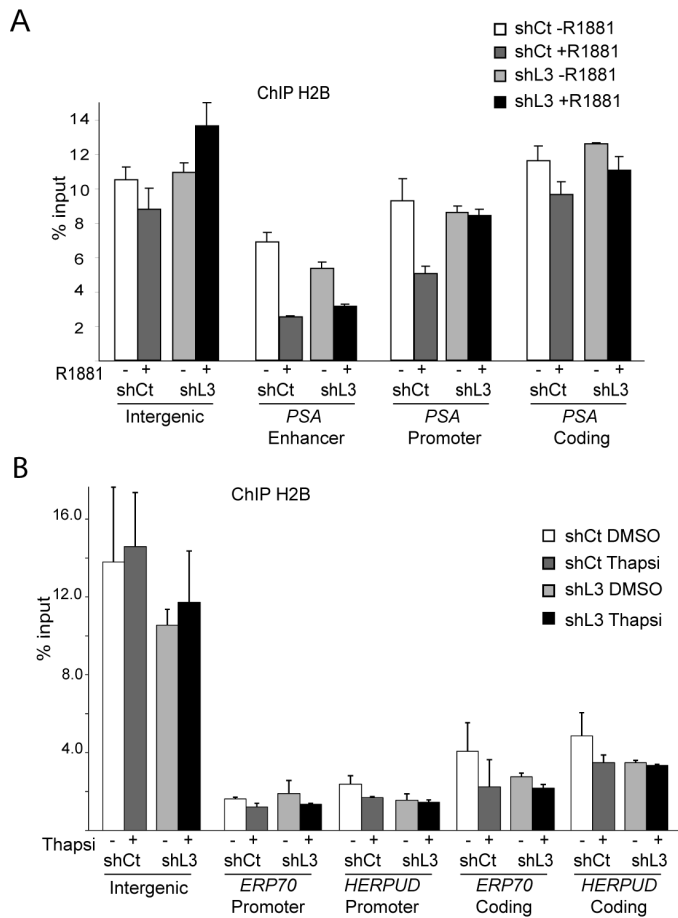
**Figure S5.**

Stoichiometry of the purified recombinant complexes as assessed by coomassie blue staining. (A,B) Coomassie blue staining of purified complexes shown in Fig3.B,C. (C) Analysis of the protein extracts used for the complex purifications presented in (B) and Fig. 3C. USP22 and ENY2 are present in similar amounts in the Sf9 cells infected by their respective baculovirus as shown by western blot. Asterisk (\*) indicates an aspecific band.



**Figure S6.**

Efficiency of *ATXN7L3* downregulation by shRNA. RNA extracted from HeLa (A) or LNCaP (B) cell lines stably expressing a shRNA against *ATXN7L3* or a control shRNA were reverse transcribed and quantified by qPCR. *ATXN7L3* mRNA levels are presented as a percentage of the mean of control cells after normalization to *CycloB* for HeLa cells or *GAPDH* for LNCaP cells. All the results are presented as mean value +/- SD of three independent experiments.



**Figure S7.**

H2B ChIP experiments on the *PSA* and *HERPUD* genes revealed that the H2B levels are not significantly changed after ATXN7L3 knock-down in LNCaP or in HeLa cells. **(A)** Control LNCaP cell lines and LNCaP cell lines expressing a shRNA against ATXN7L3 were incubated with (+) or without (-) ligand (R1881) for 210 min. Precipitated DNA was quantified by realtime qPCR using primers on enhancer/AREIII, promoter/AREI/II and coding region/exon4 of the *PSA* gene and an intergenic control region. **(B)** Control and ATXN7L3 shRNA HeLa cell lines were incubated with thapsigargin for 4 hours or with vehicle only (DMSO). Precipitated DNA was quantified by realtime qPCR using primers on *ERP70* and *HERPUD* promoters and coding regions and on an intergenic control region. **(A,B)** The values (mean +/- SD of triplicates) are expressed as percentage of the amount of

immunoprecipitated DNA normalized to the respective input DNA signal (%input).

Quantitative ChIP assays were performed for H2B. These results are representative of three independent experiments.

## Partie 6 : Discussion et Perspectives

### A. Etude genome-wide de l'activité DUB de SAGA

#### Constat et définition du projet :

Sur les quelques gènes que nous avons analysés lors de notre précédente étude (Lang et al, 2011), il apparaît que l'activité de déubiquitination (DUB) de USP22 est principalement détectée sur la région transcrite alors que l'analyse de la localisation de SAGA par CHIP ne permet d'observer le complexe qu'au niveau des promoteurs (cf. Partie Résultats). Comme je l'ai indiqué dans la partie Résultats, ce paradoxe pourrait être expliqué par une différence dans la dynamique du complexe SAGA au niveau du promoteur ou de la région transcrite d'un gène. En effet, on peut imaginer qu'au promoteur, le complexe SAGA soit stabilisé par des interactions avec des activateurs ou des modifications de la chromatine, comme H3K4me3, une marque reconnue par SGF29, et que par contre sur la région transcrite, en l'absence de ces interactions, SAGA soit présent de façon beaucoup plus transitoire. Nos précédents résultats suggèrent donc qu'une approche de CHIP classique, avec l'utilisation d'un anticorps reconnaissant une sous-unité du complexe SAGA, ne permet d'analyser que les régions où le complexe est recruté de façon stable mais ne permet pas de définir la totalité des régions où SAGA exerce son activité DUB.

Pour généraliser les observations que nous avons faites sur les gènes régulés par un stress du réticulum endoplasmique, nous voudrions donc mener des analyses sur l'ensemble du génome (genome-wide). Tout d'abord, nous souhaitons réaliser une approche de CHIP couplée à un séquençage haut débit (ChIP-seq) en utilisant notamment un anticorps dirigé contre une sous-unité du DUBm de SAGA. D'autre part, nous voulons mener une approche complémentaire pour définir tous les sites où le complexe SAGA déubiquitine l'histone H2B en comparant le niveau de H2B ubiquitiné (H2Bub) dans des cellules sauvages et des cellules dans lesquelles nous avons supprimé l'activité DUB de SAGA (lignée shATXN7L3). Pour cela, nous allons réaliser des expériences de CHIP dans ces deux lignées cellulaires avec un anticorps qui est très spécifique de H2Bub. Ce deuxième type d'approche devrait donc nous permettre de traquer de façon beaucoup plus fine le complexe SAGA en suivant son activité DUB. Nous pensons d'ailleurs que ce type d'expérience peut-être comparé à une approche

de DamID. En effet, le principe du DamID est de comparer le niveau d'une méthylation spécifique de l'ADN, différente de la méthylation des cytosines présente chez les eucaryotes, entre des cellules contrôles et des cellules exprimant une protéine fusionnée à une méthylase bactérienne (Dam méthylase). Dans notre cas, nous n'avons pas besoin d'exprimer un transgène, il nous suffit de supprimer l'activité DUB de SAGA (ce qui n'affecte pas la structure du complexe) et d'analyser le niveau de H2Bub dans des cellules contrôles et shATXN7L3.

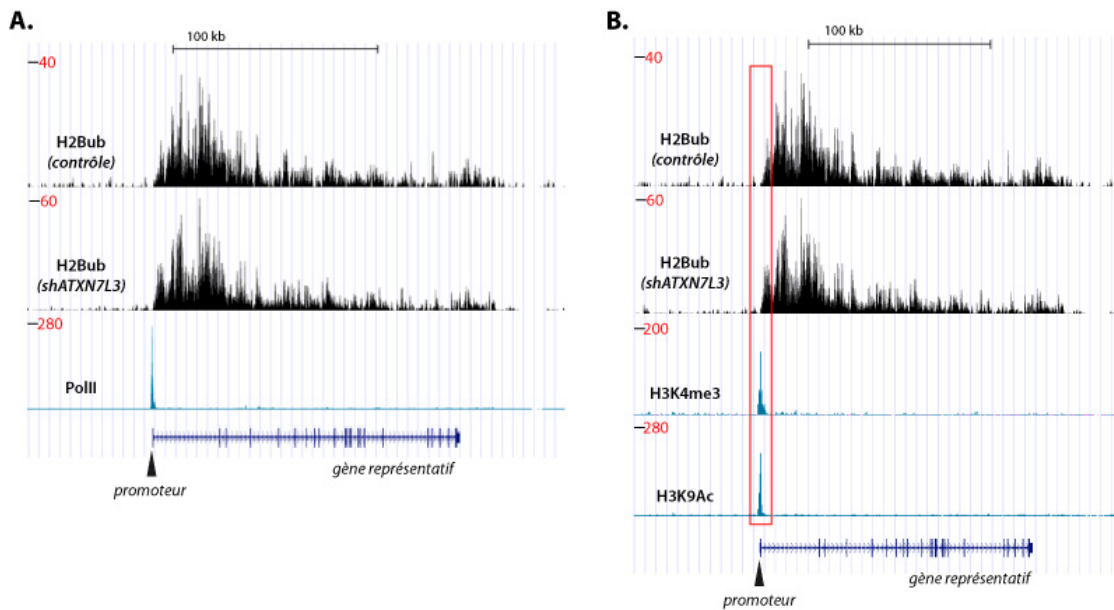
#### Résultats préliminaires :

J'ai réalisé une première série d'expériences de CHIP anti-H2Bub dans des cellules contrôles et shATXN7L3 et l'ADN ainsi immunoprécipité a été séquencé. Pour le moment, nous n'avons pu faire qu'un certain nombre de constatations, en observant au hasard de nombreuses régions génomiques, mais il nous reste encore à confirmer ces observations par des analyses bioinformatiques plus systématiques.

Tout d'abord, dans les deux cas (cellules contrôles et shATXN7L3), on ne retrouve pas de H2Bub sur les régions intergéniques, mais par contre, la marque est enrichie de manière significative sur le corps d'un très grand nombre de gènes transcrits (Figure 32 A.). Ceci est en accord avec les données de la littérature et nous permet de valider nos données. Notons également qu'à première vue, il n'y a pas de corrélation évidente entre le niveau de H2Bub observé sur la région transcrite d'un gène et le niveau d'expression de ce gène mais des analyses globales restent à mener pour apporter une conclusion définitive. Il semble donc plutôt exister une situation binaire avec présence d'H2Bub sur la région transcrite de gènes exprimés et absence de H2Bub sur les régions non transcrite. Néanmoins, ces observations sont extrêmement préliminaires et demandent à être confirmées par une analyse plus systématique.

De plus, nos résultats indiquent que le signal correspondant à H2Bub est principalement présent sur la région transcrite des gènes et qu'il est pratiquement absent au niveau du promoteur. En effet, nous observons que le signal correspondant à H2Bub apparaît en aval (3') du pic de PolII (ARN Polymérase II) ou de celui des marques d'histones H3K9Ac et H3K4me3 sur la presque totalité des gènes analysés (Figure 32).

Afin de pouvoir comparer quantitativement les profils de H2Bub dans les cellules contrôles et shATXN7L3, nous avons décidé de normaliser les résultats en définissant que le bruit de fond observé dans les régions intergéniques doit être similaire dans les deux situations. Ceci est notamment possible car ce niveau de bruit de fond est faible et



**Figure 32. SAGA exerce son activité DUB sur la région transcrite de gènes exprimés.**  
**A.** Profil d'ubiquitination de H2B (H2Bub) sur un gène représentatif dans les conditions contrôle et shATXN7L3 (suppression de l'activité DUB de SAGA). H2Bub est enrichi sur la région transcrite de ce gène, mais la marque est absente en amont et en aval du gène, et est très faible au niveau du promoteur (peu de H2Bub au niveau du pic de PolII (les données de PolII proviennent du consortium ENCODE)). Le profil de H2Bub est très similaire entre la situation contrôle et shATXN7L3 mais le niveau de H2Bub est 1,5 fois plus élevé dans la situation shATXN7L3 par rapport à la situation contrôle. **B.** H2Bub est presque absent du promoteur de ce gène représentatif. La colocalisation entre le signal de H2Bub et deux marques associées avec les promoteurs (H3K4me3 et H3K9Ac, données générées par le consortium ENCODE) est presque inexistante.

très homogène sur l'ensemble du génome et car les deux expériences de ChIP contrôles et shATXN7L3 ont été réalisées en parallèle et dans les mêmes conditions. En normalisant ainsi nos résultats, nous observons une augmentation du niveau moyen de H2Bub supérieure à 2 dans la situation shATXN7L3 pour 70% des gènes exprimés (Figure 32).

Cette première analyse succincte de nos résultats nous permet tout de même de conclure que l'activité DUB de USP22 agit sur la région transcrite de presque tous les gènes exprimés. Une approche genome-wide réalisée au sein de notre équipe a récemment montré que SAGA n'était recruté de façon stable qu'au niveau d'un petit nombre de promoteurs (Krebs et al, 2011), mais nos nouveaux résultats indiquent que l'activité DUB de SAGA agit de façon beaucoup plus large. Enfin, ces observations indiquent que le co-activateur de la transcription SAGA ne modifie pas seulement la chromatine au niveau des promoteurs de ses gènes cibles mais également au niveau de la région transcrite d'un très grand nombre de gènes.



Perspectives :

Bien entendu, il nous reste encore beaucoup de travail pour terminer cette étude et j'ai choisi d'indiquer un certain nombre des questions que nous voudrions adresser.

Existence d'un DUBm libre et indépendant de SAGA ?

Vu l'apparente contradiction entre d'un côté, un complexe SAGA recruté de façon stable sur un nombre limité de promoteurs (Krebs et al, 2011) et de l'autre, une activité de déubiquitination détectée sur la presque totalité des gènes, on peut se poser la question de savoir si un DUBm libre contenant ATXN7, ATXN7L3, ENY2 et USP22 pourrait agir indépendamment du complexe SAGA chez l'homme.

Durant ma thèse, j'ai obtenu des données biochimiques qui suggèrent qu'un DUB libre contenant au moins ATXN7L3, USP22 et ENY2 pourrait exister. Par exemple, si on compare par western blot une immuno-purification de SAGA réalisée à partir d'une sous-unité structurale ou d'une sous-unité du module d'acétylation et une autre immuno-purification de SAGA réalisée à partir du DUBm, pour un signal de même intensité pour GCN5 ou pour TRRAP, le signal correspondant à USP22 ou à ATXN7L3 est plus important dans la purification réalisée à partir du DUBm. Après avoir fait cette observation, nous avons analysé, par une colonne d'exclusion – inclusion, l'élution d'une immuno-précipitation réalisée avec un anticorps dirigé contre USP22. L'analyse des fractions obtenues à la sortie de la colonne nous a permis de mettre en évidence un pic correspondant au complexe SAGA et un autre pic, de masse apparente beaucoup plus faible, contenant au moins USP22 et ATXN7L3.

Nous restons tout de même prudent concernant l'existence d'un DUBm libre et actif dans la cellule. Tout d'abord, les tests enzymatiques que j'ai réalisés *in vitro* ont montré que pour être pleinement activée, USP22 avait besoin d'interagir non seulement avec ENY2 et ATXN7L3 mais également ATXN7. Or, dans le « DUBm libre » que nous avons observé en sortie de la colonne d'exclusion – inclusion, nous n'avons jamais pu détecter ATXN7 par western blot. Il est donc possible que ce module libre ne contienne que ATXN7L3, ENY2 et USP22. Dans ce cas, ce module ne serait pas catalytiquement actif, mais ce serait davantage un module de pré-assemblage qui serait prêt à interagir avec ATXN7 pour être activé et être intégré à SAGA.

Pour clarifier ce point, nous voudrions maintenant pouvoir déterminer si le DUBm libre peut être recruté sur la chromatine *in vivo*, de façon indépendante au complexe SAGA. Pour cela, nous voulons analyser la localisation sur tout le génome

d'une sous-unité du DUBm. Nous envisageons d'utiliser un anticorps polyclonal anti-ATXN7L3 pour réaliser un ChIP-seq dans des cellules sauvages, et pour vérifier la spécificité de cet anticorps, nous voulons réaliser la même approche dans des cellules shATXN7L3 qui sont dépourvues de la protéine ATXN7L3. Une autre alternative serait de générer une lignée cellulaire qui exprimerait de façon stable une version étiquetée de USP22 fusionnée à trois épitopes HA. Dans un second temps, nous comparerons ces résultats de ChIP avec ceux obtenus pour une autre sous-unité de SAGA, présente en dehors du DUBm. Si le nombre de sites où le DUBm est observé est beaucoup plus important que celui où SAGA est détecté, alors nous pourrions penser que le DUBm libre a une importance fonctionnelle *in vivo* et que ce module libre participe avec SAGA à la déubiquitination de H2B.

#### Quelle est la fonction jouée par le DUBm au sein de SAGA ?

Dans la partie de l'introduction consacrée au complexe SAGA, nous avons vu que chez la levure et dans les conditions standard, la perte de l'activité DUB de SAGA n'affecte que de façon très modeste la transcription des gènes. A l'inverse, la perte de l'activité d'acétylation ou la délétion de sous-unités du module de régulation de TBP provoque la dérégulation de l'expression d'un nombre important de gènes. Nous avons aussi vu que la co-délétion de Ubp8 et de Gcn5 conduisait à une aggravation des phénotypes observés dans une souche *gcn5Δ*. L'ensemble de ces observations suggère donc que contrairement au module d'acétylation, la fonction du DUBm de SAGA n'est pas reliée directement aux mécanismes d'initiation de la transcription, mais que l'activité DUB de SAGA régule d'autres processus dans la cellule que nous devons identifier.

En accord avec cette idée, on peut noter que mes résultats de ChIP anti-H2Bub dans les cellules contrôles et shATXN7L3 mettent en évidence une absence de colocalisation entre le signal de H2Bub (présent sur la région transcrite des gènes mais absent au niveau du promoteur) et le signal de H3K9 acétylé (centré sur le promoteur) (Figure 32 B.). Or, la seconde marque est mise en place par l'activité d'acétylation de SAGA alors que H2Bub est déubiquitinée par USP22. La répartition de ces deux marques suggère donc que les deux activités de SAGA ne sont pas exercées par le complexe sur les mêmes nucléosomes. On peut alors imaginer que *in vivo*, plusieurs conformations du complexe SAGA pourraient exister et qu'en fonction de l'environnement chromatinien SAGA pourrait davantage acétyler H3 ou déubiquitiner H2B. On peut supposer que l'acétylation par GCN5 pourrait nécessiter la présence d'autres marques

chromatiniennes associées avec les promoteurs ainsi qu'une stabilisation du complexe sur la chromatine par des interactions avec des activateurs, alors que la déubiquitination par USP22 pourrait simplement requérir la présence de H2Bub. Suivant la nature des modifications de la chromatine avoisinante, on peut donc émettre l'hypothèse que le complexe SAGA peut exercer de façon spécifique l'une ou l'autre de ses activités enzymatiques afin de réguler tels ou tels processus biologiques.

La déubiquitination de H2B dépend-elle de la transcription ?

Dans la partie de l'introduction consacrée à l'ubiquitination de H2B, nous avons vu que si l'on bloque la transcription sur des cellules humaines sauvages, on provoque une perte totale de H2Bub en moins d'une heure. Cette observation ainsi que d'autres résultats démontre que l'ubiquitination de H2B est un processus co-transcriptionnel et qu'en l'absence de transcription, H2B ne peut pas être ubiquitiné. La même question pourrait être posée pour ce qui est de la déubiquitination. En d'autres termes, on pourrait se demander si SAGA « suit » la PolII sur le corps des gènes ou si une partie des complexes SAGA diffuse librement sur la chromatine et déubiquitine H2B lorsque celui-ci est ubiquitiné afin de maintenir un niveau relativement bas de H2Bub sur le corps des gènes transcrits.

Pour déterminer si la perte rapide de H2Bub après un blocage de la transcription est due à l'activité DUB de SAGA, il serait sans doute possible de suivre la cinétique de disparition de H2Bub, dans des cellules sauvages et des cellules shATXN7L3 après un tel blocage de la transcription. Pour cela il faudrait réaliser des extractions acides des histones de ces deux lignées cellulaires à différents moments avant et après avoir bloqué la transcription. Le niveau global de H2Bub serait ensuite analysé par western blot. Après le blocage de la transcription, si une diminution du signal de H2Bub était d'abord visible dans les cellules contrôles sans qu'elle ne soit détectable dans les cellules shATXN7L3, cela signifierait que cette baisse est due à l'activité DUB de SAGA. On pourrait donc conclure que SAGA peut déubiquitiner H2B en l'absence de transcription. Dans ce cas, cela suggérerait qu'une fonction importante du DUBm de SAGA serait de participer à l'homéostasie de l'ubiquitine en maintenant un équilibre adéquat entre niveau de H2Bub et niveau d'ubiquitine libre.

### Déubiquitination de H2B et régulation de la quantité d'ubiquitine libre

L'ubiquitine est une molécule impliquée dans la régulation de très nombreux processus cellulaires allant de la transcription, à la dégradation des protéines en passant par la réparation de l'ADN (cf. Partie sur les USP). Une mauvaise régulation de la quantité d'ubiquitine libre dans la cellule peut donc affecter diverses voies fonctionnelles. De façon intéressante, notons qu'un blocage du protéasome provoque une perte de l'ubiquitination de H2B, ce qui illustre l'importance de la régulation de ce niveau d'ubiquitine libre.

Dans notre cas, il faudrait observer si le traitement de cellules humaines sauvages avec un ARN interférent dirigé contre ATXN7L3 cause une diminution de la quantité d'ubiquitine libre de la cellule. Si c'était le cas, il serait alors intéressant de combiner la perte de l'activité DUB de SAGA à un deuxième stress ayant aussi pour effet de diminuer le niveau d'ubiquitine libre. On pourrait par exemple éliminer par ARN interférence USP14, une déubiquitinase qui est associée au protéasome et dont la perte de fonction provoque une baisse du niveau d'ubiquitine libre. Pour évaluer les conséquences de ce double stress, nous pourrions alors mesurer le temps de duplication de cellules traitées par les deux ARN interférents contre ATXN7L3 et USP14 et le comparer à celui de cellules sauvages.

### Analyse de la distribution de H2Bub chez la levure

Comme nous l'avons déjà vu, chez l'homme, nous disposons d'un anticorps très spécifique contre H2Bub qui donne notamment d'excellents résultats en CHIP. Cependant, cet anticorps ne reconnaît pas H2Bub chez la levure *S. cerevisiae* et pour le moment, aucun anticorps spécifique de cette marque n'existe dans cet organisme, ce qui empêche d'analyser la distribution de H2Bub sur le génome de la levure.

Nous avons imaginé une stratégie qui devrait nous permettre d'utiliser l'excellent anticorps anti H2Bub chez la levure. Dans cet organisme, nous voulons donc analyser le profil de distribution de H2Bub et déterminer où SAGA déubiquitine H2B en comparant par CHIP le niveau de H2Bub dans des levures sauvages et des levures *sgf11Δ* ou *ubp8Δ* (dans lesquelles l'activité DUB de SAGA est perdue). Nous pourrions aussi imaginer de compléter ces résultats avec une analyse de la localisation du complexe SAGA sur le génome de la levure en réalisant des CHIP à partir de plusieurs sous-unités du complexe.

## **B. Analyses structure fonction du DUBm**

### **I. Analyses structure fonction des domaines SCA7**

Mon travail de thèse a donc permis de caractériser les deux domaines SCA7 de ATXN7 et ATXN7L3, deux sous-unités du module de déubiquitination (DUBm) de SAGA chez l'homme. Nous avons montré que même si ces deux domaines partagent un cœur structural commun et spécifique, il existe des divergences structurales qui distinguent les domaines SCA7 de ATXN7 et de ATXN7L3 et qui doivent, au moins en partie, expliquer les spécificités fonctionnelles de ces deux domaines. Je voudrais maintenant développer trois points visant à mieux définir les propriétés fonctionnelles de ces domaines SCA7. Dans un premier temps, nous parlerons de la zone d'interaction du domaine SCA7 de ATXN7 sur un nucléosome, puis je discuterai ensuite brièvement le fait que le domaine SCA7 de ATXN7 puisse interagir avec le complexe TREX2 et enfin, je proposerai une démarche expérimentale afin de définir la fonction du domaine SCA7 de ATXN7L3.

#### Définition de la zone d'interaction du domaine SCA7 de ATXN7 sur un nucléosome

Nous avons montré que le domaine SCA7 de ATXN7 mais pas celui de ATXN7L3 interagit avec un nucléosome via le dimère H2A-H2B. Pour davantage caractériser l'interaction entre le domaine SCA7 de ATXN7 et le nucléosome, il serait intéressant de définir plus précisément la région du nucléosome sur laquelle se lie ce domaine SCA7. Pour le moment, nous avons montré par des expériences de GST pull down que le domaine SCA7 de ATXN7 est capable de lier le dimère d'histones recombinants H2A-H2B et que cette interaction ne nécessite pas la queue N-terminale de ces histones. Ceci suggère donc que le domaine SCA7 de ATXN7 interagit avec la face latérale du nucléosome et on peut émettre l'hypothèse que cette interaction ait lieu au niveau d'une région du dimère H2A-H2B appelée « patch acide ». En effet, cette région de la face latérale du nucléosome est chargée négativement et est connue pour être la zone de liaison de plusieurs partenaires du nucléosome comme le peptide LANA (latency-associated nuclear antigen) qui provient du virus de l'herpès et qui est associé au sarcome de Kaposi. Etant donné que le domaine SCA7 de ATXN7 est chargé très positivement, il ne serait pas surprenant qu'il puisse également lier le nucléosome au niveau de ce patch acide.

Pour démontrer cette hypothèse, nous pourrions étudier si l'ajout de concentrations croissantes de peptides LANA dans des expériences de GST pull down peut inhiber par compétition pour le patch acide l'interaction entre le domaine SCA7 et le nucléosome. D'autre part, un nouveau type d'analyse du nucléosome par RMN, développé par Hugo van Ingen dans l'équipe de Lewis Kay, pourrait également nous permettre de répondre à cette question (Kato et al, 2011). Sans entrer dans les détails techniques, il est maintenant possible de « visualiser » dans un spectromètre de RMN les isoleucines, leucines et valines d'un des histones après avoir marqués ces résidus lors de la synthèse de la protéine recombinante. Cela est une grande avancée car le nucléosome est un très gros objet qui ne pouvait pas facilement être étudié par les techniques standard de RMN. Dans notre cas, on pourrait donc imaginer d'utiliser des nucléosomes marqués puis de les titrer avec des quantités croissantes de domaines SCA7 non marqués pour déterminer quels résidus des différents histones sont affectés par la liaison du domaine SCA7 sur le nucléosome ce qui nous renseignerait sur le lieu exact de l'interaction entre les deux partenaires.

Comme nous venons de le voir, nous supposons que l'interaction entre le domaine SCA7 de ATXN7 et le nucléosome pourrait avoir lieu sur la face latérale du nucléosome, non loin du lieu où celui-ci est ubiquitiné sur H2B. Cependant, nous ne pensons pas que le domaine SCA7 de ATXN7 puisse lier directement l'ubiquitine ou que l'affinité du domaine SCA7 pour le nucléosome soit augmentée si celui-ci est ubiquitiné. Aucune observation allant dans ce sens n'a été faite dans nos expériences *in vitro*. Par contre, notre modèle est que l'interaction entre la face latérale du nucléosome et le domaine SCA7 de ATXN7 pourrait permettre de positionner le DUBm de façon optimale à proximité de son substrat. Il est certain qu'à très long terme, l'obtention d'une structure d'un DUBm comprenant le domaine SCA7 avec un nucléosome ubiquitiné permettrait de comprendre quelle est la fonction exacte jouée par le domaine SCA7 de ATXN7 au sein du DUBm.

#### Lien entre le domaine SCA7 de Sgf73/ATXN7 et le complexe TREX2 ?

Dans une publication récente (Köhler et al, 2008), Kohler et al ont proposé que le domaine SCA7 de Sgf73 pourrait être impliqué dans une interaction entre le complexe SAGA et le complexe TREX2, un complexe localisé au niveau des pores nucléaires et qui est impliqué dans l'export des ARNm. Une telle observation semble contradictoire avec

nos résultats. Il serait donc intéressant de tester *in vitro* si le domaine SCA7 de Sgf73 ou celui d'ATXN7 peut interagir avec le complexe TREX2.

### Vers l'identification de la fonction du domaine SCA7 de ATXN7L3

Pour le moment, nous n'avons aucun indice sur la fonction du domaine SCA7 de ATXN7L3, et comme ce domaine est absent chez Sgf11, il n'est même pas possible d'utiliser des mutants de levure pour essayer de caractériser sa fonction, ce qui fait que je ne peux proposer que des hypothèses sur la fonction putative de ce domaine. Soit ce domaine SCA7 interagit de façon transitoire avec des protéines en dehors de SAGA, soit il interagit avec d'autres sous-unités du complexe SAGA. Dans le premier cas, on peut sans doute exclure le nucléosome ainsi que l'ubiquitine comme l'ont montré mes expériences *in vitro* mais on pourrait imaginer une interaction avec un autre co-activateur de la transcription ou avec un complexe faisant partie de la machinerie transcriptionnelle de base. Concernant la seconde hypothèse, on pourrait penser que l'interaction du domaine SCA7 de ATXN7L3 avec d'autres sous-unités du complexe SAGA permette de renforcer l'ancrage du DUBm au sein du complexe. Une telle interaction permettrait en effet d'apporter un niveau d'organisation structurale supplémentaire par rapport à la situation de la levure en limitant les degrés de liberté du DUBm par rapport au reste du complexe SAGA.

Au début de ma thèse, j'ai essayé de mettre en évidence une éventuelle interaction entre le domaine SCA7 de ATXN7L3 et d'autres sous-unités de SAGA. En utilisant des protéines recombinantes et surexprimées, aucune interaction significative n'avait pu être mise en évidence. Mais je n'avais pas testé toutes les sous-unités de SAGA et de plus, je m'étais limité à rechercher si le domaine SCA7 de ATXN7L3 pouvait former un dimère avec une autre sous-unité de SAGA. Or, on ne peut pas exclure que deux sous-unités ou plus soient requises pour stabiliser une possible interaction avec le domaine SCA7 de ATXN7L3 et c'est d'ailleurs ce type d'interactions que j'ai observé en analysant l'architecture du DUBm.

Je pense qu'une façon simple pour mieux adresser la question de la fonction de ce domaine SCA7 serait de générer une lignée cellulaire exprimant de façon stable le domaine SCA7 de ATXN7L3 fusionné aux étiquettes HA et Flag. Nous pourrions ainsi tenter d'identifier les éventuels partenaires de ce domaine en analysant par spectrométrie de masse la nature des protéines co-purifiées par le domaine. Pour gagner en spécificité et pouvoir comparer différentes purifications, il serait intéressant de

générer trois lignées cellulaires qui exprimeraient différentes versions étiquetées de ATXN7L3 : la protéine entière (composée par un domaine N-terminal et le domaine SCA7), le domaine N-terminal de ATXN7L3 (ce domaine se trouve au cœur du DUBm et correspond à la région homologue à Sgf11) et le domaine SCA7 seul. Enfin, l'idéal serait d'exprimer dans ces diverses lignées un ARN interférent contre l'ARNm endogène de ATXN7L3 mais qui ne ciblerait pas les transgènes afin de ne permettre que la synthèse de la forme recombinante de ATXN7L3. Pour ne pas risquer de manquer des partenaires d'interactions très transitoires, il faudrait également laver les expériences de co-IP avec différentes stringences.

Si le domaine SCA7 de ATXN7L3 interagit avec des partenaires en dehors de SAGA, ces éventuelles protéines devraient alors être détectées dans les purifications réalisées à partir du domaine SCA7 seul et de la protéine entière mais devraient être absentes de la purification réalisée avec le domaine N-terminal de ATXN7L3 ou dans une purification contrôle. Par contre, si le domaine SCA7 de ATXN7L3 interagit avec des sous-unités de SAGA, les mêmes protéines seront purifiées par les trois formes de la protéine. Dans ce cas, il faudrait déterminer qui sont le ou les sous-unités de SAGA qui interagissent avec ce domaine SCA7. Je crois qu'une hypothèse intéressante serait que le domaine SCA7 de ATXN7L3 puisse interagir avec des sous-unités du module d'acétylation ce qui créerait un lien physique entre les deux modules enzymatiques de SAGA.

## **II. Régulation enzymatique de l'activité DUB de SAGA par les sous-unités du DUBm**

### Comparaison du DUBm chez la levure et chez l'homme

Bien qu'ayant été identifié plus tardivement, le DUBm est le module de SAGA qui a été le mieux caractérisé concernant les aspects structure - fonction. Chez la levure, il a en effet été montré que les interactions entre Ubp8 et les autres sous-unités du DUBm permettaient de stimuler l'activité de l'enzyme. De plus, les deux structures cristallographiques du DUBm de levure publiées en 2010 expliquent l'importance de ce réseau complexe d'interactions pour assurer l'activation de Ubp8. En accord avec ces résultats, mon travail de thèse a permis de montrer que ce mécanisme de régulation allostérique était conservé chez l'homme. Dans les deux espèces, il apparaît en effet clair que le doigt de zinc de Sgf11/ATXN7L3 joue un rôle direct dans l'activation de l'enzyme.



De plus, chez la levure et chez l'homme, le domaine N-terminal de Sgf73/ATXN7 est requis pour la pleine activation de l'enzyme. On peut toutefois noter une différence importante entre les deux espèces au niveau de la fonction du domaine ZnF-UBP porté par Ubp8 ou USP22. Chez la levure, une forme tronquée de Ubp8 ne contenant que ce domaine est suffisante pour interagir avec les trois autres sous-unités du module et former un complexe très stable. Dans cet organisme, il semble donc que la formation d'un lobe d'assemblage autour du domaine ZnF-UBP serve à positionner le ZnF-Sgf11 à proximité du site catalytique de Ubp8. Chez l'homme, par contre, le domaine ZnF-UBP de USP22 seul n'est pas suffisant pour permettre la formation d'un complexe stable. L'enzyme USP22 entière est nécessaire pour obtenir une interaction forte avec ATXN7, ATXN7L3 et ENY2. Ceci indique donc qu'au cours de l'évolution, la séparation fonctionnelle entre lobe d'assemblage et lobe catalytique s'est amoindrie. Le fait que des éléments structuraux du domaine catalytique de USP22 soient devenus nécessaires à la formation d'un DUBm stable suggère donc que chez l'homme, le positionnement du doigt de zinc de ATXN7L3 à proximité du site catalytique de USP22 est encore d'avantage contrôlé que chez la levure. La résolution de la structure atomique du module humain permettra sans doute d'apporter une réponse claire sur ce sujet.

#### Mécanisme d'activation allostérique de l'enzyme

De très nombreux éléments sont donc connus au sujet de l'organisation structurale du DUBm et de ses conséquences fonctionnelles. Cependant, certaines questions restent encore sans réponse. Il serait par exemple intéressant de disséquer d'avantage le mécanisme d'activation allostérique en déterminant notamment quelles régions de Ubp8 sont affectées par la liaison des autres sous-unités du module. Par exemple, on peut imaginer que la formation du DUBm stabilise la poche de liaison de l'ubiquitine du domaine catalytique de l'enzyme ou bien qu'elle permette de positionner de façon optimale les résidus de la triade catalytique. Pour répondre à cette question, il faudrait pouvoir obtenir une structure cristallographique de l'enzyme inactive pour la comparer avec celle de la forme active qui est connue. Cela pourrait être Ubp8 ou USP22 seule, mais il serait également intéressant de déterminer le rôle exact joué par le doigt de zinc de Sgf11 pour activer l'enzyme en essayant de résoudre la structure d'un module inactif délété de ce seul domaine.

Reconnaissance spécifique du substrat : rôle du doigt de zinc de Sgf11/ATXN7L3

Une autre question qui reste sans réponse claire concerne la reconnaissance spécifique du nucléosome ubiquitiné sur H2B par le DUBm. Chez la levure, cette marque est portée par la lysine 123, au niveau de l'extrémité C-terminale de H2B. L'ubiquitine se situe donc sur la face latérale du nucléosome. De plus, la lysine 123 de H2B est insérée au sein d'une hélice  $\alpha$  ce qui doit réduire la flexibilité de l'ubiquitine par rapport au nucléosome. Pour déubiquitiner H2B, le DUBm doit donc vraisemblablement se trouver en contact non seulement avec l'ubiquitine mais également avec un ou plusieurs éléments structuraux du nucléosome. Les auteurs des deux études qui ont déterminé la structure du DUBm de levure proposent que le doigt de zinc de Sgf11 pourrait jouer un rôle d'adaptateur entre le nucléosome et le DUBm (Kohler et al, 2010; Samara et al, 2010). En effet, en plus de se trouver au contact du site actif de Ubp8 et d'être requis pour l'activité de l'enzyme, ce doigt de zinc possède plusieurs arginines strictement conservées au cours de l'évolution et qui sont exposées vers le solvant dans les structures cristallographiques (Kohler et al, 2010; Samara et al, 2010) (cf. Partie sur les USP). Un modèle proposé par une des équipes suggère que le doigt de zinc de Sgf11 reconnaisse le nucléosome en réalisant des interactions électrostatiques avec les groupements phosphate de l'ADN (Samara et al, 2010). L'autre équipe a montré que chez la levure, des mutations de certaines des arginines du doigt de zinc de Sgf11 provoquent une forte augmentation du niveau de H2B ubiquitiné chez ces souches mutantes, similaire à celle observée en l'absence de Ubp8 (Kohler et al, 2010) (cf. Partie sur les USP). Nous avons également obtenu des résultats montrant la nécessité du doigt de zinc de ATXN7L3 pour déubiquitiner un nucléosome *in vitro* alors que l'absence de ce doigt de zinc dans un module recombinant provoque simplement une diminution partielle de la capacité de liaison du substrat suicide Ub-VS sur USP22. Il serait donc intéressant de démontrer expérimentalement que le doigt de zinc de Sgf11 interagit avec l'ADN du nucléosome ainsi que de préciser quels résidus du domaine sont impliqués dans cette interaction. Ce travail est mené en collaboration avec l'équipe de Bruno Kieffer à l'IGBMC. En utilisant une approche de RMN, il est en effet possible de mettre en évidence des interactions de faible intensité et très dynamiques comme doit l'être celle entre le ce doigt de zinc et le nucléosome. Des résultats préliminaires indiquent que ce domaine peut interagir avec des fragments d'ADN libres ainsi qu'avec des nucléosomes purifiés à partir de cellules HeLa. De plus, les analyses de RMN

montrent que certaines des arginines conservées ainsi que d'autres résidus du doigt de zinc de Sgf11 pourraient être impliqués dans cette interaction (Communication de l'équipe de Bruno Kieffer). Il sera donc intéressant de générer des souches de levures dans lesquels ces résidus du doigt de zinc de Sgf11 sont mutés et d'observer l'effet de ces mutations sur le niveau global H2B ubiquitiné. En conclusion de cette partie, je voudrais souligner que le doigt de zinc de Sgf11 peut sans doute être considéré comme partie intégrante du domaine catalytique de Ubp8. En effet, sur le plan structural il est situé dans le prolongement du domaine catalytique de Ubp8 et sur le plan fonctionnel, il est non seulement requis pour l'activité de l'enzyme mais il lui permet en plus de reconnaître son substrat *in vivo*.

#### Définition structurale du DUBm

Pour finir, je voudrais proposer une définition du module de déubiquitination de SAGA chez l'homme. Bien entendu, USP22 se situe au cœur de ce module. Il faut y ajouter ENY2, le domaine N-terminal de ATXN7L3 (la région homologue à Sgf11) ainsi que le domaine N-terminal de ATXN7. L'ensemble de ces domaines correspond donc au module qui a été cristallisé chez la levure. Vu les résultats que j'ai obtenu durant ma thèse, je pense que l'on peut intégrer le domaine SCA7 de ATXN7 au sein du DUBm. Enfin, je propose que le domaine C-terminal de ATXN7 et peut être le domaine SCA7 de ATXN7L3 n'appartiennent pas au DUBm mais fassent plutôt le lien avec d'autres modules de SAGA qui devront encore être identifiés.

# **Annexes**

## Références

Allard S, Utlely RT, Savard J, Clarke A, Grant P, Brandl CJ, Pillus L, Workman JL, Cote J (1999) NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO J* **18**(18): 5108-5119

Amerik A, Swaminathan S, Krantz BA, Wilkinson KD, Hochstrasser M (1997) In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J* **16**(16): 4826-4838

Anderson C, Crimmins S, Wilson JA, Korbel GA, Ploegh HL, Wilson SM (2005) Loss of Usp14 results in reduced levels of ubiquitin in ataxia mice. *J Neurochem* **95**(3): 724-731

Atanassov BS, Evrard YA, Multani AS, Zhang Z, Tora L, Devys D, Chang S, Dent SY (2009) Gcn5 and SAGA regulate shelterin protein turnover and telomere maintenance. *Mol Cell* **35**(3): 352-364

Atanassov BS, Koutelou E, Dent SY (2011) The role of deubiquitinating enzymes in chromatin regulation. *FEBS Lett* **585**(13): 2016-2023

Avvakumov GV, Walker JR, Xue S, Allali-Hassani A, Asinas A, Nair UB, Fang X, Zuo X, Wang YX, Wilkinson KD, Dhe-Paganon S (2012) Two ZnF-UBP domains in isopeptidase T (USP5). *Biochemistry*

Avvakumov GV, Walker JR, Xue S, Finerty PJ, Jr., Mackenzie F, Newman EM, Dhe-Paganon S (2006) Amino-terminal dimerization, NRDP1-rhodanese interaction, and inhibited catalytic domain conformation of the ubiquitin-specific protease 8 (USP8). *J Biol Chem* **281**(49): 38061-38070

Babour A, Dargemont C, Stutz F (2012) Ubiquitin and assembly of export competent mRNP. *Biochim Biophys Acta*

Barlev NA, Candau R, Wang L, Darpino P, Silverman N, Berger SL (1995) Characterization of physical interactions of the putative transcriptional adaptor, ADA2, with acidic activation domains and TATA-binding protein. *J Biol Chem* **270**(33): 19337-19344

Behrends C, Harper JW (2011) Constructing and decoding unconventional ubiquitin chains. *Nat Struct Mol Biol* **18**(5): 520-528

Belgareh-Touze N, Leon S, Erpapazoglou Z, Stawiecka-Mirota M, Urban-Grimal D, Haguener-Tsapis R (2008) Versatile role of the yeast ubiquitin ligase Rsp5p in intracellular trafficking. *Biochem Soc Trans* **36**(Pt 5): 791-796

Berger SL, Cress WD, Cress A, Triezenberg SJ, Guarente L (1990) Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: evidence for transcriptional adaptors. *Cell* **61**(7): 1199-1208

Berger SL, Pina B, Silverman N, Marcus GA, Agapite J, Regier JL, Triezenberg SJ, Guarente L (1992) Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* **70**(2): 251-265

Bhaumik S, Green M (2001) SAGA is an essential in vivo target of the yeast acidic activator Gal4p. *Genes & development* **15**(15): 1935-1980

Bhaumik SR, Raha T, Aiello DP, Green MR (2004) In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev* **18**(3): 333-343

Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Barsyte D, Li J, Wu M, Zhou BO, Fleharty BE, Paulson A, Allali-Hassani A, Zhou JQ, Mer G, Grant PA, Workman JL, Zang J, Min J (2011) Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J* **30**(14): 2829-2842

Borodovsky A, Kessler BM, Casagrande R, Overkleeft HS, Wilkinson KD, Ploegh HL (2001) A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. *EMBO J* **20**(18): 5187-5196

Brand M, Leurent C, Mallouh V, Tora L, Schultz P (1999a) Three-dimensional structures of the TAFII-containing complexes TFIID and TFIIIC. *Science* **286**(5447): 2151-2153

Brand M, Moggs JG, Oulad-Abdelghani M, Lejeune F, Dilworth FJ, Stevenin J, Almouzni G, Tora L (2001) UV-damaged DNA-binding protein in the TFIIIC complex links DNA damage recognition to nucleosome acetylation. *EMBO J* **20**(12): 3187-3196

Brand M, Yamamoto K, Staub A, Tora L (1999b) Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. *J Biol Chem* **274**(26): 18285-18289

Bremm A, Komander D (2011) Emerging roles for Lys11-linked polyubiquitin in cellular regulation. *Trends Biochem Sci* **36**(7): 355-363

Brown CE, Howe L, Sousa K, Alley SC, Carrozza MJ, Tan S, Workman JL (2001) Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science* **292**(5525): 2333-2337

Brownell JE, Allis CD (1995) An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. *Proc Natl Acad Sci U S A* **92**(14): 6364-6368

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD (1996) *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**(6): 843-851

Bryant GO, Ptashne M (2003) Independent recruitment in vivo by Gal4 of two complexes required for transcription. *Mol Cell* **11**(5): 1301-1309

- Cabal GG, Genovesio A, Rodriguez-Navarro S, Zimmer C, Gadal O, Lesne A, Buc H, Feuerbach-Fournier F, Olivo-Marin JC, Hurt EC, Nehrbass U (2006) SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature* **441**(7094): 770-773
- Cai L, Sutter B, Li B, Tu B (2011) Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Molecular cell* **42**(4): 426-463
- Chernova TA, Allen KD, Wesoloski LM, Shanks JR, Chernoff YO, Wilkinson KD (2003) Pleiotropic effects of Ubp6 loss on drug sensitivities and yeast prion are due to depletion of the free ubiquitin pool. *J Biol Chem* **278**(52): 52102-52115
- Choukrallah MA, Kobi D, Martianov I, Pijnappel WW, Mischerikow N, Ye T, Heck AJ, Timmers HT, Davidson I (2011) Interconversion between active and inactive TATA-binding protein transcription complexes in the mouse genome. *Nucleic Acids Res*
- Clarke AS, Lowell JE, Jacobson SJ, Pillus L (1999) Esa1p is an essential histone acetyltransferase required for cell cycle progression. *Mol Cell Biol* **19**(4): 2515-2526
- Daniel J, Torok M, Sun Z-W, Schieltz D, Allis C, Yates J, Grant P (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *The Journal of biological chemistry* **279**(3): 1867-1938
- Dikic I, Wakatsuki S, Walters KJ (2009) Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol* **10**(10): 659-671
- Doyon Y, Cote J (2004) The highly conserved and multifunctional NuA4 HAT complex. *Curr Opin Genet Dev* **14**(2): 147-154
- Dudley AM, Rougeulle C, Winston F (1999) The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step in vivo. *Genes Dev* **13**(22): 2940-2945
- Dye BT, Schulman BA (2007) Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins. *Annu Rev Biophys Biomol Struct* **36**: 131-150
- Eisenmann DM, Arndt KM, Ricupero SL, Rooney JW, Winston F (1992) SPT3 interacts with TFIID to allow normal transcription in *Saccharomyces cerevisiae*. *Genes Dev* **6**(7): 1319-1331
- Eisenmann DM, Chapon C, Roberts SM, Dollard C, Winston F (1994) The *Saccharomyces cerevisiae* SPT8 gene encodes a very acidic protein that is functionally related to SPT3 and TATA-binding protein. *Genetics* **137**(3): 647-657
- Eisenmann DM, Dollard C, Winston F (1989) SPT15, the gene encoding the yeast TATA binding factor TFIID, is required for normal transcription initiation in vivo. *Cell* **58**(6): 1183-1191

Ellisdon AM, Jani D, Kohler A, Hurt E, Stewart M (2010) Structural basis for the interaction between yeast Spt-Ada-Gcn5 acetyltransferase (SAGA) complex components Sgf11 and Sus1. *J Biol Chem* **285**(6): 3850-3856

Faesen AC, Dirac AM, Shanmugham A, Ovaa H, Perrakis A, Sixma TK (2011) Mechanism of USP7/HAUSP activation by its C-terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Mol Cell* **44**(1): 147-159

Fernandez-Montalvan A, Bouwmeester T, Joberty G, Mader R, Mahnke M, Pierrat B, Schlaeppli JM, Worpenberg S, Gerhartz B (2007) Biochemical characterization of USP7 reveals post-translational modification sites and structural requirements for substrate processing and subcellular localization. *FEBS J* **274**(16): 4256-4270

Finley D (2009) Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* **78**: 477-513

Fishburn J, Mohibullah N, Hahn S (2005) Function of a eukaryotic transcription activator during the transcription cycle. *Mol Cell* **18**(3): 369-378

Frontini M, Soutoglou E, Argentini M, Bole-Feysot C, Jost B, Scheer E, Tora L (2005) TAF9b (formerly TAF9L) is a bona fide TAF that has unique and overlapping roles with TAF9. *Mol Cell Biol* **25**(11): 4638-4649

Gamper AM, Kim J, Roeder RG (2009) The STAGA subunit ADA2b is an important regulator of human GCN5 catalysis. *Mol Cell Biol* **29**(1): 266-280

Gangloff YG, Sanders SL, Romier C, Kirschner D, Weil PA, Tora L, Davidson I (2001) Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. *Mol Cell Biol* **21**(5): 1841-1853

Gangloff YG, Werten S, Romier C, Carre L, Poch O, Moras D, Davidson I (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. *Mol Cell Biol* **20**(1): 340-351

Garcia-Oliver E, Garcia-Molinero V, Rodriguez-Navarro S (2011) mRNA export and gene expression: The SAGA-TREX-2 connection. *Biochim Biophys Acta*

Gatchel JR, Zoghbi HY (2005) Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet* **6**(10): 743-755

Georgiev PG, Gerasimova TI (1989) Detection of new genes participating in the trans-regulation of locus yellow of MDG-4 in *Drosophila melanogaster*. *Genetika* **25**(8): 1409-1419

Georgieva S, Nabirochkina E, Dilworth FJ, Eickhoff H, Becker P, Tora L, Georgiev P, Soldatov A (2001) The novel transcription factor e(y)2 interacts with TAF(II)40 and



potentiates transcription activation on chromatin templates. *Mol Cell Biol* **21**(15): 5223-5231

Giles J (2004) Chemistry Nobel for trio who revealed molecular death-tag. *Nature* **431**(7010): 729

Glinsky GV, Berezovska O, Glinskii AB (2005) Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. *J Clin Invest* **115**(6): 1503-1521

Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG (2007) Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Mol Cell* **25**(1): 31-42

Grabbe C, Husnjak K, Dikic I (2011) The spatial and temporal organization of ubiquitin networks. *Nat Rev Mol Cell Biol* **12**(5): 295-307

Grant P, Schieltz D, Pray-Grant M, Yates J, Workman J (1998a) The ATM-related cofactor Tra1 is a component of the purified SAGA complex. *Molecular cell* **2**(6): 863-870

Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL, Workman JL (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev* **11**(13): 1640-1650

Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR, 3rd, Workman JL (1998b) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94**(1): 45-53

Guelman S, Kozuka K, Mao Y, Pham V, Solloway MJ, Wang J, Wu J, Lill JR, Zha J (2009) The double-histone-acetyltransferase complex ATAC is essential for mammalian development. *Mol Cell Biol* **29**(5): 1176-1188

Guelman S, Sukanuma T, Florens L, Weake V, Swanson SK, Washburn MP, Abmayr SM, Workman JL (2006) The essential gene *wda* encodes a WD40 repeat subunit of *Drosophila* SAGA required for histone H3 acetylation. *Mol Cell Biol* **26**(19): 7178-7189

Hahn S, Buratowski S, Sharp PA, Guarente L (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): 1173-1181

Hanna J, Hathaway NA, Tone Y, Crosas B, Elsasser S, Kirkpatrick DS, Leggett DS, Gygi SP, King RW, Finley D (2006) Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation. *Cell* **127**(1): 99-111

Hanna J, Leggett DS, Finley D (2003) Ubiquitin depletion as a key mediator of toxicity by translational inhibitors. *Mol Cell Biol* **23**(24): 9251-9261

- Hansen JC (2002) Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu Rev Biophys Biomol Struct* **31**: 361-392
- Harper JW, Schulman BA (2006) Structural complexity in ubiquitin recognition. *Cell* **124**(6): 1133-1136
- Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, Workman JL (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111**(3): 369-379
- Helmlinger D, Abou-Sleymane G, Yvert G, Rousseau S, Weber C, Trottier Y, Mandel JL, Devys D (2004a) Disease progression despite early loss of polyglutamine protein expression in SCA7 mouse model. *J Neurosci* **24**(8): 1881-1887
- Helmlinger D, Hardy S, Abou-Sleymane G, Eberlin A, Bowman AB, Gansmuller A, Picaud S, Zoghbi HY, Trottier Y, Tora L, Devys D (2006) Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol* **4**(3): e67
- Helmlinger D, Hardy S, Sasorith S, Klein F, Robert F, Weber C, Miguet L, Potier NI, Van-Dorsseleer A, Wurtz J-M, Mandel J-L, Tora L, Devys D (2004b) Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Human molecular genetics* **13**(12): 1257-1322
- Helmlinger D, Marguerat S, Villén J, Gygi S, Bähler J, Winston F (2008) The *S. pombe* SAGA complex controls the switch from proliferation to sexual differentiation through the opposing roles of its subunits Gcn5 and Spt8. *Genes & development* **22**(22): 3184-3279
- Helmlinger D, Marguerat S, Villén J, Swaney D, Gygi S, Bähler J, Winston F (2011) Tra1 has specific regulatory roles, rather than global functions, within the SAGA co-activator complex. *The EMBO journal* **30**(14): 2843-2895
- Henry K, Wyce A, Lo W-S, Duggan L, Emre N, Kao C-F, Pillus L, Shilatifard A, Osley M, Berger S (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes & development* **17**(21): 2648-2711
- Hershko A, Ciechanover A, Varshavsky A (2000) Basic Medical Research Award. The ubiquitin system. *Nat Med* **6**(10): 1073-1081
- Hochstrasser M (2006) Lingering mysteries of ubiquitin-chain assembly. *Cell* **124**(1): 27-34
- Hochstrasser M (2007) Ubiquitin ligation without a ligase. *Dev Cell* **13**(1): 4-6
- Horiuchi J, Silverman N, Marcus GA, Guarente L (1995) ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Mol Cell Biol* **15**(3): 1203-1209

- Horiuchi J, Silverman N, Piñón B, Marcus G, Guarente L (1997) ADA1, a novel component of the ADA/GCN5 complex, has broader effects than GCN5, ADA2, or ADA3. *Molecular and cellular biology* **17**(6): 3220-3228
- Hu M, Gu L, Li M, Jeffrey PD, Gu W, Shi Y (2006) Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: implications for the regulation of the p53-MDM2 pathway. *PLoS Biol* **4**(2): e27
- Hu M, Li P, Li M, Li W, Yao T, Wu JW, Gu W, Cohen RE, Shi Y (2002) Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**(7): 1041-1054
- Hu M, Li P, Song L, Jeffrey PD, Chenova TA, Wilkinson KD, Cohen RE, Shi Y (2005) Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J* **24**(21): 3747-3756
- Huisinga KL, Pugh BF (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol Cell* **13**(4): 573-585
- Hurley JH, Lee S, Prag G (2006) Ubiquitin-binding domains. *Biochem J* **399**(3): 361-372
- Ingvarsdottir K, Krogan N, Emre N, Wyce A, Thompson N, Emili A, Hughes T, Greenblatt J, Berger S (2005) H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Molecular and cellular biology* **25**(3): 1162-1234
- Jani D, Lutz S, Hurt E, Laskey RA, Stewart M, Wickramasinghe VO (2012) Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Res*
- Jani D, Lutz S, Marshall N, Fischer T, Köhler A, Ellisdon A, Hurt E, Stewart M (2009) Sus1, Cdc31, and the Sac3 CID region form a conserved interaction platform that promotes nuclear pore association and mRNA export. *Molecular cell* **33**(6): 727-764
- Jenuwein T, Allis CD (2001) Translating the histone code. *Science* **293**(5532): 1074-1080
- Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, Wang C, Brindle PK, Dent SY, Ge K (2011) Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J* **30**(2): 249-262
- Kato H, van Ingen H, Zhou BR, Feng H, Bustin M, Kay LE, Bai Y (2011) Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc Natl Acad Sci U S A* **108**(30): 12283-12288
- Khorasanizadeh S (2004) The nucleosome: from genomic organization to genomic regulation. *Cell* **116**(2): 259-272

Knutson B, Hahn S (2011) Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Molecular and cellular biology* **31**(4): 818-849

Kohler A, Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* **8**(10): 761-773

Kohler A, Pascual-Garcia P, Llopis A, Zapater M, Posas F, Hurt E, Rodriguez-Navarro S (2006) The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Mol Biol Cell* **17**(10): 4228-4236

Köhler A, Schneider M, Cabal GG, Nehrbass U, Hurt E (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* **10**(6): 707-715

Kohler A, Zimmerman E, Schneider M, Hurt E, Zheng N (2010) Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell* **141**(4): 606-617

Komander D (2009) The emerging complexity of protein ubiquitination. *Biochem Soc Trans* **37**(Pt 5): 937-953

Komander D, Clague MJ, Urbe S (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* **10**(8): 550-563

Komander D, Lord CJ, Scheel H, Swift S, Hofmann K, Ashworth A, Barford D (2008) The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. *Mol Cell* **29**(4): 451-464

Koulich E, Li X, DeMartino GN (2008) Relative structural and functional roles of multiple deubiquitylating proteins associated with mammalian 26S proteasome. *Mol Biol Cell* **19**(3): 1072-1082

Kouranti I, McLean JR, Feoktistova A, Liang P, Johnson AE, Roberts-Galbraith RH, Gould KL (2010) A global census of fission yeast deubiquitinating enzyme localization and interaction networks reveals distinct compartmentalization profiles and overlapping functions in endocytosis and polarity. *PLoS Biol* **8**(9)

Koutelou E, Hirsch C, Dent S (2010) Multiple faces of the SAGA complex. *Current opinion in cell biology* **22**(3): 374-456

Krebs AR, Demmers J, Karmodiya K, Chang NC, Chang AC, Tora L (2010) ATAC and Mediator coactivators form a stable complex and regulate a set of non-coding RNA genes. *EMBO Rep* **11**(7): 541-547

Krebs AR, Karmodiya K, Lindahl-Allen M, Struhl K, Tora L (2011) SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol Cell* **44**(3): 410-423

Krishna SS, Grishin NV (2004) The finger domain of the human deubiquitinating enzyme HAUSP is a zinc ribbon. *Cell Cycle* **3**(8): 1046-1049

Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, Roth SY, Allis CD (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* **383**(6597): 269-272

Kurabe N, Katagiri K, Komiya Y, Ito R, Sugiyama A, Kawasaki Y, Tashiro F (2007) Deregulated expression of a novel component of TFTC/STAGA histone acetyltransferase complexes, rat SGF29, in hepatocellular carcinoma: possible implication for the oncogenic potential of c-Myc. *Oncogene* **26**(38): 5626-5660

Kurshakova M, Krasnov A, Kopytova D, Shidlovskii Y, Nikolenko J, Nabirochkina E, Spohner D, Schultz P, Tora L, Georgieva S (2007) SAGA and a novel Drosophila export complex anchor efficient transcription and mRNA export to NPC. *The EMBO journal* **26**(24): 4956-5021

Kusch T, Guelman S, Abmayr SM, Workman JL (2003) Two Drosophila Ada2 homologues function in different multiprotein complexes. *Mol Cell Biol* **23**(9): 3305-3319

Lam YA, DeMartino GN, Pickart CM, Cohen RE (1997) Specificity of the ubiquitin isopeptidase in the PA700 regulatory complex of 26 S proteasomes. *J Biol Chem* **272**(45): 28438-28446

Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, Martin A (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature*

Lang G, Bonnet J, Umlauf D, Karmodiya K, Koffler J, Stierle M, Devys D, Tora L (2011) The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements. *Mol Cell Biol* **31**(18): 3734-3744

Laprade L, Rose D, Winston F (2007) Characterization of new Spt3 and TATA-binding protein mutants of *Saccharomyces cerevisiae*: Spt3 TBP allele-specific interactions and bypass of Spt8. *Genetics* **177**(4): 2007-2024

Larschan E, Winston F (2001) The *S. cerevisiae* SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. *Genes & development* **15**(15): 1946-2002

Lee BH, Lee MJ, Park S, Oh DC, Elsasser S, Chen PC, Gartner C, Dimova N, Hanna J, Gygi SP, Wilson SM, King RW, Finley D (2010) Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467**(7312): 179-184

Lee K, Florens L, Swanson S, Washburn M, Workman J (2005) The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Molecular and cellular biology* **25**(3): 1173-1255

- Lee K, Sardi M, Swanson S, Gilmore J, Torok M, Grant P, Florens L, Workman J, Washburn M (2011a) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Molecular systems biology* **7**: 503
- Lee K, Swanson S, Florens L, Washburn M, Workman J (2009) Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. *Epigenetics & chromatin* **2**(1): 2
- Lee MJ, Lee BH, Hanna J, King RW, Finley D (2011b) Trimming of ubiquitin chains by proteasome-associated deubiquitinating enzymes. *Mol Cell Proteomics* **10**(5): R110003871
- Leggett DS, Hanna J, Borodovsky A, Crosas B, Schmidt M, Baker RT, Walz T, Ploegh H, Finley D (2002) Multiple associated proteins regulate proteasome structure and function. *Mol Cell* **10**(3): 495-507
- Lenstra T, Benschop J, Kim T, Schulze J, Brabers N, Margaritis T, van de Pasch L, van Heesch SA, Brok M, Groot Koerkamp M, Ko C, van Leenen D, Sameith K, van Hooff S, Lijnzaad P, Kemmeren P, Hentrich T, Kobor M, Buratowski S, Holstege F (2011) The specificity and topology of chromatin interaction pathways in yeast. *Molecular cell* **42**(4): 536-585
- Leurent C, Sanders S, Ruhlmann C, Mallouh V, Weil PA, Kirschner DB, Tora L, Schultz P (2002) Mapping histone fold TAFs within yeast TFIID. *EMBO J* **21**(13): 3424-3433
- Leurent C, Sanders SL, Demeny MA, Garbett KA, Ruhlmann C, Weil PA, Tora L, Schultz P (2004) Mapping key functional sites within yeast TFIID. *EMBO J* **23**(4): 719-727
- Luna-Vargas MP, Faesen AC, van Dijk WJ, Rape M, Fish A, Sixma TK (2011) Ubiquitin-specific protease 4 is inhibited by its ubiquitin-like domain. *EMBO Rep* **12**(4): 365-372
- Ma J, Martin JD, Xue Y, Lor LA, Kennedy-Wilson KM, Sinnamon RH, Ho TF, Zhang G, Schwartz B, Tummino PJ, Lai Z (2010) C-terminal region of USP7/HAUSP is critical for deubiquitination activity and contains a second mdm2/p53 binding site. *Arch Biochem Biophys* **503**(2): 207-212
- Marcus GA, Horiuchi J, Silverman N, Guarente L (1996) ADA5/SPT20 links the ADA and SPT genes, which are involved in yeast transcription. *Mol Cell Biol* **16**(6): 3197-3205
- Marcus GA, Silverman N, Berger SL, Horiuchi J, Guarente L (1994) Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. *EMBO J* **13**(20): 4807-4815
- Martinez E (2002) Multi-protein complexes in eukaryotic gene transcription. *Plant Mol Biol* **50**(6): 925-947
- Martinez E, Kundu TK, Fu J, Roeder RG (1998) A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. *J Biol Chem* **273**(37): 23781-23785

- Martinez E, Palhan VB, Tjernberg A, Lymar ES, Gamper AM, Kundu TK, Chait BT, Roeder RG (2001) Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo. *Mol Cell Biol* **21**(20): 6782-6795
- Maurice T, Duclot F, Meunier J, Naert G, Givalois L, Meffre J, Celerier A, Jacquet C, Copois V, Mechti N, Ozato K, Gongora C (2008) Altered memory capacities and response to stress in p300/CBP-associated factor (PCAF) histone acetylase knockout mice. *Neuropsychopharmacology* **33**(7): 1584-1602
- McMahon SB, Van Buskirk HA, Dugan KA, Copeland TD, Cole MD (1998) The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell* **94**(3): 363-374
- McMahon SJ, Pray-Grant MG, Schieltz D, Yates JR, 3rd, Grant PA (2005) Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. *Proc Natl Acad Sci U S A* **102**(24): 8478-8482
- Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M (2008) Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* **10**(4): 483-488
- Mischerikow N, Spedale G, Altelaar A, Timmers H, Pijnappel W, Heck A (2009) In-depth profiling of post-translational modifications on the related transcription factor complexes TFIID and SAGA. *Journal of proteome research* **8**(11): 5020-5050
- Mohibullah N, Hahn S (2008) Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. *Genes & development* **22**(21): 2994-6000
- Muratoglu S, Georgieva S, Papai G, Scheer E, Enunlu I, Komonyi O, Cserpan I, Lebedeva L, Nabirochkina E, Udvardy A, Tora L, Boros I (2003) Two different Drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol Cell Biol* **23**(1): 306-321
- Mutiu AI, Hoke SM, Genereaux J, Hannam C, MacKenzie K, Jobin-Robitaille O, Guzzo J, Cote J, Andrews B, Haniford DB, Brandl CJ (2007) Structure/function analysis of the phosphatidylinositol-3-kinase domain of yeast tra1. *Genetics* **177**(1): 151-166
- Nagy Z, Riss A, Fujiyama S, Krebs A, Orpinell M, Jansen P, Cohen A, Stunnenberg HG, Kato S, Tora L (2010) The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cell Mol Life Sci* **67**(4): 611-628
- Nagy Z, Riss A, Romier C, le Guezennec X, Dongre A, Orpinell M, Han J, Stunnenberg H, Tora L (2009) The human SPT20-containing SAGA complex plays a direct role in the regulation of endoplasmic reticulum stress-induced genes. *Molecular and cellular biology* **29**(6): 1649-1709

- Nagy Z, Tora L (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**(37): 5341-5357
- Nijman SM, Luna-Vargas MP, Velds A, Brummelkamp TR, Dirac AM, Sixma TK, Bernards R (2005) A genomic and functional inventory of deubiquitinating enzymes. *Cell* **123**(5): 773-786
- Ogryzko VV, Kotani T, Zhang X, Schiltz RL, Howard T, Yang XJ, Howard BH, Qin J, Nakatani Y (1998) Histone-like TAFs within the PCAF histone acetylase complex. *Cell* **94**(1): 35-44
- Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P, Travers AA (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J* **19**(22): 6141-6149
- Palhan VB, Chen S, Peng GH, Tjernberg A, Gamper AM, Fan Y, Chait BT, La Spada AR, Roeder RG (2005) Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc Natl Acad Sci U S A* **102**(24): 8472-8477
- Pankotai T, Komonyi O, Bodai L, Ujfaludi Z, Muratoglu S, Ciurciu A, Tora L, Szabad J, Boros I (2005) The homologous *Drosophila* transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions. *Mol Cell Biol* **25**(18): 8215-8227
- Pankotai T, Popescu C, Martin D, Grau B, Zsindely N, Bodai L, Tora L, Ferrus A, Boros I (2010) Genes of the ecdysone biosynthesis pathway are regulated by the dATAC histone acetyltransferase complex in *Drosophila*. *Mol Cell Biol* **30**(17): 4254-4266
- Papai G, Tripathi MK, Ruhlmann C, Layer JH, Weil PA, Schultz P (2010) TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature* **465**(7300): 956-960
- Papai G, Tripathi MK, Ruhlmann C, Werten S, Crucifix C, Weil PA, Schultz P (2009) Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* **17**(3): 363-373
- Peth A, Besche HC, Goldberg AL (2009) Ubiquitinated proteins activate the proteasome by binding to Usp14/Ubp6, which causes 20S gate opening. *Mol Cell* **36**(5): 794-804
- Pickart CM, Fushman D (2004) Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* **8**(6): 610-616
- Pijnappel W, Timmers H (2008) Dubbing SAGA unveils new epigenetic crosstalk. *Molecular cell* **29**(2): 152-156



- Pina B, Berger S, Marcus GA, Silverman N, Agapite J, Guarente L (1993) ADA3: a gene, identified by resistance to GAL4-VP16, with properties similar to and different from those of ADA2. *Mol Cell Biol* **13**(10): 5981-5989
- Powell D, Weaver C, Jennings J, McAfee K, He Y, Weil P, Link A (2004) Cluster analysis of mass spectrometry data reveals a novel component of SAGA. *Molecular and cellular biology* **24**(16): 7249-7308
- Pray-Grant MG, Schieltz D, McMahon SJ, Wood JM, Kennedy EL, Cook RG, Workman JL, Yates JR, 3rd, Grant PA (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol Cell Biol* **22**(24): 8774-8786
- Qiu XB, Ouyang SY, Li CJ, Miao S, Wang L, Goldberg AL (2006) hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37. *EMBO J* **25**(24): 5742-5753
- Reeves WM, Hahn S (2005) Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* **25**(20): 9092-9102
- Renatus M, Parrado SG, D'Arcy A, Eidhoff U, Gerhartz B, Hassiepen U, Pierrat B, Riedl R, Vinzenz D, Worpenberg S, Kroemer M (2006) Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. *Structure* **14**(8): 1293-1302
- Reyes-Turcu FE, Horton JR, Mullally JE, Heroux A, Cheng X, Wilkinson KD (2006) The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* **124**(6): 1197-1208
- Reyes-Turcu FE, Shanks JR, Komander D, Wilkinson KD (2008) Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T. *J Biol Chem* **283**(28): 19581-19592
- Reyes-Turcu FE, Ventii KH, Wilkinson KD (2009) Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu Rev Biochem* **78**: 363-397
- Robert F, Hardy S, Nagy Z, Baldeyron C, Murr R, Dery U, Masson JY, Papadopoulo D, Herceg Z, Tora L (2006) The transcriptional histone acetyltransferase cofactor TRRAP associates with the MRN repair complex and plays a role in DNA double-strand break repair. *Mol Cell Biol* **26**(2): 402-412
- Roberts SM, Winston F (1996) SPT20/ADA5 encodes a novel protein functionally related to the TATA-binding protein and important for transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* **16**(6): 3206-3213
- Rodriguez-Navarro S (2009) Insights into SAGA function during gene expression. *EMBO reports* **10**(8): 843-893
- Rodriguez-Navarro S, Fischer Ts, Luo M-J, Antón J, Brettschneider S, Lechner J, Pöschel J, Reed R, Hurt E (2004) Sus1, a functional component of the SAGA

histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**(1): 75-161

Rojas JR, Trievel RC, Zhou J, Mo Y, Li X, Berger SL, Allis CD, Marmorstein R (1999) Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* **401**(6748): 93-98

Rondon AG, Jimeno S, Aguilera A (2010) The interface between transcription and mRNP export: from THO to THSC/TREX-2. *Biochim Biophys Acta* **1799**(8): 533-538

Saleh A, Schieltz D, Ting N, McMahon SB, Litchfield DW, Yates JR, 3rd, Lees-Miller SP, Cole MD, Brandl CJ (1998) Tra1p is a component of the yeast Ada.Spt transcriptional regulatory complexes. *J Biol Chem* **273**(41): 26559-26565

Samara NL, Datta AB, Berndsen CE, Zhang X, Yao T, Cohen RE, Wolberger C (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* **328**(5981): 1025-1029

Samara NL, Wolberger C (2011) A new chapter in the transcription SAGA. *Curr Opin Struct Biol* **21**(6): 767-774

Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA (2002) Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol Cell Biol* **22**(13): 4723-4738

Sanso M, Vargas-Perez I, Quintales L, Antequera F, Ayte J, Hidalgo E (2011) Gcn5 facilitates Pol II progression, rather than recruitment to nucleosome-depleted stress promoters, in *Schizosaccharomyces pombe*. *Nucleic Acids Res* **39**(15): 6369-6379

Saridakis V, Sheng Y, Sarkari F, Holowaty MN, Shire K, Nguyen T, Zhang RG, Liao J, Lee W, Edwards AM, Arrowsmith CH, Frappier L (2005) Structure of the p53 binding domain of HAUSP/USP7 bound to Epstein-Barr nuclear antigen 1 implications for EBV-mediated immortalization. *Mol Cell* **18**(1): 25-36

Sarkari F, La Delfa A, Arrowsmith CH, Frappier L, Sheng Y, Saridakis V (2010) Further insight into substrate recognition by USP7: structural and biochemical analysis of the HdmX and Hdm2 interactions with USP7. *J Mol Biol* **402**(5): 825-837

Sarkari F, Sanchez-Alcaraz T, Wang S, Holowaty MN, Sheng Y, Frappier L (2009) EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. *PLoS Pathog* **5**(10): e1000624

Sellam A, Askew C, Epp E, Lavoie H, Whiteway M, Nantel A (2009) Genome-wide mapping of the coactivator Ada2p yields insight into the functional roles of SAGA/ADA complex in *Candida albicans*. *Molecular biology of the cell* **20**(9): 2389-2789

Semple CA (2003) The comparative proteomics of ubiquitination in mouse. *Genome Res* **13**(6B): 1389-1394

Sermwittayawong D, Tan S (2006) SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. *EMBO J* **25**(16): 3791-3800

Sheng Y, Saridakis V, Sarkari F, Duan S, Wu T, Arrowsmith CH, Frappier L (2006) Molecular recognition of p53 and MDM2 by USP7/HAUSP. *Nat Struct Mol Biol* **13**(3): 285-291

Shevchenko A, Roguev A, Schaft D, Buchanan L, Habermann B, Sakalar C, Thomas H, Krogan NJ, Stewart AF (2008) Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment. *Genome Biol* **9**(11): R167

Smith ER, Eisen A, Gu W, Sattah M, Pannuti A, Zhou J, Cook RG, Lucchesi JC, Allis CD (1998) ESA1 is a histone acetyltransferase that is essential for growth in yeast. *Proc Natl Acad Sci U S A* **95**(7): 3561-3565

Soldatov A, Nabirochkina E, Georgieva S, Belenkaja T, Georgiev P (1999) TAFII40 protein is encoded by the e(y)1 gene: biological consequences of mutations. *Mol Cell Biol* **19**(5): 3769-3778

Song EJ, Werner SL, Neubauer J, Stegmeier F, Aspden J, Rio D, Harper JW, Elledge SJ, Kirschner MW, Rape M (2010) The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* **24**(13): 1434-1447

Sowa ME, Bennett EJ, Gygi SP, Harper JW (2009) Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**(2): 389-403

Spedale G, Meddens CA, Koster MJ, Ko CW, van Hooff SR, Holstege FC, Timmers HT, Pijnappel WW (2012) Tight cooperation between Mot1p and NC2beta in regulating genome-wide transcription, repression of transcription following heat shock induction and genetic interaction with SAGA. *Nucleic Acids Res* **40**(3): 996-1008

Spedale G, Mischerikow N, Heck AJ, Timmers HT, Pijnappel WW (2010) Identification of Pep4p as the protease responsible for formation of the SAGA-related SLIK protein complex. *J Biol Chem* **285**(30): 22793-22799

Starita LM, Parvin JD (2006) Substrates of the BRCA1-dependent ubiquitin ligase. *Cancer Biol Ther* **5**(2): 137-141

Sterner D, Belotserkovskaya R, Berger S (2002a) 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): 11622-11629

Sterner D, Grant P, Roberts S, Duggan L, Belotserkovskaya R, Pacella L, Winston F, Workman J, Berger S (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): 86-184

- Sterner DE, Wang X, Bloom MH, Simon GM, Berger SL (2002b) The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex. *J Biol Chem* **277**(10): 8178-8186
- Stewart M (2010) Nuclear export of mRNA. *Trends Biochem Sci* **35**(11): 609-617
- Suganuma T, Gutierrez JL, Li B, Florens L, Swanson SK, Washburn MP, Abmayr SM, Workman JL (2008) ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nat Struct Mol Biol* **15**(4): 364-372
- Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**(3): 105-178
- Timmers HT, Tora L (2005) SAGA unveiled. *Trends Biochem Sci* **30**(1): 7-10
- van der Knaap JA, Kumar BR, Moshkin YM, Langenberg K, Krijgsveld J, Heck AJ, Karch F, Verrijzer CP (2005) GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7. *Mol Cell* **17**(5): 695-707
- Venters BJ, Pugh BF (2009) How eukaryotic genes are transcribed. *Crit Rev Biochem Mol Biol* **44**(2-3): 117-141
- Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamon AJ, Jain P, Roller NS, Jiang C, Hemeryck-Walsh C, Pugh BF (2011) A comprehensive genomic binding map of gene and chromatin regulatory proteins in *Saccharomyces*. *Mol Cell* **41**(4): 480-492
- Ventii KH, Wilkinson KD (2008) Protein partners of deubiquitinating enzymes. *Biochem J* **414**(2): 161-175
- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**(6): 967-980
- Wang L, Liu L, Berger S (1998) Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo. *Genes & development* **12**(5): 640-693
- Wang YL, Faiola F, Xu M, Pan S, Martinez E (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. *J Biol Chem* **283**(49): 33808-33815
- Warfield L, Ranish JA, Hahn S (2004) Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. *Genes Dev* **18**(9): 1022-1034
- Weake V, Dyer J, Seidel C, Box A, Swanson S, Peak A, Florens L, Washburn M, Abmayr S, Workman J (2011) Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation. *Genes & development* **25**(14): 1499-2008

- Weake V, Lee K, Guelman S, Lin C-H, Seidel C, Abmayr S, Workman J (2008) SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the *Drosophila* visual system. *The EMBO journal* **27**(2): 394-799
- Weake VM, Swanson SK, Mushegian A, Florens L, Washburn MP, Abmayr SM, Workman JL (2009) A novel histone fold domain-containing protein that replaces TAF6 in *Drosophila* SAGA is required for SAGA-dependent gene expression. *Genes Dev* **23**(24): 2818-2823
- Weake VM, Workman JL (2008) Histone ubiquitination: triggering gene activity. *Mol Cell* **29**(6): 653-663
- Weake VM, Workman JL (2011) SAGA function in tissue-specific gene expression. *Trends Cell Biol*
- Wieczorek E, Brand M, Jacq X, Tora L (1998) Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* **393**(6681): 187-191
- Wilkinson KD (2009) DUBs at a glance. *J Cell Sci* **122**(Pt 14): 2325-2329
- Wilson SM, Bhattacharyya B, Rachel RA, Coppola V, Tessarollo L, Householder DB, Fletcher CF, Miller RJ, Copeland NG, Jenkins NA (2002) Synaptic defects in ataxia mice result from a mutation in *Usp14*, encoding a ubiquitin-specific protease. *Nat Genet* **32**(3): 420-425
- Winget JM, Mayor T (2010) The diversity of ubiquitin recognition: hot spots and varied specificity. *Mol Cell* **38**(5): 627-635
- Winston F, Chaleff D, Valent B, Fink G (1984a) Mutations affecting Ty-mediated expression of the *HIS4* gene of *Saccharomyces cerevisiae*. *Genetics* **107**(2): 179-276
- Winston F, Dollard C, Malone EA, Clare J, Kapakos JG, Farabaugh P, Minehart PL (1987) Three genes are required for trans-activation of Ty transcription in yeast. *Genetics* **115**(4): 649-656
- Winston F, Durbin KJ, Fink GR (1984b) The *SPT3* gene is required for normal transcription of Ty elements in *S. cerevisiae*. *Cell* **39**(3 Pt 2): 675-682
- Wu P-YJ, Ruhlmann C, Winston F, Schultz P (2004) Molecular architecture of the *S. cerevisiae* SAGA complex. *Molecular cell* **15**(2): 199-407
- Wu P-YJ, Winston F (2002) Analysis of Spt7 function in the *Saccharomyces cerevisiae* SAGA coactivator complex. *Molecular and cellular biology* **22**(15): 5367-5446
- Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR, Roth SY (2000) Loss of *Gcn5l2* leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet* **26**(2): 229-232

- Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, Nakatani Y (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**(6589): 319-324
- Ye Y, Akutsu M, Reyes-Turcu F, Enchev RI, Wilkinson KD, Komander D (2011) Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. *EMBO Rep* **12**(4): 350-357
- Ye Y, Scheel H, Hofmann K, Komander D (2009) Dissection of USP catalytic domains reveals five common insertion points. *Mol Biosyst* **5**(12): 1797-1808
- Zhang XY, Varthi M, Sykes SM, Phillips C, Warzecha C, Zhu W, Wyce A, Thorne AW, Berger SL, McMahon SB (2008) The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol Cell* **29**(1): 102-111
- Zhao Y, Lang G, Ito S, Bonnet J, Metzger E, Sawatsubashi S, Suzuki E, Le Guezennec X, Stunnenberg HG, Krasnov A, Georgieva SG, Schule R, Takeyama K, Kato S, Tora L, Devys D (2008) A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell* **29**(1): 92-101
- Zohn IE, Li Y, Skolnik EY, Anderson KV, Han J, Niswander L (2006) p38 and a p38-interacting protein are critical for downregulation of E-cadherin during mouse gastrulation. *Cell* **125**(5): 957-969

# A TFTC/STAGA Module Mediates Histone H2A and H2B Deubiquitination, Coactivates Nuclear Receptors, and Counteracts Heterochromatin Silencing

Yue Zhao,<sup>2</sup> Guillaume Lang,<sup>1</sup> Saya Ito,<sup>2</sup> Jacques Bonnet,<sup>1</sup> Eric Metzger,<sup>3</sup> Shun Sawatsubashi,<sup>2</sup> Eriko Suzuki,<sup>2</sup> Xavier Le Guezennec,<sup>4</sup> Hendrik G. Stunnenberg,<sup>4</sup> Aleksey Krasnov,<sup>5</sup> Sofia G. Georgieva,<sup>5</sup> Roland Schüle,<sup>3</sup> Ken-Ichi Takeyama,<sup>2</sup> Shigeaki Kato,<sup>2</sup> László Tora,<sup>1,\*</sup> and Didier Devys<sup>1,\*</sup>

<sup>1</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR 7104, INSERM U 596, Université Louis Pasteur de Strasbourg, BP 10142-67404 ILLKIRCH Cedex, CU de Strasbourg 67404, France

<sup>2</sup>Laboratory of Nuclear Signaling, Institute of Molecular and Cellular Biosciences, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-0032, Japan

<sup>3</sup>Universitäts-Frauenklinik und Zentrum für Klinische Forschung, Klinikum der Universität Freiburg, Breisacherstrasse 66, 79106 Freiburg, Germany

<sup>4</sup>Department of Molecular Biology, NCMLS, Radboud University Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

<sup>5</sup>Institute of Gene Biology, Russian Academy of Sciences, Moscow 119334, Russia

\*Correspondence: laszlo@igbmc.u-strasbg.fr (L.T.), devys@igbmc.u-strasbg.fr (D.D.)

DOI 10.1016/j.molcel.2007.12.011

## SUMMARY

Transcriptional activators, several different coactivators, and general transcription factors are necessary to access specific loci in the dense chromatin structure to allow precise initiation of RNA polymerase II (Pol II) transcription. Histone acetyltransferase (HAT) complexes were implicated in loosening the chromatin around promoters and thus in gene activation. Here we demonstrate that the 2 MDa GCN5 HAT-containing metazoan TFTC/STAGA complexes contain a histone H2A and H2B deubiquitinase activity. We have identified three additional subunits of TFTC/STAGA (ATXN7L3, USP22, and ENY2) that form the deubiquitination module. Importantly, we found that this module is an enhancer of position effect variegation in *Drosophila*. Furthermore, we demonstrate that ATXN7L3, USP22, and ENY2 are required as cofactors for the full transcriptional activity by nuclear receptors. Thus, the deubiquitinase activity of the TFTC/STAGA HAT complex is necessary to counteract heterochromatin silencing and acts as a positive cofactor for activation by nuclear receptors *in vivo*.

## INTRODUCTION

Transcription initiation by RNA polymerase II (Pol II) is a highly regulated process that requires the coordinated action of numerous factors. Coactivators are multisubunit complexes that are recruited to promoters by gene-specific activators to facilitate transcription initiation either by direct interaction with general transcription factors (GTFs) or Pol II or indirectly through modification of chromatin structure (reviewed in Li et al., 2007 and Martinez, 2002).

The yeast 2 MDa SAGA coactivator complex and its *Drosophila* and human homologs, the TBP-free TAF complex (TFTC) and the SPT3/TAF9/GCN5 acetyltransferase complex (STAGA), are multisubunit complexes that facilitate access of GTFs to DNA through histone acetylation mediated by the catalytic activity of the GCN5 subunit (Martinez, 2002). Human homologs of most SAGA subunits have been identified in STAGA and TFTC, and the three-dimensional structure of TFTC can be superimposed to that of SAGA (Wu et al., 2004). In addition, biochemical and functional characterization demonstrated that the described human complexes are almost identical (hereafter called TFTC/STAGA complex; reviewed in Nagy and Tora, 2007). Thus, genetic and structural studies point to an evolutionarily conserved function of these complexes with modular organization that relates to their multiple activities in regulating gene expression.

Gcn5 and other SAGA or TFTC/STAGA subunits have been shown to play a role in ligand-dependent gene activation by nuclear receptors (Wallberg et al., 2000 and Yanagisawa et al., 2002, and references therein). These observations suggested a model in which nuclear receptors recruit, in addition to other coactivator complexes, the SAGA/TFTC/STAGA complex, and that this contributes to chromatin remodeling via a mechanism involving histone acetylation.

Biochemical studies identified new components of the yeast SAGA complex that confer additional functions to the complex (Powell et al., 2004; Rodriguez-Navarro et al., 2004; Sanders et al., 2002). A ubiquitin-specific protease, Ubp8, was shown to be incorporated into SAGA through interaction with Sgf11 and to confer a histone H2B ubiquitin protease activity to the complex (Daniel et al., 2004; Henry et al., 2003; Ingvarsdotir et al., 2005; Lee et al., 2005). H2B has been shown to be monoubiquitinated (H2Bub1) on Lys 123 in *S. cerevisiae* by the E2-conjugating enzyme Rad6 and the E3-ligase enzyme Bre1 (reviewed in Zhang, 2003). Rad6-mediated H2Bub1 was found to regulate subsequent methylation of histone H3 at Lys 4 and Lys 79 through

a “trans tail” mechanism and was shown to be required for subtelomeric gene silencing (Zhang, 2003). In addition to its role in gene silencing, H2B ubiquitination plays an important role in activation of specific genes. H2Bub1 seems to be transient at the promoter of SAGA-dependent genes, and both ubiquitination and deubiquitination of H2B are required for optimal gene activation (Daniel et al., 2004; Henry et al., 2003). Interestingly, Sus1, which was shown to be a component of both SAGA and the Sac3-Thp1 mRNA export complex (Rodriguez-Navarro et al., 2004), interacts with Sgf11 and Ubp8 and regulates the deubiquitination activity of the complex (Kohler et al., 2006). In higher eukaryotes, histone H2A is also monoubiquitinated (H2Aub1), a modification which has not been reported in yeast. H2Aub1 is more prevalent (5%–15%) when compared with that of H2Bub1 (1% to 2%). H2A ubiquitination is mediated by the Polycomb repressive complex 1 (PRC1) and is crucial for Polycomb-mediated gene silencing and X inactivation (de Napoles et al., 2004; Joo et al., 2007; Wang et al., 2004).

In this study we identified ATXN7L3, USP22, and ENY2 as the human orthologs of yeast Sgf11, Ubp8, and Sus1, respectively, and show that they are integral components of TFTC/STAGA complex. These three proteins together form a module of the TFTC/STAGA complex, which specifically removes the ubiquitin moiety from monoubiquitinated histones H2A and H2B. We demonstrate that the deubiquitination module of the *Drosophila* TFTC/STAGA complex counteracts heterochromatin silencing in vivo and both the *Drosophila* and the human deubiquitination module are required for full transcriptional activation by the androgen receptor (AR). These findings suggest that the deubiquitinase activity of the TFTC/STAGA complex, in addition to its histone acetyltransferase (HAT) activity, participates in the dynamic turnover of histone modifications during the transcription cycle and functions as a positive cofactor for activation by nuclear receptors.

## RESULTS

### USP22, ATXN7L3, and ENY2 Are Components of TFTC/STAGA Complex

Homologs of recently identified yeast SAGA subunits (i.e., Ubp8, Sgf11, and Sus1) remained to be identified in the mammalian TFTC/STAGA complex. We analyzed a highly purified TFTC fraction (Wieczorek et al., 1998) by MS-MS mass spectrometry and identified peptides corresponding to USP22, ATXN7L3, and ENY2 (see the Experimental Procedures). Human USP22 is the closest homolog of yeast Ubp8 having a similar organization to the yeast ubiquitin protease possessing an N-terminal zinc finger domain (ZnF-UBP) and a C-terminal ubiquitin carboxyl-terminal hydrolase motif (UCH-2; Figure 1A). Human ATXN7L3 is the homolog of the yeast Sgf11 protein containing a highly conserved zinc finger domain at its N-terminal end (ZnF-Sgf11; Figure 1B). We have previously described ATXN7L3, an unknown protein at that time, based on its homology to ATXN7, another subunit of the TFTC/STAGA complex (Helmlinger et al., 2004). Indeed, ATXN7 and ATXN7L3 share a second C-terminal domain, the so-called SCA7 domain (InterPro: IPR013243), which is highly conserved through all members of the ATXN7 gene family (Figure 1B). In addition, we identified a peptide corresponding to

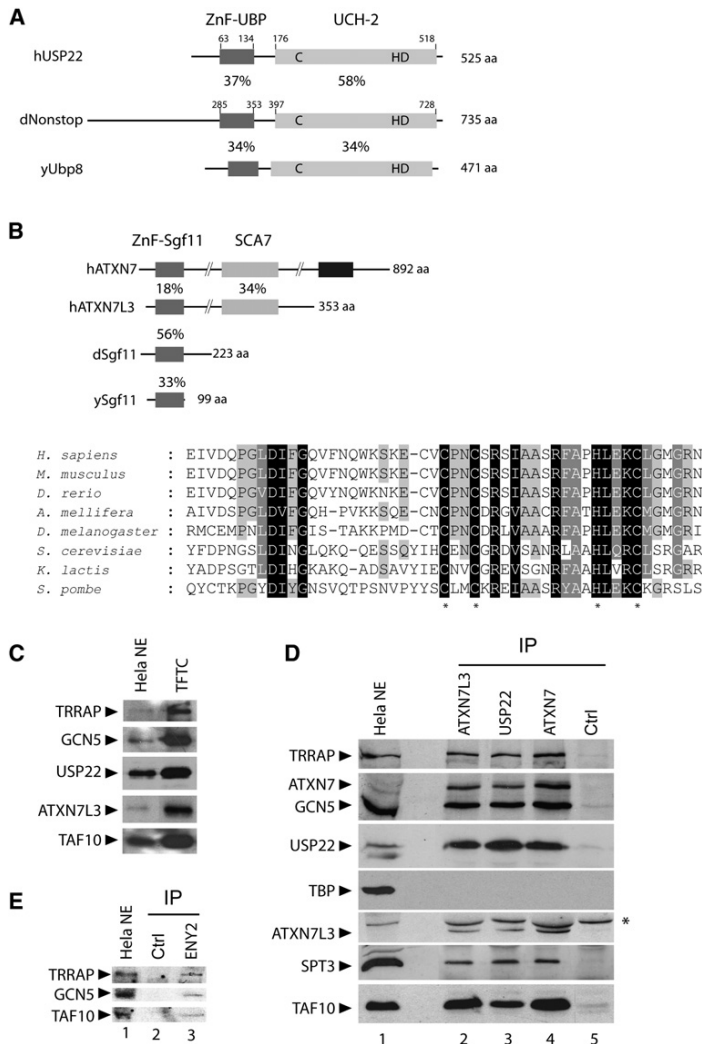
ENY2, a 101 amino acid protein that is homologous to yeast Sus1.

To test whether the identified polypeptides are indeed TFTC/STAGA subunits, polyclonal antibodies against synthetic peptides corresponding to USP22, ATXN7L3, and ENY2 were raised. These antibodies efficiently recognized the corresponding recombinant proteins overexpressed in Sf9 cells at the expected size, 60 kDa for USP22, 42 kDa for ATXN7L3, and 13 kDa for ENY2 (see Figure S1 available online). We thus analyzed whether USP22 and ATXN7L3 would be present in a highly purified TFTC preparation. Western blot analysis using anti-USP22 and anti-ATXN7L3 detected a specific signal at the expected size in the HeLa cell nuclear extract and showed that USP22 and ATXN7L3 were enriched in the purified TFTC similarly to known components of TFTC/STAGA such as TRRAP, GCN5, and TAF10 (Figure 1C). In order to confirm that USP22, ATXN7L3, and ENY2 are integral components of the TFTC/STAGA complex, we performed a series of coimmunoprecipitation experiments on HeLa cell nuclear extract using antibodies against USP22, ATXN7L3, or ENY2. Following these immunoprecipitations, we analyzed USP22 and ATXN7L3-associated proteins by western blot and identified different TFTC subunits, such as TRRAP, ATXN7, GCN5, SPT3, or TAF10 (Figure 1D, lanes 2 and 3). Similarly, TRRAP, GCN5, and TAF10 were detected in an anti-ENY2 immunoprecipitation (Figure 1E). None of the TFTC subunits were detected in a control immunoprecipitation using preimmune sera (Figure 1D, lane 5 and Figure 1E, lane 2). TBP, a component of TFIID, but not of TFTC/STAGA, was not found associated with USP22, ATXN7L3, or ENY2. We previously showed that ATXN7 is a specific subunit of TFTC/STAGA (Helmlinger et al., 2004). Western blot analysis of complexes obtained after an ATXN7 immunoprecipitation showed that USP22 and ATXN7L3 associate with ATXN7 together with other components of TFTC/STAGA (Figure 1D, lane 4). Taken together, these results indicated that USP22, ATXN7L3, and ENY2 are bona fide subunits of the TFTC/STAGA complex.

### Identification of the TFTC/STAGA Deubiquitination Subcomplex

Having established that ATXN7L3, USP22, and ENY2 are components of the TFTC/STAGA complex, we next asked how these subunits interact and incorporate into TFTC/STAGA. In order to identify putative interactions between these proteins, 293T cells were transfected with expression vectors encoding Flag-tagged ATXN7L3 or USP22, or HA-tagged ENY2, and proteins were immunoprecipitated using anti-ATXN7L3 or anti-HA antibodies. Immunoblot analysis of purified complexes revealed an interaction between ATXN7L3 and USP22 (Figure 2A, lane 1) and between ATXN7L3 and ENY2 (Figure 2A, lanes 3 and 7), but not between ENY2 and USP22 (Figure 2A, lanes 8). When the three proteins were coexpressed, they formed a stable complex that could be purified both with anti-ATXN7L3 or anti-HA antibodies (Figure 2A, lanes 2 and 6). Formation of a stable TFTC module composed of ATXN7L3, USP22, and ENY2 was further confirmed by coinfection of Sf9 cells with different combinations of baculoviruses and immunoprecipitation of ATXN7L3-associated proteins (Figure 2B). Note that when USP22 and ENY2 were coexpressed in the absence of recombinant ATXN7L3, no signal corresponding





**Figure 1. USP22, ATXN7L3, and ENY2 Are Integral Components of TFTC/STAGA Complexes**

(A) Schematic representation of Ubp8 orthologs in human (hUSP22), *Drosophila* (dNonstop), and *S. cerevisiae* (yUbp8). Sequence identities in the N-terminal zinc finger domain (ZnF-UBP, dark gray) and in the C-terminal catalytic site (UCH-2, light gray) are indicated relative to the human USP22 sequence. Cysteine, histidine, and aspartic residues forming the catalytic triad were positioned in the proteins. SwissProt/TrEmbl/RefSeq accession numbers are as follows: *Homo sapiens* USP22 (Q9UPT9), *Drosophila melanogaster* Nonstop (NP\_524140), and *Saccharomyces cerevisiae* Ubp8 (P50102).

(B) Schematic representation of ATXN7 and Sgf11 orthologs in *S. cerevisiae* (ySgf11), *Drosophila* (dSgf11), and human (hATXN7L3). Sequence identities in the N-terminal zinc finger (ZnF-Sgf11, dark gray) or in the more C-terminal SCA7 domain (light gray) are shown when compared to ATXN7L3. Lower panel, sequence alignment of the ZnF-Sgf11 from different species. Conserved residues with 100%, 80%, and 60%–80% identity are indicated with black shading, dark gray shading, and light gray shading, respectively. Histidine and cysteine residues that coordinate zinc are labeled (\*). SwissProt/TrEmbl/RefSeq accession numbers are as follows: *Homo sapiens* ATXN7L3 (Q14CW9), *Mus musculus* Atxn7l3 (XP\_997968.1), *Brachydanio rerio* (NP\_001005396.1), *Apis mellifera* (XP\_393291.2), *Drosophila melanogaster* (NP\_649050.1), *Saccharomyces cerevisiae* (NP\_015278.1), *Kluyveromyces lactis* (XP\_453582.1), and *Schizosaccharomyces pombe* (NP\_001018231.1).

(C) Proteins in HeLa cell nuclear extract (HeLa NE) or in a purified TFTC fraction were detected by western blot with the indicated antibodies.

(D) HeLa cell NEs were immunoprecipitated using anti-ATXN7L3, anti-USP22, or anti-ATXN7 antibodies or preimmune serum as a control (Ctrl). Purified complexes were analyzed by immunoblotting using antibodies against the indicated TFTC/STAGA subunits. A fraction (5%) of the input nuclear extract before immunoprecipitation was loaded as a control. IgG heavy chain migrating just above ATXN7L3 is indicated (\*).

(E) Complexes associated with ENY2 were analyzed by immunoprecipitation of HeLa NE using an anti-ENY2 antibody or preimmune serum as a control (Ctrl). Purified complexes and a fraction (5%) of the input nuclear extract were analyzed by immunoblotting revealed with the indicated antibodies.

to these proteins could be detected after an ATXN7L3 immunoprecipitation (Figure 2A, lane 4 and Figure 2B, lane 5).

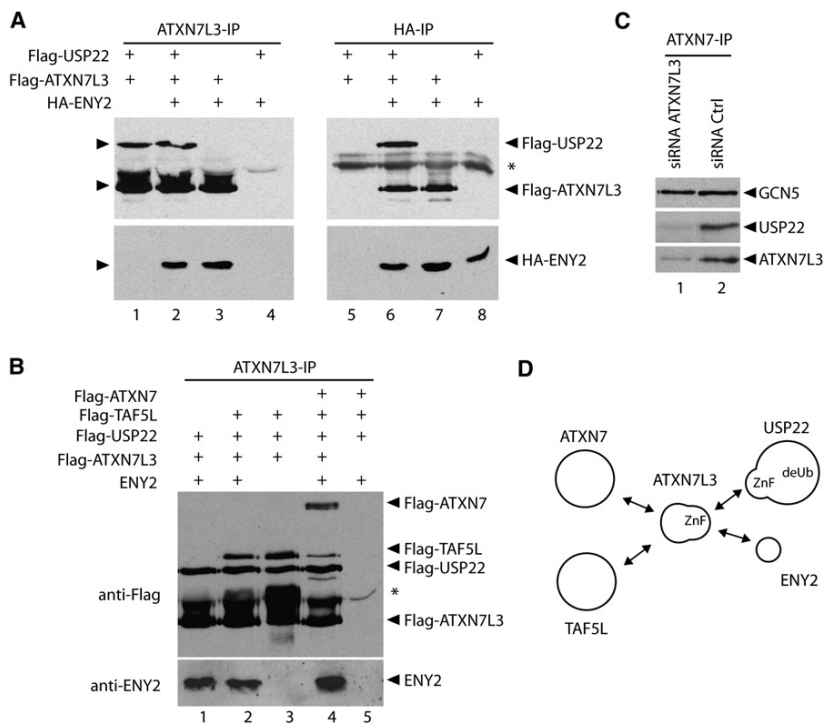
To analyze whether ATXN7L3 is required for USP22 association with TFTC/STAGA, we have downregulated ATXN7L3 expression by an siRNA treatment. HeLa cells were transfected with siRNA against ATXN7L3, cells were grown for 48 hr, TFTC/STAGA was purified by an ATXN7 immunoprecipitation, and the levels of TFTC/STAGA subunits were compared to those of complexes purified from control cells (Figure 2C). While the amount of GCN5 did not change in the purified complexes, USP22 was significantly reduced in the complex prepared from the cells where ATXN7L3 was knocked down (Figure 2C). These results together suggest that ATXN7L3, USP22, and ENY2 form a stable subcomplex and that USP22 and ENY2 are recruited into TFTC/STAGA by ATXN7L3 (Figure 2D).

We further asked which domains of ATXN7L3 and USP22 are involved in the formation of this subcomplex. Cotransfection and coimmunoprecipitation experiments revealed that the ATXN7L3 ZnF-Sgf11 domain alone is able to interact with both ENY2 and

USP22 as efficiently as the full-length ATXN7L3 (see Figure S3B). Similar experiments using USP22 deletion mutants showed that its interaction with ATXN7L3 required the ZnF-UBP domain, but not the catalytic domain of USP22 (data not shown). Thus, the formation of this module would be driven by interactions between ATXN7L3 and USP22 through their respective ZnF-Sgf11 and ZnF-UBP domains, and between ATXN7L3 ZnF-Sgf11 and ENY2 (Figure 2D). Finally, we showed that two different TFTC/STAGA subunits, TAF5L and ATXN7, interacted with this module and may mediate its association with TFTC/STAGA (Figure 2B and Figures S2 and S3).

#### Deubiquitination Activity of TFTC/STAGA

It has been shown that the yeast SAGA specifically removes ubiquitin moiety from H2Bub1 (Henry et al., 2003). We thus tested whether the human TFTC/STAGA complex has a similar activity on human H2Bub1. Although it plays an important role in transcription regulation, H2Bub1 is expected to represent less than 2% of H2B in human cells. To detect H2Bub1, HEK293T cells



**Figure 2. Interaction between Subunits of the TFTC/STAGA Deubiquitination Module**

(A) 293T cells were transfected with different combinations of plasmids expressing epitope-tagged USP22, ATXN7L3, and ENY2 as indicated. Whole-cell extracts were made and proteins were immunoprecipitated using anti-ATXN7L3 or anti-HA antibodies. Purified proteins were analyzed by western blotting using anti-Flag (upper panels) or anti-HA antibodies (lower panels). IgG heavy chain is indicated (\*).

(B) Sf9 cells were infected with different combinations of baculoviruses expressing ATXN7, TAF5L, USP22, ATXN7L3, and ENY2 as indicated. Immunoprecipitations using an anti-ATXN7L3 antibody were analyzed by western blotting revealed with the indicated antibodies. IgG heavy chain migrating above ATXN7L3 is indicated (\*).

(C) HeLa cells were transfected with siRNA against ATXN7L3 or a control (Ctrl) siRNA. TFTC/STAGA complexes were purified using an anti-ATXN7 antibody and analyzed by western blotting.

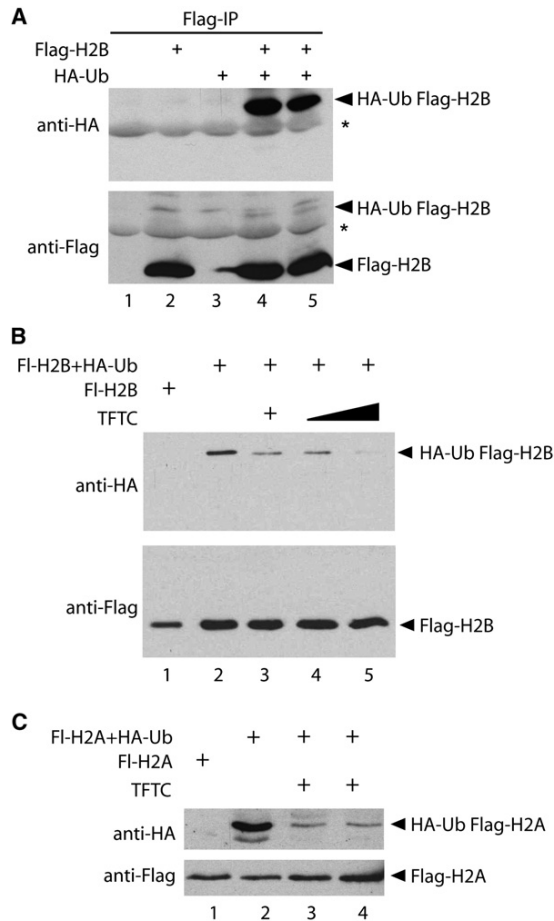
(D) Schematic representation of interactions between different subunits of the TFTC deubiquitination module.

were cotransfected with expression vectors encoding a Flag-tagged histone H2B (Flag-H2B) and/or an HA-tagged ubiquitin (HA-Ub). Cells were grown for 48 hr, and recombinant histone H2B was immunoprecipitated using an anti-Flag antibody. When testing the purified proteins by immunoblotting, the anti-HA antibody specifically detected H2Bub1 (Figure 3A, lanes 4 and 5), while the anti-Flag antibody revealed the unmodified histone H2B and a faint signal corresponding to H2Bub1. The specific detection of H2Bub1 was also confirmed by the absence of signal when Flag-H2B or HA-Ub was omitted from the transfection reactions (Figure 3A, lanes 2 and 3).

To characterize the deubiquitination activity of TFTC/STAGA, we used the immunopurified fraction of H2B, which contains both unmodified H2B and H2Bub1 as a substrate in the *in vitro* deubiquitination assay. Two independent TFTC preparations decreased the level of H2Bub1 significantly, when analyzed with an anti-HA antibody, while the level of unmodified H2B was equal in each reaction (Figure 3B, compare lane 2 to lanes 3 and 4). The levels of H2Bub1 were further decreased when the amount of purified TFTC was increased in the deubiquitination assay, suggesting a dose-dependent effect (Figure 3B, lanes 4 and 5). Because in mammalian cells an important fraction of histone H2A is monoubiquitinated on Lys 119, we investigated the deubiquitination activity of TFTC/STAGA on the H2Aub1. Recombinant Flag-H2A-HA-Ub was prepared as described above for H2Bub1 and used as a substrate in an *in vitro* deubiquitination assay. Two different TFTC preparations could deubiquitinate H2Aub1 with a high efficiency (Figure 3C). From these results we conclude that the human TFTC/STAGA complex deubiquitinates H2Aub1 and H2Bub1.

### Both *nonstop*, the *Drosophila* Homolog of USP22, and *dSgf11*, the *Drosophila* Homolog of ATXN7L3, Counteract Genomic Silencing Mediated by Heterochromatin

To characterize the *in vivo* function of the TFTC/STAGA deubiquitination activity, we followed a genetic approach in *Drosophila*. The *Drosophila* homolog of USP22 is called *nonstop* (Figure 1A) and was shown to be required for glial cell and neuronal development (Poeck et al., 2001). To determine a possible role of *nonstop* in the modulation of chromatin structure *in vivo*, we have used the well-defined *In(1)W<sup>m4h</sup>* system that allows the identification of modifiers of position effect variegation (PEV) (Tartof et al., 1984). The *In(1)W<sup>m4h</sup>* inversion brings the *white* locus close to a heterochromatic breakpoint leading to variegated expression of the *white* gene and a mosaic pattern of eye pigmentation. We found that mutation of one *nonstop* copy (*not<sup>d2069</sup>*, 75D4) or a deficient line, lacking the chromosomal region containing *nonstop* (*Df[3L]JED225, 75C1-75D4*), gave significant decrease in pigment area in this PEV model, resulting in a 2-fold decrease of pigment levels in *In(1)W<sup>m4h</sup>* (Figure 4A). In good agreement, we also showed that overexpression of full-length (FL) Nonstop increased eye pigmentation to wild-type levels in the *In(1)W<sup>m4h</sup>* PEV model (Figure 4B). To further investigate which domain of Nonstop is required for the effect on PEV modification, we constructed Nonstop N- or C-terminal truncation mutations. Transgenic fly lines expressing only the deubiquitinase catalytic domain of Nonstop (UCH-2) had significant increase in pigment area in the eye when tested in the *In(1)W<sup>m4h</sup>* PEV model (although weaker than FL), whereas there was no modification of PEV in transgenic fly lines expressing only the ZnF-UBP domain of Nonstop



**Figure 3. Deubiquitination Activity of the TFTC/SAGA Complex**

(A) Purification of recombinant H2B and monoubiquitinated H2B (H2Bub1). HEK293T cells were cotransfected with plasmids encoding Flag-H2B and HA-Ub. Proteins from whole-cell extracts were immunopurified with an anti-Flag antibody and analyzed by immunoblotting with the indicated antibodies. H2Bub1 is specifically detected with the anti-HA antibody just above the IgG light chain (\*).

(B) The purified TFTC complex removes ubiquitin from H2Bub1. Recombinant H2B, unmodified and monoubiquitinated, was eluted by peptide competition and used as a substrate in an in vitro deubiquitination assay using two independent TFTC preparations (in lane 3 a different preparation was used than in lanes 4 and 5). Recombinant H2B and H2Bub1 were analyzed by immunoblotting revealed with anti-Flag and anti-HA antibodies as indicated.

(C) Recombinant ubiquitinated histone H2A was purified as H2Bub1 in (A) and used in an in vitro deubiquitination assay with two independent TFTC preparations (lanes 3 and 4) as in (B).

(Figure 4B). These data strongly suggest that the deubiquitination activity of Nonstop is required for the modification of PEV in the *ln(1)W<sup>m4h</sup>* system.

Because both yUbp8/ySgf11 and their human homologs USP22/ATXN7L3 seem to exert their activity in a common sub-complex, next we set out to identify the *Drosophila* homolog of ySgf11 or ATXN7L3. The *CG13379* gene product (hereafter called dSgf11) was identified as a homolog of yeast Sgf11 and human ATXN7L3 (Figure 1B). We then investigated the effect

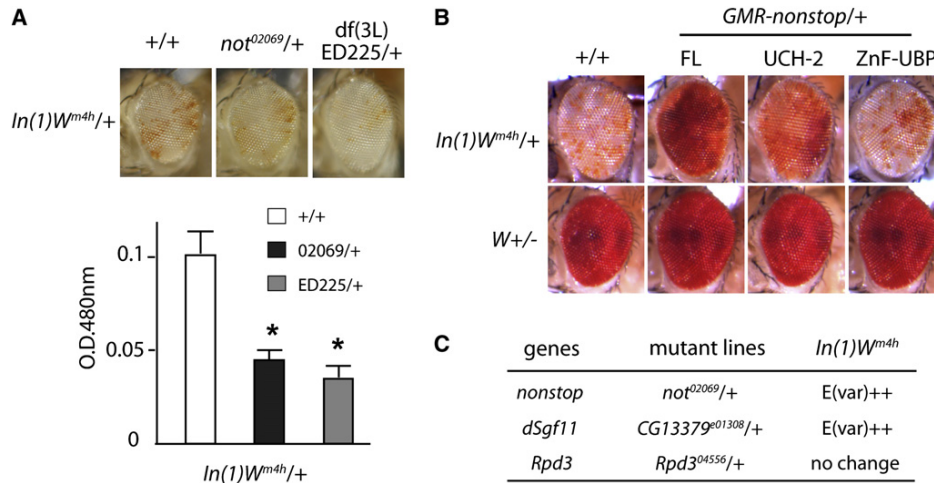
of mutations of the *dSgf11* gene on the *ln(1)W<sup>m4h</sup>* PEV system. *dSgf11*, similarly to *nonstop*, acted as an enhancer of PEV as its mutations led to decreased *white* expression assessed by eye pigmentation (Figure 4C). In contrast, mutation of *dRpd3*, a histone deacetylase, had no visible effect on PEV in the *ln(1)W<sup>m4h</sup>* system (Figure 4C). These results together suggest that both Nonstop and its partner dSgf11 are required to counteract heterochromatic gene silencing in vivo.

### The Subunits of the *Drosophila* TFTC/STAGA Deubiquitination Module Are Required for Androgen Receptor-Mediated Transactivation

The fact that TFTC/STAGA has been implicated as a chromatin-modifying cofactor for activated transcription and that both Nonstop and dSgf11 counteract heterochromatic gene silencing suggested that the deubiquitination module of dTFTC/STAGA may play a coactivator role for transcriptional activators by modifying chromatin. To test this possibility, we developed an androgen receptor (AR)-associated position effect variegation (AR-PEV) model in *Drosophila*. In this system, the AR or a truncated AR harboring the ligand-independent AF-1 domain (AR[AF-1]) is expressed in the *Drosophila* eye using the *glass multimer reporter* (*GMR*) gene promoter (Figure 5A). A reporter construct that contains the *white* gene and a gene encoding the green fluorescent protein (GFP) controlled by eight AR-responsive elements (AREs) was inserted into a heterochromatic region leading to a mosaic red eye phenotype (Figures 5A and 5B). In this system, the AR counteracted silencing by heterochromatin spreading in a ligand-dependent manner (data not shown). The ligand-independent AR(AF-1) led to higher transactivation of the reporter genes and was hereafter used to analyze the AR-dependent alteration of chromatin structure. Two different mutations of *nonstop* (*not<sup>02069</sup>* and *Df[3L]ED225*) gave significant decrease (2.2-fold) in the pigment area in this AR-dependent PEV model (Figure 5B). In good agreement, we also show that overexpression of Nonstop increased eye pigmentation to wild-type levels in the AR(AF-1)-PEV model (data not shown).

We further analyzed the effects of Nonstop on AR-dependent gene activation in vivo by testing AR(AF-1)-mediated transactivation on the GFP reporter gene expression system. In order to be able to detect GFP expression, we used additional transgenic lines expressing AR(AF-1) in the *Drosophila* eye using a GMR GAL4 driver in a GAL4-UAS system. The ARE-GFP reporter was inserted in euchromatic region and further mobilized into pericentric heterochromatin leading to variegated expression of GFP reporter (Y.Z., K.-I.T., and S.K., unpublished data). In agreement with our results obtained with the above described AR-dependent PEV model, this reporter system confirmed again that *nonstop* loss-of-function mutations, *not<sup>02069</sup>* and *Df[3L]ED225*, dominantly reduce AR(AF-1) transactivation of the reporter transgene inserted in pericentric heterochromatin (Figure 5C). In contrast, but consistent with the loss-of-function effects, *nonstop* overexpression (*UAS-nonstop*) significantly increased GFP transactivation by AR(AF-1) (Figure 5C).

Having established that overexpression of Nonstop-mediated AR-dependent gene activation in pericentric region, we investigated whether this activity requires the association of Nonstop with TFTC/STAGA complex. To this end, we examined how



**Figure 4. *Drosophila* Nonstop Counteracts Heterochromatic Gene Silencing**

(A) Two different fly lines carrying *nonstop* loss-of-function mutation (*not*<sup>02069</sup>) or deficiency line including *nonstop* gene location (*Df(3L)ED225*) or wild-type flies (+/+) were crossed with *In(1)W<sup>m4h</sup>* flies. Modification of PEV was analyzed by areas of eye pigmentation in the progeny (upper panels) and by optical density (OD) measurement at 480 nm (lower panel). Average values of more than three independent measurements are shown with SD. \**p* < 0.01 when compared to +/+ control (Student's *t* test).

(B) The catalytic domain of Nonstop is required for its modification effect on PEV. *In(1)W<sup>m4h</sup>* flies were crossed with wild-type flies (+/+) or with flies overexpressing full-length Nonstop (FL), Nonstop catalytic domain (UCH-2), or Nonstop ZnF-UBP (ZnF-UBP) (upper panels). Effect of Nonstop overexpression was assessed by modification of eye pigmentation. Flies carrying the *white* gene in its normal locus were used as a control (lower panels).

(C) Both Nonstop and Sgf11 are enhancers of PEV. Heterozygous males from the lines bearing mutations of *dSgf11* and *Rpd3* were mated to (*In(1)W<sup>m4h</sup>*) females. Eye pigmentation was analyzed in 2- to 3-day-old male progeny.

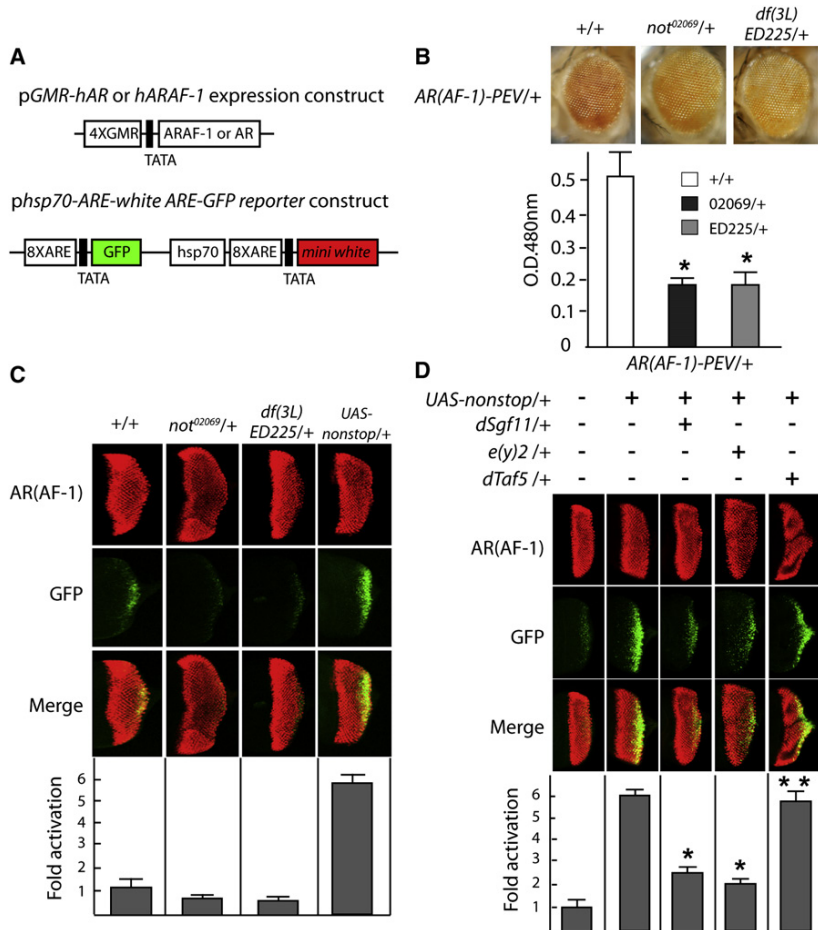
mutations of different TFTC/STAGA components affect Nonstop-mediated AR(AF-1) transcriptional activation in vivo using the second GFP transactivation system. As shown in Figure 5D, Nonstop overexpression enhanced AR(AF-1) transactivation of the ARE-GFP reporter gene expression. In good agreement with our biochemical interaction studies, mutations of *dSgf11* (CG13379<sup>e01308</sup>) and *e(y)2* (*Df(1)m259-4*) significantly reduced GFP expression (Figure 5D). Importantly, the coactivation activity of Nonstop was not modified by mutation of *dTAF5* (*Taf5*<sup>EY01764b</sup>), which is a specific component of TFIID. Taken together, these results suggest that the deubiquitination activity of *Drosophila* TFTC/STAGA, mediated by Nonstop, requires the components of the TFTC/STAGA deubiquitination module to play an important role in AR-dependent transactivation in vivo.

#### The Human Deubiquitination Module and Its Activity Are Positive Modulators of AR-Mediated Transactivation in Human Cells

The fact that *Drosophila* Nonstop, Sgf11, and E(y)2 increased in vivo transactivation activity of the AR on silenced reporter transgenes suggested that these proteins play a coactivator role in AR activation. To further investigate this hypothesis, we analyzed putative interactions of USP22 or ATXN7L3 with AR in mammalian cells. Cotransfection in HEK293T cells and coimmunoprecipitation experiments revealed that AR interacted with ATXN7L3 in a ligand-dependent manner, similarly to GCN5, while interaction between AR and USP22 was detected also in the absence of ligand (see Figure S6A). As AR and USP22 interacted in vivo, we next tested whether AR could be a substrate of the TFTC/STAGA deubiquitination activity. Coexpression of

increasing amounts of USP22 with AR in 293T cells revealed that the AR steady-state levels were not modified, indicating that AR is not a substrate of USP22 (see Figure S6B).

In order to determine whether USP22, ATXN7L3, and ENY2 are required for AR-dependent transcriptional activity in human cells, we performed AR-dependent reporter assays. 293T cells were transfected with an AR-dependent reporter gene together with expression vectors encoding AR alone, or together with USP22, ATXN7L3, and ENY2 in the presence or absence of ligand (Figures 6A and 6B). USP22, ATXN7L3, and ENY2 significantly enhanced AR-mediated transactivation in the presence of ligand (2.5- to 3.5-fold compared to AR alone), similarly to the effect of the GCN5 HAT (Figures 6A and 6B). Furthermore, we observed that USP22 dominantly increased also ER- or GR-mediated transactivation in similar reporter assay analyses (data not shown). These results suggested that USP22, ATXN7L3, and ENY2 can function as coactivators for nuclear receptor-mediated transactivation. We next examined whether the deubiquitination activity of USP22 is required for coactivation of AR transcriptional activity. To this end we constructed an expression vector encoding a USP22 mutant in which the two conserved histidine residues in the catalytic site (His box) were mutated to alanines. Such mutations in the active site of ubiquitin-specific proteases were shown to result in undetectable levels of catalytic activity (Hu et al., 2002). Expression of this inactive USP22 catalytic mutant did not change the ligand-induced activation of the AR-dependent luciferase reporter (Figure 6A), further demonstrating that USP22 deubiquitination activity is required for its nuclear receptor coactivation function. In good agreement with our observations in *Drosophila*, ATXN7L3, ENY2, and USP22,



**Figure 5. The Androgen Receptor Coactivation Function of Nonstop Requires dSgf11 and dE(y)2**

(A) Schematic representation of the expression and reporter constructs. The expression constructs include the full-length human androgen receptor (hAR) or hAR(AF-1) driven by four GMR binding sites. The reporter construct harbors GFP and *white* reporter genes driven by the hsp70 promoter in which eight AREs were introduced.

(B) AR(AF-1)-PEV flies were crossed with *nonstop* mutant lines, and progeny were analyzed as in Figure 4A.

(C) Flies expressing AR(AF-1) in the eye using a GMR GAL4 driver and carrying an ARE-GFP reporter in pericentric heterochromatin were crossed with lines harboring *nonstop* loss-of-function (*not<sup>02069</sup>*, *Df[3L]ED225*) or gain-of-function (*UAS-nonstop*) mutations as indicated. Expression of AR(AF-1) was assessed by immunostaining using an anti-AR antibody (upper panels). The effect of *nonstop* mutations or overexpression on AR(AF-1)-mediated transactivation was assessed by examination of GFP expression (middle panels). Merge images are shown in lower panels. (D) Flies carrying ARE-GFP reporter and AR(AF-1) expression transgenes were crossed with flies carrying *UAS-nonstop* and loss-of-function mutations of *dSgf11* (*CG13379<sup>e01308</sup>*), *e(y)2* (*Df[1]m259-4*), and *dTaf5* (*Taf5<sup>EY01764b</sup>*). Analysis of AR(AF-1) transactivation was performed on eye disc of the third instar larvae from the progeny. Average values of four independent measurements are shown with SD. \**p* < 0.001 and \*\**p* > 0.10 when compared to *UAS-nonstop*/+ control (Student's *t* test).

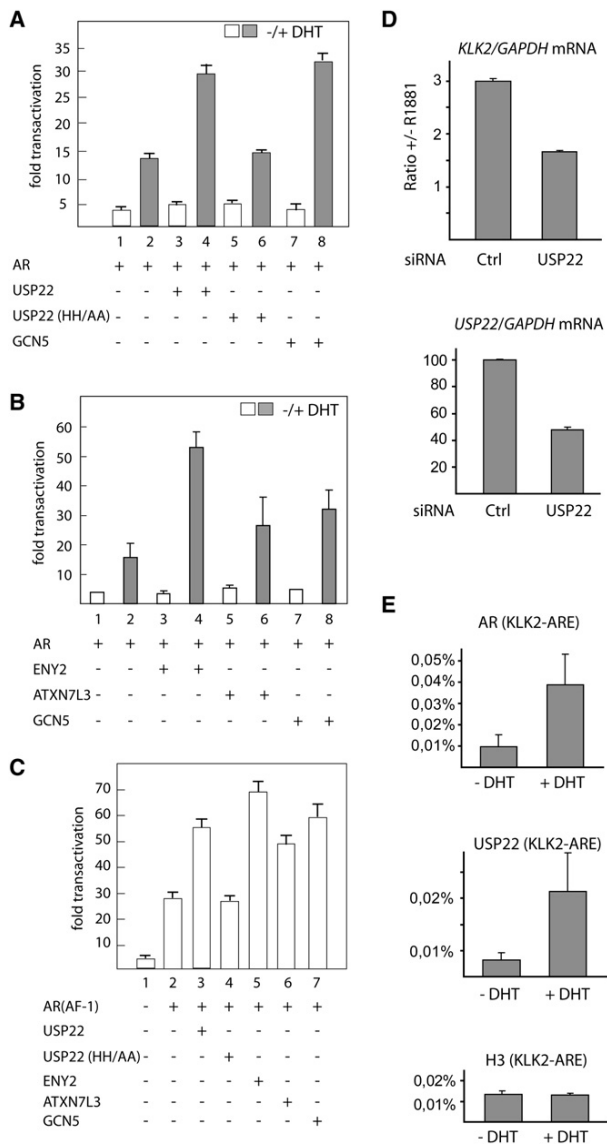
but not the USP22 catalytic mutant, increased transactivation also by the ligand-independent AR(AF-1) in similar reporter assays (Figure 6C).

We then asked whether the TFTC/STAGA deubiquitination activity is required for endogenous AR-dependent gene activation in vivo. To this end, we turned to LNCaP prostate tumor cells in which the AR is constitutively expressed and analyzed the androgen-induced expression of the endogenous AR target gene *KLK2* after downregulation of USP22. Transfection of LNCaP cells with siRNA against *USP22* led to a 55% reduction of *USP22* mRNA and protein levels (Figure 6D and data not shown). Induction of *KLK2* expression by the synthetic AR agonist R1881 in LNCaP cells transfected with control siRNA was 3-fold and significantly reduced in cells transfected with siRNA against *USP22* (1.6-fold) (Figure 6D). This result together with the observed interaction between the AR and both ATXN7L3 and USP22 suggested that these proteins are recruited together to the promoters of AR-dependent genes upon ligand induction. To test this hypothesis, chromatin was prepared from LNCaP cells treated with or without ligand and subjected to chromatin immunoprecipitation (ChIP) using anti-AR and anti-USP22 antibodies. Recruitment of the AR was significantly increased at the promoter region of *KLK2* gene after ligand addition

(Figure 6E). Similarly, USP22 associated with the chromatin of *KLK2* promoter in a ligand-dependent manner (Figure 6E). No significant association of AR and USP22 with chromatin at an intergenic region could be detected, and the amount of histone H3 at *KLK2* promoter remained constant, further demonstrating specificity (Figure 6E). Taken together, these results indicate that upon ligand induction TFTC/STAGA is recruited to the promoter of AR-dependent genes where its deubiquitination activity facilitates gene expression.

## DISCUSSION

In this study, we identified three subunits of the 2 MDa GCN5-containing TFTC/STAGA complex. We show that the ubiquitin protease USP22 forms a subcomplex with ATXN7L3 (*ySgf11* homolog) and ENY2 (*ySus1* homolog). Importantly, we demonstrate that the catalytic domain of the ubiquitin protease together with *dSgf11* is necessary to modify PEV in the *Drosophila* eye, pointing to the involvement of the deubiquitination module in chromatin remodeling. Furthermore, we show that the subunits of the deubiquitination subcomplex are required for the full activity of the ligand-induced AR and other nuclear receptors in vivo. Thus, to our knowledge we established for the first time the



**Figure 6. Androgen Receptor-Dependent Transactivation Requires the Deubiquitination Activity of USP22, ATXN7L3, and ENY2 in Human Cells**

(A and B) 293T cells were cotransfected with plasmids expressing the full-length androgen receptor (AR) together with USP22, USP22(HH/AA), GCN5 (A), ATXN7L3, or ENY2 (B) expression plasmids with AR-dependent reporter in the absence (white bars) or presence (gray bars) of ligand (DHT). GCN5, ATXN7L3, ENY2, and USP22, but not a USP22 catalytic mutant, mediate AR-induced transcriptional activity.

(C) Similar results were obtained using the ligand-independent AF-1 domain (AR[AF-1]). In (A)–(C), the error bars represent mean  $\pm$  SD ( $n \geq 4$ ).

(D) LNCaP cells were transfected with either control (Ctrl) or USP22 siRNA and treated with or without ligand (R1881). Induction of *KLK2* expression is expressed as the ratio of *KLK2* mRNA levels normalized to *GAPDH* levels between cells treated and untreated with R1881 (upper panel). *USP22* downregulation was assessed by quantitative RT-PCR. *USP22* levels are presented as a percentage of the mean of control cells after normalization to *GAPDH* mRNA levels (lower panel).

deubiquitination activity of the metazoan TFTC/STAGA HAT complexes as a second positive chromatin-modifying function.

### TFTC/STAGA Complex Has a Modular Organization that Links the Deubiquitination Function to HAT, Activator Binding, and Transcription Initiation

An overlay of the yeast SAGA and the human TFTC structures indicated, in good agreement with their similar subunit composition (Nagy and Tora, 2007), a high degree of structural conservation in size and shape (Wu et al., 2004). Several SAGA subunits have been localized in one of the five domains of the complex (Wu et al., 2004). Gcn5 was detected in the central region of SAGA and colocalized with Ada2 and Ada3, indicating that this central domain harbors the HAT activity. We have now identified a module of TFTC/STAGA, which is responsible for the deubiquitination activity of this complex on H2Aub1 and H2Bub1. This module is composed of three subunits, among which ATXN7L3 (hSGF11) interacts with TAF5L and ATXN7 and allows the recruitment of both USP22 and ENY2 to TFTC/STAGA. Because yTAF5 localizes close to the Gcn5-containing HAT module, it is tempting to speculate that a functional connection may exist between the HAT and the deubiquitination modules. Along these lines it is interesting to note that neither free USP22 nor a stable recombinant subcomplex, composed of TAF5L, ATXN7L3, ENY2, and USP22, can deubiquitinate H2Aub1 or H2Bub1 in vitro (see Figure S4), suggesting that additional TFTC subunits are required for the regulation of this activity.

Interestingly, the two different TFTC/STAGA subunits, ATXN7 (ySgf73 homolog) and ATXN7L3 (ySgf11 homolog) share a conserved domain and may have a related function within the complex. A polyglutamine expansion in ATXN7 has been identified as the mutation responsible for an inherited neurodegenerative disorder, Spinocerebellar ataxia type 7 (SCA7). It remains to be investigated whether a polyglutamine expansion in ATXN7 deregulates the TFTC/STAGA deubiquitination activity and whether this would account for chromatin modifications and transcriptional alterations previously identified in SCA7 models (Helmlinger et al., 2006).

The yeast and the *Drosophila* homolog of hENY2 (Sus1 and E[y]2, respectively) were identified as subunits of SAGA and an anchoring and mRNA export (AMEX) complex, homolog of the yeast Sac3-Thp1-Sus1 complex (Kurshakova et al., 2007; Rodriguez-Navarro et al., 2004). Sus1 or dE(y)2 as subunits of AMEX function together with SAGA/TFTC in the anchoring of a subset of transcription sites to the nuclear pore complexes (NPCs) to achieve efficient transcription and mRNA export (Cabal et al., 2006; Kurshakova et al., 2007). The association of USP22 and ENY2 in the same TFTC/STAGA module would therefore

(E) LNCaP cells were incubated with or without ligand (DHT) and subjected to ChIP using antibodies against AR (upper panel), USP22 (middle panel), or unmodified histone H3 (lower panel). Precipitated DNA was quantified by real-time qPCR using primers flanking the promoter region of *KLK2* gene (–343 to –90). The values are expressed as percentage of the amount of immunoprecipitated DNA normalized to the respective input DNA signal. Background signals obtained from an intergenic region were subtracted for the quantification of AR and USP22 association with the specific region. All results are representative of at least two independent experiments. In (D) and (E), the error bars represent mean  $\pm$  SD ( $n = 2$ ).

suggest that two activities, histone deubiquitination and mRNA export, may also be coregulated within this complex. Although we showed that TFTC/STAGA deubiquitinates H2Aub1 and H2Bub1, USP22 could act on other nonhistone targets as suggested by the accumulation of several ubiquitinated proteins after inactivation of *Drosophila nonstop* (Poeck et al., 2001).

Different histone deubiquitinases involved in transcriptional regulation have been identified in metazoan. *Drosophila* USP7 specifically deubiquitinates H2Bub1 and regulates Polycomb-mediated silencing (van der Knaap et al., 2005). Recently, two ubiquitin proteases, Ubp-M/USP16 and 2A-DUB/KIAA1915/MYSM1, were identified as specific H2A deubiquitinases regulating *Hox* gene expression and AR-dependent gene activation (Joo et al., 2007; Zhu et al., 2007). In contrast to the specific recognition of either H2Aub1 or H2Bub1 by these enzymes, we showed that TFTC/STAGA is active on both H2Aub1 and H2Bub1 in vitro. Thus, it seems that, in addition to these enzymes that seem to be active only at specific target genes, TFTC/STAGA has a broader specificity to counteract heterochromatin silencing and to regulate gene activation by nuclear receptors. Alternatively, the action of TFTC/STAGA on either H2Aub1 or H2Bub1 could be restricted to specific genes in vivo.

### The Functional Link between the HAT and the Deubiquitination Activities of TFTC/STAGA in Gene Activation

Genetic and biochemical analysis in yeast and *Drosophila* suggested that the 2 MDa GCN5-containing complexes have multiple roles in the cells: global deposition of histone H3 acetylation marks along the entire chromosomes and targeted acetylation of histones at specific loci. Our results obtained with the ubiquitin protease mutants in both the *In(1)W<sup>m4h</sup>* PEV system or the AR-dependent PEV model suggest that, similarly to its HAT activity, the deubiquitination module of SAGA/TFTC is influencing both the global chromatin architecture, as a modifier of PEV, and the gene-specific function of the complex, as a nuclear receptor coactivator. This intriguing parallel raises the following questions: is the TFTC/STAGA HAT activity in vivo functionally connected with deubiquitination activity within the complex, and which activity is needed first? Interestingly, all the subunits of the deubiquitination module of TFTC/STAGA are able to function as coactivators during the activation process of AR, similarly to what has been demonstrated for individual subunits of the HAT module (Meng et al., 2004; vom Baur et al., 1998) or for the entire 2 MDa HAT complexes (Yanagisawa et al., 2002). In this respect the HAT and the deubiquitination modules seem to fulfill similar functions suggesting that the two activities together are necessary to achieve full activation of nuclear receptors on chromatin templates. Histone acetylation is thought to open up the chromatin, which may suggest that the deubiquitination activity of the complex would work on templates, which have been previously acetylated by the complex.

Our study pinpoints the monoubiquitination of histone H2A and H2B as a mark that is timely regulated for appropriate transcription in metazoan organisms and further emphasizes that the precise turnover of chromatin modifications is crucial for the control of gene expression. The association of both HAT and deubiquitination activities in the TFTC/STAGA complex provides

an attractive mechanism by which the so called “cross-talk” between given histone marks is coordinated within the same regulatory complex. Further mechanistic studies will be required to address the exact link between these activities and the other chromatin-modifying complexes to understand how these sequential events participate in chromatin remodeling and gene activation.

### EXPERIMENTAL PROCEDURES

#### Plasmids, Fly Stocks, and Genetics

Plasmid construction, generation of transgenic flies, fly stocks, and crossing schemes are presented in the Supplemental Experimental Procedures.

#### Mass Spectrometry

Mass spectrometry was performed using LTQ-FT mass spectrometer (Thermo Electron, Bremen, Germany) essentially as described previously (Olsen et al., 2004). Peptides identified were as following: SPHIPYK, DFFLSDR, and ITSNC TIGLR corresponding to USP22; SHLSLVGTASGLGSNKK and SLLTTQCGVI SEHTK corresponding to ATXN7L3; and TFLAQHASL corresponding to ENY2.

#### Antibodies

The anti-USP22 (2391 and 2392) and anti-ENY2 (2584 and 2585) polyclonal antibodies were obtained by immunization of rabbits with peptides corresponding to residues 2–15 (VSRPEPEGEAMDAE) of human USP22 and to residues 50–66 (CKEVIKEKLEHVTVD) of human ENY2. Rabbits and mice immunization with ATXN7L3 peptides corresponding to residues 278–296 (LQWDGSSDLSPSDSGSSKT) and to residues 335–353 (KKKPKPPAPPTPS IYDDIN) produced the 2326 polyclonal and the 2ATX-2B1 monoclonal antibodies and the 2325 polyclonal and 1ATX-2D7 monoclonal antibodies, respectively.

#### In Vitro Deubiquitination Assay

293T cells transfected with Flag-H2A or Flag-H2B (gift from R.G. Roeder) and/or HA-Ub (gift from I. Dikic) plasmids were collected in phosphate buffer saline 1X (PBS1X) containing 20 mM N-ethylmaleimide (NEM) and 1 mM PMSF. Whole-cell extracts prepared in modified RIPA buffer (50 mM Tris-HCl [pH 7.9], 150 mM NaCl, 5 mM EDTA, 15 mM MgCl<sub>2</sub>, 1% NP40, 0.5% sodium deoxycholate, 1 mM DTT, 1X protease inhibitor, 1 mM PMSF, and 20 mM NEM) were immunoprecipitated using M2 beads (Sigma). After extensive washing, bound proteins were eluted with an excess of free Flag peptide. Eluted proteins were used as a substrate in a deubiquitination reaction with different amounts of highly purified TFTC complex prepared as described (Wieczorek et al., 1998). The reaction was performed in 100 mM Tris-HCl (pH 7.9), 5% glycerol, 1 mM DTT, and 1 mM EDTA for 2 hr at 37°C. Unmodified and monoubiquitinated recombinant H2A or H2B were analyzed by immunoblotting using anti-Flag and anti-HA antibodies.

#### Supplemental Data

Supplemental Data include six figures, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at <http://www.molecularcell.org/cgi/content/full/29/1/92/DC1/>.

#### ACKNOWLEDGMENTS

We thank D. Helmlinger, O. Bombarde, M. Vermeulen, M. Oulad-Abdelghani, and G. Duval for their contribution to this work. We thank Z. Nagy and M.E. Torres Padilla for critically reading the manuscript, and R.G. Roeder, I. Dikic, T. Nagase, and E. Martinez for generous gifts of different reagents. Research in the authors' laboratory is supported by grants from the the Agence Nationale de la Recherche (ANR-05-MRAR-O29-01), from the Fondation de la Recherche Médicale (FRM, DLC20060206408), and from Connaître les Syndromes Cerebelleux to D.D. and from ANR (05-BLAN-0396-01; Regulome), European Community (HPRN-CT 00504228 and STREP LSHG-CT-2004-502950) to L.T.

Received: June 19, 2007  
 Revised: November 2, 2007  
 Accepted: December 15, 2007  
 Published: January 17, 2008

## REFERENCES

- Cabal, G.G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J.C., Hurt, E.C., and Nehrbass, U. (2006). SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature* **441**, 770–773.
- Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates, J.R., III, and Grant, P.A. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J. Biol. Chem.* **279**, 1867–1871.
- de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676.
- Helmlinger, D., Hardy, S., Sasorith, S., Klein, F., Robert, F., Weber, C., Miguet, L., Potier, N., Van-Dorsselaer, A., Wurtz, J.M., et al. (2004). Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum. Mol. Genet.* **13**, 1257–1265.
- Helmlinger, D., Hardy, S., Abou-Sleymane, G., Eberlin, A., Bowman, A.B., Gansmuller, A., Picaud, S., Zoghbi, H.Y., Trotter, Y., Tora, L., and Devys, D. (2006). Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol.* **4**, e67. 10.1371/journal.pbio.0040067.
- Henry, K.W., Wyce, A., Lo, W.S., Duggan, L.J., Emre, N.C., Kao, C.F., Pillus, L., Shilatifard, A., Osley, M.A., and Berger, S.L. (2003). Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev.* **17**, 2648–2663.
- Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J.W., Gu, W., Cohen, R.E., and Shi, Y. (2002). Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**, 1041–1054.
- Ingvarsdottir, K., Krogan, N.J., Emre, N.C., Wyce, A., Thompson, N.J., Emili, A., Hughes, T.R., Greenblatt, J.F., and Berger, S.L. (2005). H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol. Cell. Biol.* **25**, 1162–1172.
- Joo, H.Y., Zhai, L., Yang, C., Nie, S., Erdjument-Bromage, H., Tempst, P., Chang, C., and Wang, H. (2007). Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* **449**, 1068–1072.
- Kohler, A., Pascual-Garcia, P., Llopis, A., Zapater, M., Posas, F., Hurt, E., and Rodriguez-Navarro, S. (2006). The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitylation through its interaction with Ubp8 and Sgf11. *Mol. Biol. Cell* **17**, 4228–4236.
- Kurshakova, M.M., Krasnov, A.N., Kopytova, D.V., Shidlovskii, Y.V., Nikolenko, J.V., Nabirochikina, E.N., Spehner, D., Schultz, P., Tora, L., and Georgieva, S.G. (2007). SAGA and a novel *Drosophila* export complex anchor efficient transcription and mRNA export to NPC. *EMBO J.* **26**, 4956–4965.
- Lee, K.K., Florens, L., Swanson, S.K., Washburn, M.P., and Workman, J.L. (2005). The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol. Cell. Biol.* **25**, 1173–1182.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* **128**, 707–719.
- Martinez, E. (2002). Multi-protein complexes in eukaryotic gene transcription. *Plant Mol. Biol.* **50**, 925–947.
- Meng, G., Zhao, Y., Nag, A., Zeng, M., Dimri, G., Gao, Q., Wazer, D.E., Kumar, R., Band, H., and Band, V. (2004). Human ADA3 binds to estrogen receptor (ER) and functions as a coactivator for ER-mediated transactivation. *J. Biol. Chem.* **279**, 54230–54240.
- Nagy, Z., and Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**, 5341–5357.
- Olsen, J.V., Ong, S.E., and Mann, M. (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **3**, 608–614.
- Poeck, B., Fischer, S., Gunning, D., Zipursky, S.L., and Salecker, I. (2001). Glial cells mediate target layer selection of retinal axons in the developing visual system of *Drosophila*. *Neuron* **29**, 99–113.
- Powell, D.W., Weaver, C.M., Jennings, J.L., McAfee, K.J., He, Y., Weil, P.A., and Link, A.J. (2004). Cluster analysis of mass spectrometry data reveals a novel component of SAGA. *Mol. Cell. Biol.* **24**, 7249–7259.
- Rodriguez-Navarro, S., Fischer, T., Luo, M.J., Antunez, O., Brettschneider, S., Lechner, J., Perez-Ortin, J.E., Reed, R., and Hurt, E. (2004). Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**, 75–86.
- Sanders, S.L., Jennings, J., Canutescu, A., Link, A.J., and Weil, P.A. (2002). Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol. Cell. Biol.* **22**, 4723–4738.
- Tartof, K.D., Hobbs, C., and Jones, M. (1984). A structural basis for variegating position effects. *Cell* **37**, 869–878.
- van der Knaap, J.A., Kumar, B.R., Moshkin, Y.M., Langenberg, K., Krijgsveld, J., Heck, A.J., Karch, F., and Verrijzer, C.P. (2005). GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7. *Mol. Cell* **17**, 695–707.
- vom Baur, E., Harbers, M., Um, S.J., Benecke, A., Chambon, P., and Losson, R. (1998). The yeast Ada complex mediates the ligand-dependent activation function AF-2 of retinoid X and estrogen receptors. *Genes Dev.* **12**, 1278–1289.
- Wallberg, A.E., Flinn, E.M., Gustafsson, J.A., and Wright, A.P. (2000). Recruitment of chromatin remodelling factors during gene activation via the glucocorticoid receptor N-terminal domain. *Biochem. Soc. Trans.* **28**, 410–414.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878.
- 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.
- Wu, P.Y., Ruhlmann, C., Winston, F., and Schultz, P. (2004). Molecular architecture of the *S. cerevisiae* SAGA complex. *Mol. Cell* **15**, 199–208.
- Yanagisawa, J., Kitagawa, H., Yanagida, M., Wada, O., Ogawa, S., Nakagomi, M., Oishi, H., Yamamoto, Y., Nagasawa, H., McMahon, S.B., et al. (2002). Nuclear receptor function requires a TFTC-type histone acetyl transferase complex. *Mol. Cell* **9**, 553–562.
- Zhang, Y. (2003). Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev.* **17**, 2733–2740.
- Zhu, P., Zhou, W., Wang, J., Puc, J., Ohgi, K.A., Erdjument-Bromage, H., Tempst, P., Glass, C.K., and Rosenfeld, M.G. (2007). A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol. Cell* **27**, 609–621.



## Hsp70 and Hsp40 Chaperones Do Not Modulate Retinal Phenotype in SCA7 Mice\* ♦

Received for publication, August 9, 2004, and in revised form, September 17, 2004  
Published, JBC Papers in Press, October 19, 2004, DOI 10.1074/jbc.M409062200

Dominique Helmlinger‡§, Jacques Bonnet‡, Jean-Louis Mandel‡, Yvon Trottier‡, and Didier Devys‡¶

From the ‡Department of Molecular Pathology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP/Collège de France, BP 10142, 67404 Illkirch Cedex, France

Nine neurodegenerative diseases, including spinocerebellar ataxia type 7 (SCA7), are caused by the expansion of polyglutamine stretches in the respective disease-causing proteins. A hallmark of these diseases is the aggregation of expanded polyglutamine-containing proteins in nuclear inclusions that also accumulate molecular chaperones and components of the ubiquitin-proteasome system. Manipulation of HSP70 and HSP40 chaperone levels has been shown to suppress aggregates in cellular models, prevent neuronal death in *Drosophila*, and improve to some extent neurological symptoms in mouse models. An important issue in mammals is the relative expression levels of toxic and putative rescuing proteins. Furthermore, overexpression of both HSP70 and its co-factor HSP40/HDJ2 has never been investigated in mice. We decided to address this question in a SCA7 transgenic mouse model that progressively develops retinopathy, similar to SCA7 patients. To co-express HSP70 and HDJ2 with the polyglutamine protein, in the same cell type, at comparable levels and with the same time course, we generated transgenic mice that express the heat shock proteins specifically in rod photoreceptors. While co-expression of HSP70 with its co-factor HDJ2 efficiently suppressed mutant ataxin-7 aggregation in transfected cells, they did not prevent either neuronal toxicity or aggregate formation in SCA7 mice. Furthermore, nuclear inclusions in SCA7 mice were composed of a cleaved mutant ataxin-7 fragment, whereas they contained the full-length protein in transfected cells. We propose that differences in the aggregation process might account for the different effects of chaperone overexpression in cellular and animal models of polyglutamine diseases.

neuronal loss in cerebellum, brainstem, and retina (1). SCA7 is caused by an abnormal expansion of a polyglutamine (poly(Q)) tract (38–460 repeats) in the SCA7 gene product ataxin-7 (2), a 892-amino acid protein recently identified as a new component of a multisubunit transcriptional complex (3). Eight other neurodegenerative diseases are caused by a CAG/poly(Q) repeat expansion, including Huntington disease (HD), dentatorubro-pallidolysian atrophy, spinobulbar muscular atrophy (SBMA) and spinocerebellar ataxia types 1, 2, 3, 6, and 17. Genetic and molecular studies indicate that poly(Q) tracts confer a novel toxic function to the otherwise unrelated proteins, but the mechanisms by which poly(Q) expansions lead to neurodegeneration remain unclear (4, 5).

All of these diseases are characterized by continuous accumulation of mutant proteins in insoluble aggregates, typically forming nuclear inclusions (NIs) (6). NIs have been shown to stain positively for ubiquitin, chaperones (mainly heat shock proteins), and proteasome subunits, in various cellular and animal models of poly(Q) diseases, as well as in post-mortem patient brains (7, 8). Several SCA7 mouse models, which reproduce many features of the human situation, display numerous NIs consistently stained for ubiquitin, proteasome subunits, and chaperones (HSP70, HSC70, and HDJ2) (9–12). In SCA7 patient brains, HDJ2, ubiquitin, and 19 S proteasome subunits co-localize with NIs (13–15). Furthermore, the levels of several chaperones, including the HSP70 and the HSP40 chaperones HDJ1 and HDJ2, were shown to be reduced in SCA7 and/or HD mouse models (16, 17). Altogether, these findings suggested that protein misfolding and impaired clearance might be a common pathogenic event in poly(Q) diseases. Accordingly, several studies showed that overexpression of chaperones in cells reduces aggregate formation and, in some cases, suppresses poly(Q) toxicity (reviewed in Refs. 8 and 18). Genetic screens, in *Drosophila* models overexpressing mutant ataxin-1 or a pure 127Q stretch, identified molecular chaperones and components of the ubiquitin-proteasome pathway as strong modulators of poly(Q)-induced toxicity (19, 20). In particular, deletion of HSP70 genes exacerbated SCA1 or SBMA fly phenotypes, whereas overexpression of HSP70 or HDJ1 suppressed neurodegeneration in various *Drosophila* models (19–24). Furthermore, Chan *et al.* (25) showed that HSP70 and HDJ1 synergistically suppressed poly(Q) toxicity in a SCA3 fly model. These chaperones work together in an ATP-dependent manner, notably to help refold denatured and aggregated proteins. HSP40 co-chaperones recognize abnormally folded polypeptide substrates, present them to HSP70, and stimulate HSP70 ATPase activity (26).

Spinocerebellar ataxia type 7 (SCA7)<sup>1</sup> is a dominantly inherited neurodegenerative disorder, characterized by late-onset

\* This work was supported by grants from the Hereditary Disease Foundation and from the European Community (EUROSCA integrated Project LSHM-CT-2004-503304), funds from the Institut National de la Santé et de la Recherche Médicale, the Centre National de la Recherche Scientifique, and the Hôpital Universitaire de Strasbourg (HUS). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

♦ This article was selected as a Paper of the Week.

§ Supported by a fellowship from the Fondation pour la Recherche Médicale.

¶ To whom correspondence should be addressed. Tel.: 33-388-65-34-13; Fax: 33-388-65-32-46; E-mail: devys@igbmc.u-strasbg.fr.

<sup>1</sup> The abbreviations used are: SCA7, spinocerebellar ataxia type 7; poly(Q), polyglutamine; HD, Huntington disease; SBMA, spinobulbar muscular atrophy; NI, nuclear inclusion; HSP (or Hsp), heat shock

protein; ERG, electroretinogram; HA, hemagglutinin; DAPI, 4,6-diamidino-2-phenylindole; RT, reverse transcription; WT, wild-type.

Overexpression of HSP70 had a more mitigated outcome in modulating poly(Q) toxicity in mouse models (17, 27, 28). The authors hypothesized that this might be explained by the lack of HSP40 induction that would be required for HSP70 to modulate neurodegeneration. Overexpression of both HSP70 and HSP40 has not been investigated in mouse models of poly(Q) diseases. We decided to address this question in a SCA7 transgenic mouse model that presents with a progressive and severe retinopathy (9, 29). In this model, neuronal dysfunction of a single cell type (rod photoreceptors) can be quantified on living animals by recording electroretinograms (ERGs). To co-express mutant ataxin-7 and the molecular chaperones in the same cell type, we generated transgenic mice overexpressing HSP70 and HDJ2 specifically in rod photoreceptors. These heat shock proteins efficiently suppressed mutant ataxin-7 aggregation in transfected cells but were not able to prevent either neuronal toxicity or aggregate formation in photoreceptors from SCA7 mice. We show here for the first time that expressing both HSP70 and its co-factor HDJ2 has no beneficial effect on phenotype progression in a poly(Q) mouse model.

#### EXPERIMENTAL PROCEDURES

**Generation of Transgenic Animals**—A 2.2-kb XbaI-SmaI fragment of the human rhodopsin promoter, encompassing the distal enhancer region, was subcloned from the R7N construct (9) into the pGS<sub>2</sub> vector, kindly provided by Dr. D. Metzger, upstream of the rabbit  $\beta$ -globin intron II, to generate the pGS-Rho vector. The 1.9-kb human HSP70 cDNA was amplified by PCR from clone pH2.3, kindly provided by Dr. R. Morimoto, using a 5' primer encoding a FLAG epitope. The 1.2-kb human HDJ2 cDNA was amplified by PCR from clone J3, kindly provided by Dr. T. Mohanakumar, using a 3' primer encoding a HA epitope. These PCR fragments were XhoI-NheI-digested and cloned into the pGS-Rho vector to generate pGS-Rho-RH70 and pGS-Rho-RH40 transgene constructs, respectively. All final plasmids were verified by sequencing and subsequently SalI-ApaLI-digested to remove bacterial sequences, purified on sucrose gradients, and microinjected into FVB/N-fertilized eggs, as described (9). Mouse tail DNA was screened by PCR for the presence of the transgenes. Founder mice were backcrossed on the inbred C57BL/6 background to get F3 RH40 or RH70 transgenic mice with a mixed (87.5% C57BL/6; 12.5% FVB/N) genetic background. The loss of the *rd* allele from the FVB/N background was checked by PCR followed by DpnI digestion. For genotyping of R7E, RH70, and RH40 transgenic mice, mouse tail DNA was screened by PCR using specific primers (5'-AGGCGATGACAAAGAAGAG-3' and 5'-AGGGGATGAGGATGAAGAG-3' for the SCA7 transgene, 470-bp product; 5'-CAACAAGATCACCATCACAAC-3' and 5'-GCTATTGCTTTATTGT-AACC-3' for the HSP70 transgene, 500-bp product; 5'-GGGGTCAAA-CCCAATGCTACTC-3' and 5'-CATCCTTCCTCCTCCTCAAAA-3' for the HDJ2 transgene, 260-bp product). For genotyping of double and triple transgenic mice, two independent PCR reactions were performed on the same mouse tail DNA sample; one including both HSP70 and HDJ2 primer pairs and another using primers specific for the SCA7 transgene.

**Plasmid Construction and Cell Transfection**—For cell transfection experiments, a full-length SCA7 cDNA with 128 CAGs was obtained by PCR amplification of the B7E2B construct (10), using a 5' primer encoding a FLAG epitope. This fragment was subcloned into pcDNA3.1(+) (Invitrogen) to generate pc7E2FL. A truncated SCA7 construct (pc7E2-456) was generated by PCR amplification of pc7E2FL and introduced into pcDNA3.1(+) vector (Invitrogen). Tagged full-length HDJ2 and HSP70 cDNAs were subcloned from pGS-Rho-RH70 and pGS-Rho-RH40 constructs into pcDNA3.1(+) (Invitrogen) to generate pcHDJ2 and pcHSP70 constructs, respectively.  $2 \times 10^5$  HEK293T cells were transiently transfected by calcium phosphate precipitation, washed 16 h later, and harvested or fixed 24 or 48 h after washing.

**Immunoblot Blot Analysis**—Whole cell extracts from dissected retina or from transfected HEK293T cells were obtained by homogenization in lysis buffer containing 50 mM Tris-HCl, pH 8.0, 10% glycerol, 5 mM EDTA, 150 mM KCl, 0.5% Nonidet P-40, 1 mM phenylmethylsulfonyl fluoride, and a mixture of protease inhibitors. Homogenates were then incubated for 20 min on ice and centrifuged for 30 min at  $10,000 \times g$  and 4°C. Supernatants were resolved on 8% SDS-PAGE gels. Primary antibodies used were mouse monoclonal anti-FLAG M2 (Sigma), anti-HA 12CA5, anti-HSP70 SPA-810 (Stressgen, Victoria, Canada), and anti-HDJ2/Hsp40 (Neomarkers, Fremont, CA).

**Antibody Production**—The anti-ataxin-7 polyclonal 1599 antibody was generated by immunization of rabbits with a peptide corresponding to amino acids 226–245 (KEKLQLRGNTRPMHPHPIQQSR) of human ataxin-7 and purified as described previously (9).

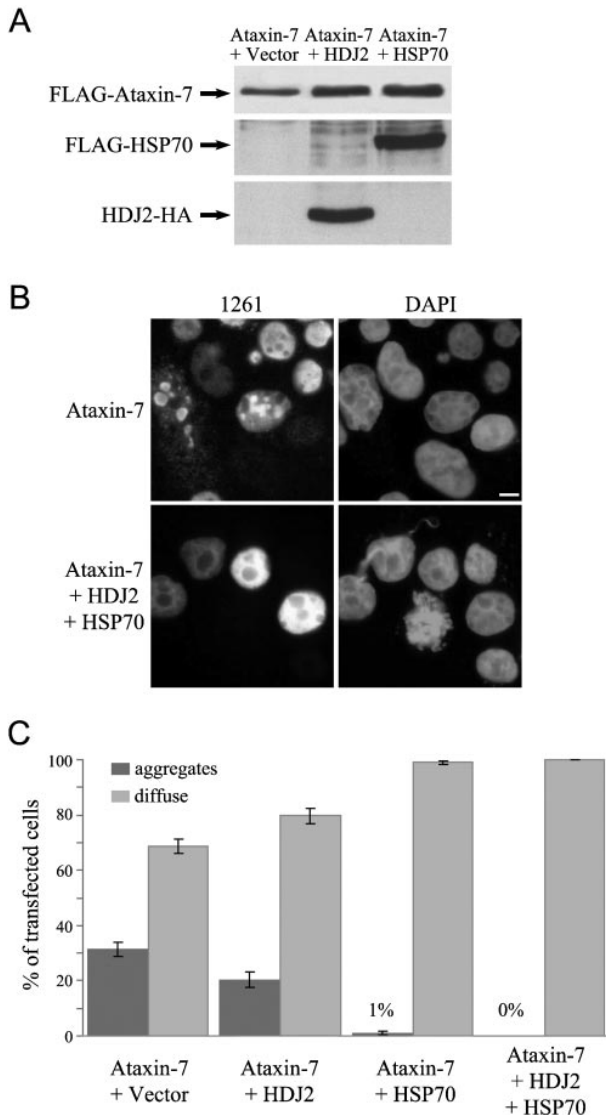
**Immunohistology and Light Microscopy**—For histology, retina were dissected from enucleated eyes and then fixed by immersion in 2.5% glutaraldehyde. Fixed eyes were rinsed in phosphate-buffered saline, dehydrated with a graded alcohol series, and embedded in Epon. Semithin (1  $\mu$ m) sections were stained with toluidin blue. For immunocytofluorescence, transfected HEK293T cells were fixed in 4% paraformaldehyde and stained with anti-ataxin-7 2A10 monoclonal and affinity-purified 1261 and 1598 polyclonal antibodies, as described (9, 10). Immunohistochemistry was performed as described previously (29). Briefly, dissected retina were fixed in 4% paraformaldehyde, cryoprotected, and frozen in Cryomatrix (Thermo Shandon). 10- $\mu$ m cryostat sections were permeabilized, blocked, and then immunostained using the following primary antibodies: anti-ataxin-7 2A10 monoclonal and affinity-purified 1261, 1599, and 1598 polyclonal antibodies, as described (9, 10), anti-FLAG M2 (Sigma), anti-HA 12CA5, anti-HSP70 SPA-810 (Stressgen), and anti-HDJ2/Hsp40 (Neomarkers, Fremont, CA) mouse monoclonal antibodies. Secondary antibodies used were CY3- and Oregon Green-conjugated goat anti-rabbit and anti-mouse IgG (Jackson ImmunoResearch, West Grove, PA). Nuclei were counterstained with 0.5  $\mu$ g/ml 4,6-diaminido-2-phenylindole (DAPI).

**Real-time RT-PCR Analysis of Retinal RNA**—Real-time RT-PCR analysis was performed as described previously (29). Briefly, reverse-transcribed total retinal RNA was amplified with SYBR Green I (Sigma) on a Light-Cycler instrument (Roche Applied Science, Basel, Switzerland) using *Rho*- and *Ppia*-specific primers.

**Electroretinogram Recordings**—Electroretinographic analysis was performed as described previously (29). Briefly, dark-adapted mice were anesthetized, pupils were dilated, and the cornea was locally anesthetized. Scotopic rod ERGs were recorded by using a gold loop electrode placed on the corneal surface and calibrated to a reference electrode inserted subcutaneously on the head. Light stimulation was increased from  $-4$  to  $1.4$  log candela s m<sup>-2</sup>, as measured with a luxmeter at eye level. The duration of the light stimulus was constant. Responses were digitized using a data acquisition labmaster board (Multiliner Vision).

#### RESULTS

**Molecular Chaperones Suppress Formation of Aggregates in Cells Expressing Mutant Ataxin-7**—We first evaluated whether overexpression of HSP70 and/or HSP40/HDJ2 could modulate mutant ataxin-7 aggregation in mammalian cells. We transfected HEK293T cells with truncated or full-length mutant SCA7 cDNA constructs (128 CAG) either alone or together with human HSP70- and HDJ2-tagged cDNA constructs. Fourty hours after transfection, Western blots revealed high and comparable expression of FLAG-tagged mutant ataxin-7, FLAG-tagged HSP70, and HA-tagged HDJ2 (Fig. 1A). We then performed immunofluorescence staining of transfected cells using an N-terminal anti-ataxin-7 antibody (1261 (9)). While truncated (residues 1–456) and full-length normal ataxin-7 displayed a diffuse nuclear staining (3), mutant ataxin-7 either showed a similar localization or accumulated into large NIs (Fig. 1B, upper panels). Co-transfection of both HDJ2 and HSP70 cDNAs with truncated (residues 1–456) or full-length mutant SCA7 completely suppressed formation of NIs (Fig. 1B, lower panels and data not shown). All transfected cells displayed a diffuse nuclear staining suggesting that overexpression of these chaperones inhibited mutant ataxin-7 aggregation. We then quantified this effect by counting cells displaying either NIs or homogenous nuclear staining after transfecting either truncated mutant SCA7 cDNA alone or together with HDJ2, with HSP70, or with both HDJ2 and HSP70 cDNAs. Overexpression of HDJ2 had a mild effect on the number of aggregates ( $20.3\% \pm 2.6\%$  versus  $31.3\% \pm 2.7\%$ ) that was not statistically significant (one-way analysis of variance,  $F_{(2,14)} = 1.6$ ,  $p > 0.05$ ). Drastic suppression of aggregate formation was achieved by overexpressing HSP70 alone ( $1.0\% \pm 0.5\%$  versus  $31.3\% \pm 2.7\%$ ; one-way analysis of variance,  $F_{(3,14)} = 4.8$ ,  $p < 0.05$ ), or together with its co-factor HDJ2 (Fig. 1C). These



**FIG. 1. HSP70 and HSP40/HDJ2 chaperones can prevent mutant ataxin-7 aggregation in transfected cells.** *A*, HEK293T cells were co-transfected with a full-length mutant *SCA7* cDNA construct (pc7E2FL) and either an empty vector, a human *HDJ2* HA-tagged cDNA construct (pcHDJ2), or a human *HSP70* FLAG-tagged cDNA construct (pcHSP70). Whole cell extracts were prepared 40 h after transfection and analyzed by immunoblotting using an anti-FLAG antibody (upper and middle panels) or an anti-HA antibody (lower panel). *B*, HEK293T cells were co-transfected with pc7E2FL and with either an empty vector or both pcHDJ2 and pcHSP70 constructs (ratio pc7E2FL:pcHDJ2:pcHSP70 1:1:2). Immunolocalization of transiently expressed mutant ataxin-7 (128 Gln) using an anti-ataxin-7 antibody (1261, left panels) was performed 24 h after transfection and showed that a significant proportion of transfected cells had nuclear inclusions. When both HDJ2 and HSP70 chaperones were overexpressed, mutant ataxin-7 resulted in homogenous nuclear staining, whereas almost no aggregates could be detected. Nuclei were counterstained with DAPI (right panels). The scale bar represents 5  $\mu$ m in each panel. *C*, HEK293T cells were co-transfected with a truncated mutant *SCA7* cDNA construct (pc7E2-456) and either with an empty vector or pcHDJ2, pcHSP70, or both constructs. Immunolocalization of ataxin-7 was performed 24 h after transfection and the percentage of transfected cells with nuclear inclusions (dark gray bars) or displaying diffuse nuclear staining (light gray bars) was counted in each condition. Each bar represents the mean  $\pm$  S.E. of at least three independent experiments.

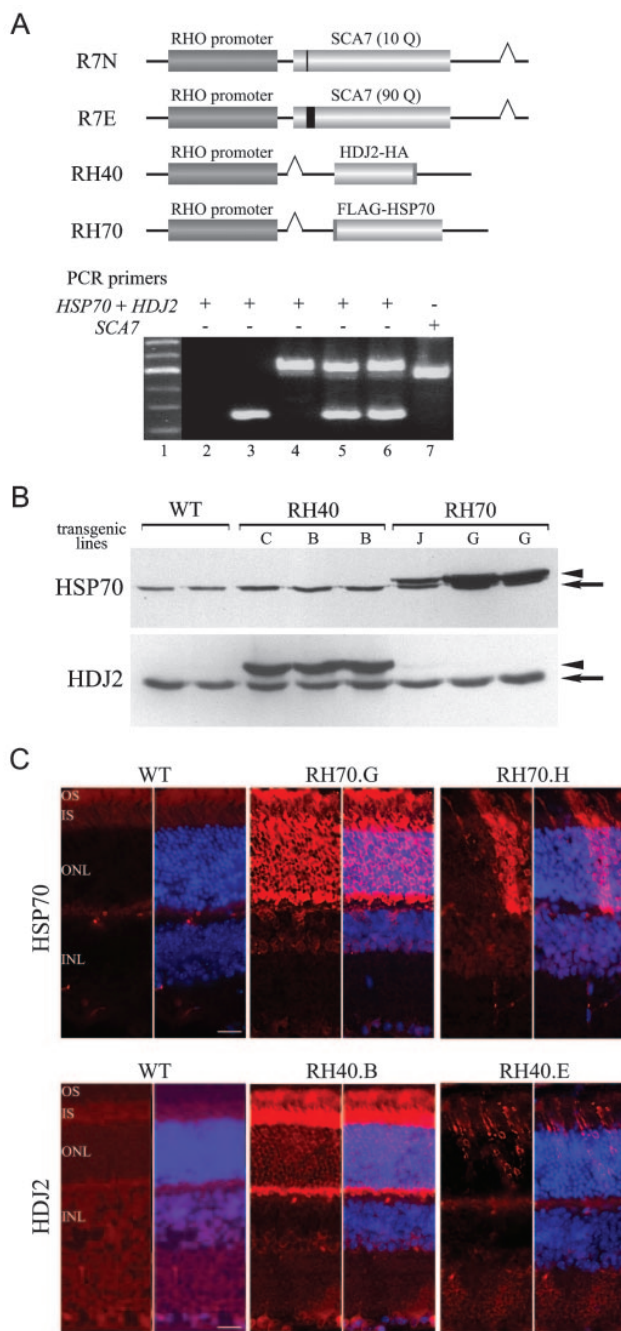
results show that mutant ataxin-7 aggregation can be suppressed in transfected cells by manipulating HSP70 and HDJ2 chaperone expression levels.

**Generation of Transgenic Lines Overexpressing HDJ2 or HSP70 in the Retina**—We then wanted to address whether HSP70 and/or HDJ2 overexpression could modulate mutant ataxin-7 aggregation and neuronal toxicity in a *SCA7* mouse model. We previously generated *SCA7* transgenic mice (R7E) that highly and specifically overexpress mutant ataxin-7 (90Q) in a single cell type, rod photoreceptors. These mice (from the R7E.A line) have an early and severe retinal dysfunction, as revealed by a progressive decrease of rod photoreceptor electrophysiological activity (9, 29).

We thus made new transgenic mice that co-express molecular chaperones and mutant ataxin-7 in the same cells, at comparable levels and with the same time course. We constructed expression vectors with human *HSP70* or *HDJ2* cDNAs downstream from the same 2.2 kb of the human rhodopsin promoter used to drive mutant ataxin-7 expression. To assess recombinant chaperone expression, we FLAG-tagged *HSP70* and HA-tagged *HDJ2* cDNAs (Fig. 2A). The FLAG tag was inserted at the N-terminal end of HSP70 to preserve the integrity of its C-terminal co-chaperone binding domain, while the HA tag was inserted at the C-terminal end of HDJ2 to avoid interference with the structure of its N-terminal J domain (30, 31). We generated several transgenic lines containing *HSP70* or *HDJ2* transgenes and called RH70 and RH40, respectively. They express the recombinant chaperones at different levels relative to the endogenous proteins (Fig. 2B and Table I), as determined by Western blot analysis of retinal homogenates using anti-HSP70 or anti-HDJ2 antibodies. RH40 line B and RH70 line G showed the highest expression, reaching 5–10-fold the endogenous levels. Expression levels of endogenous Hsp70 and Hdj2 were not modified in transgenic mice compared with wild type (Fig. 2B). Immunostaining with either anti-tag or anti-chaperone antibodies indicated that expression was high and clearly restricted to rod photoreceptors (Fig. 2C). In RH40 line E and RH70 line H, we observed strong expression of recombinant chaperones confined to a small subset of rods, accounting for moderate levels detected by Western blots. Finally, chaperone overexpression did not affect retinal structure and function, as revealed by histological examination and electroretinogram recordings (Fig. 2C and data not shown).

**HSP70/HDJ2 Overexpression, Alone or in Combination, Does Not Modulate SCA7 Retinopathy**—F2 hemizygous animals from RH70 line G and RH40 line B were cross-bred with hemizygous R7E mice to generate double transgenic mice R7E/RH70 and R7E/RH40, respectively. This breeding scheme results in all genotypes (wild-type, R7E, RH70, RH40, and double transgenic mice) from the same litter. Rod photoreceptor function was quantified on living animals at 9 weeks of age by measuring the amplitude of the a-wave from electroretinogram recordings performed on dark-adapted mice (scotopic ERGs). At this age, RH70 and RH40 mice displayed similar ERG recordings to that of wild-type littermates (data not shown). We previously showed that rod activity was reduced by around 65% in 9-week-old R7E mice compared with wild-type littermates. These results were obtained with mice maintained on a pure C57BL/6 genetic background (29). As our breeding scheme generated mice with a mixed background (87.5% C57BL/6, 12.5% FVB/N, see “Experimental Procedures”), we first performed ERG recordings on R7E and control mice from these litters. Rod response was only reduced by 40% in 9-week-old R7E mice as compared with control (CT) littermates, indicating a slightly less severe disease progression on this mixed background (Fig. 3).

A comparable reduction of rod response was observed between age-matched R7E ( $106 \pm 11 \mu$ V), R7E/RH40 ( $120 \pm 16 \mu$ V), and R7E/RH70 ( $94 \pm 14 \mu$ V) mice as compared with



**FIG. 2. Expression of recombinant HSP70 and HDJ2 in transgenic mouse retina.** *A*, schematic view of the transgene constructs generated and used in this study. A representative PCR genotyping of the different transgenic mice is shown below and demonstrates unambiguous determination of all genotypes analyzed in this study (lane 2, wild type; lane 3, RH40; lane 4, RH70; lane 5, RH40/RH70 double transgenic; lanes 6 and 7, R7E/RH40/RH70 triple transgenic animal; lane 1, 100-bp DNA ladder). *B*, analysis of protein expression in 1-month-old WT, RH40 (lines C and B), and RH70 (lines J and G) transgenic mice. Western blots of retinal homogenates were probed with anti-HSP70 or anti-HDJ2 antibodies. Recombinant tagged HSP70 or HDJ2 (arrowheads) were detected above endogenous Hsp70 or Hdj2 (arrows) allowing to estimate recombinant proteins expression levels relative to endogenous chaperones levels. *C*, immunofluorescence of retina from 1-month-old WT, RH40 (lines B and E), and RH70 (lines G and H) transgenic animals using anti-HSP70 or anti-HDJ2 antibodies. Faint staining was detected in wild-type retina (left panels). In contrast, strong HSP70 and HDJ2 immunoreactivities were found exclusively in photoreceptors of RH70.G and RH40.B retina (middle panels). Finally, strong HSP70 and HDJ2 immunoreactivities were found in a very restricted

**TABLE I**  
Summary of the different HSP70/HDJ2 transgenic mouse lines generated

Construct	No. founders	Lines <sup>a</sup>	Protein expression <sup>b</sup>		
RH40	12	A	++ <sup>c</sup>		
		B	+++		
		C	+++		
		D	+		
		E	+		
		F	++		
RH70	16	G	+++		
		H	+		
		I	++		
		J	++		

<sup>a</sup> Only lines that stably transmitted the transgene over three generations and expressed detectable levels of protein are presented.

<sup>b</sup> Estimated by Western blot analysis using anti-HDJ2 and anti-HA antibodies or anti-HSP70 and anti-FLAG antibodies.

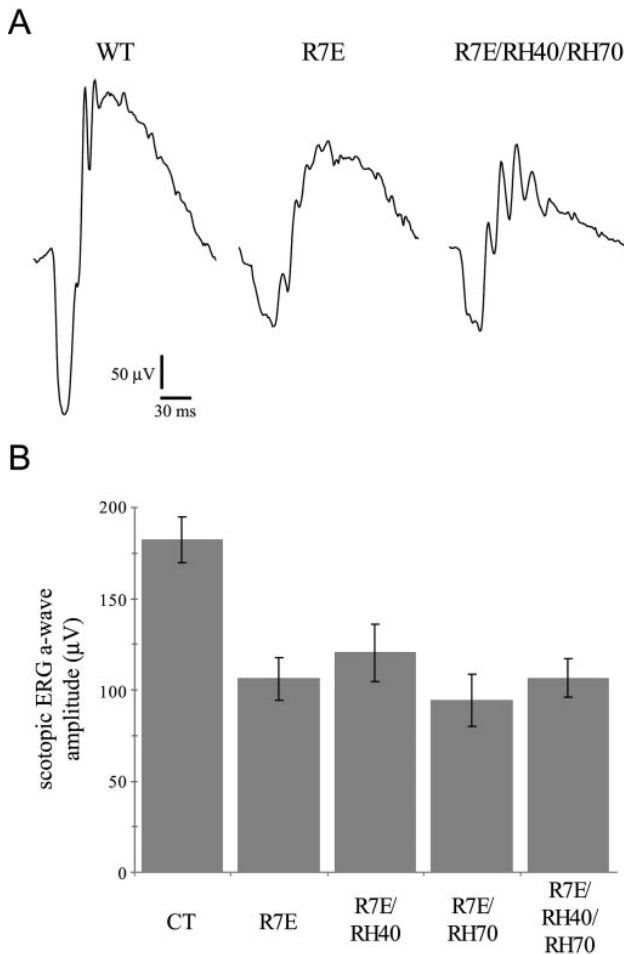
<sup>c</sup> Recombinant/endogenous chaperones expression ratio: +++, >5-fold; ++, around 1-fold; +, <1-fold.

control littermates ( $182 \pm 13 \mu\text{V}$ ) (Fig. 3B). As it has been previously shown that HSP70 and HDJ2 act synergistically in preventing polyQ-induced toxicity (25), we then generated triple transgenic mice to overexpress both HSP70 and its co-factor HDJ2 in rod photoreceptors of R7E mice. At 9 weeks of age, ERG recordings showed that rod photoreceptor responses were comparable in R7E/RH40/RH70 ( $107 \pm 11 \mu\text{V}$ ) and R7E ( $106 \pm 11 \mu\text{V}$ ) mice (Fig. 3).

We previously reported that progressive reduction of rod photoreceptor activity was accompanied by dramatic thinning of the segment layer in R7E mice. We assessed progression of retinopathy by histological examination of retina from the 10-week-old R7E, R7E/RH40/RH70 triple transgenic, and WT mice that had been used for the ERG recordings. Toluidin blue-stained retinal sections from triple transgenic R7E/RH40/RH70 or R7E mice were indistinguishable (Fig. 4A). At this age, mice of both genotypes displayed heterogeneous pattern of segment loss, accounting for the wavy shape adopted by the photoreceptor nuclear layer (ONL). Massive loss of the outer segments is likely due to a failure of rods to renew their segments. Indeed, we previously reported an early down-regulation of rhodopsin expression in R7E mice (29). We thus quantified *Rho* mRNA levels by performing quantitative RT-PCR on RNA extracted from R7E, R7E/RH40/RH70 triple transgenic, and WT retina. At 10 weeks of age, *Rho* mRNA levels were reduced to  $9\% \pm 2\%$  in retinal extracts from R7E mice compared with wild-type mice. A similar down-regulation of *Rho* expression was observed in retina from R7E/RH40/RH70 mice ( $10\% \pm 4\%$ ) (Fig. 4B). Taken together, these results indicate that high and specific overexpression of either HSP70, HDJ2, or both molecular chaperones does not overcome rod photoreceptor dysfunction and is not sufficient to counteract poly(Q)-induced toxicity in this SCA7 model.

We previously showed that the mutant SCA7 transgene in R7E mice, which is controlled by the rhodopsin promoter, is transiently expressed with less than 5% residual expression from 6 weeks of age (29). We wondered whether recombinant chaperone expression would be affected in R7E/RH40 or in R7E/RH70 mice. Western blot analysis of retinal extracts showed a 2-fold reduction of recombinant HDJ2 levels in 10-

number of photoreceptors in RH70.H and RH40.E retina (right panels). Both recombinant and endogenous HSP70 and HDJ2 were found in the photoreceptors cytoplasm, particularly within the outer and the inner segments, respectively. Each panel is divided in two showing HDJ2/HSP70 immunoreactivities (left) or merged image after counterstaining with DAPI (right). OS, outer segment; IS, inner segment; ONL, outer nuclear layer; INL, inner nuclear layer. The scale bar represents 20  $\mu\text{m}$  in each panel.



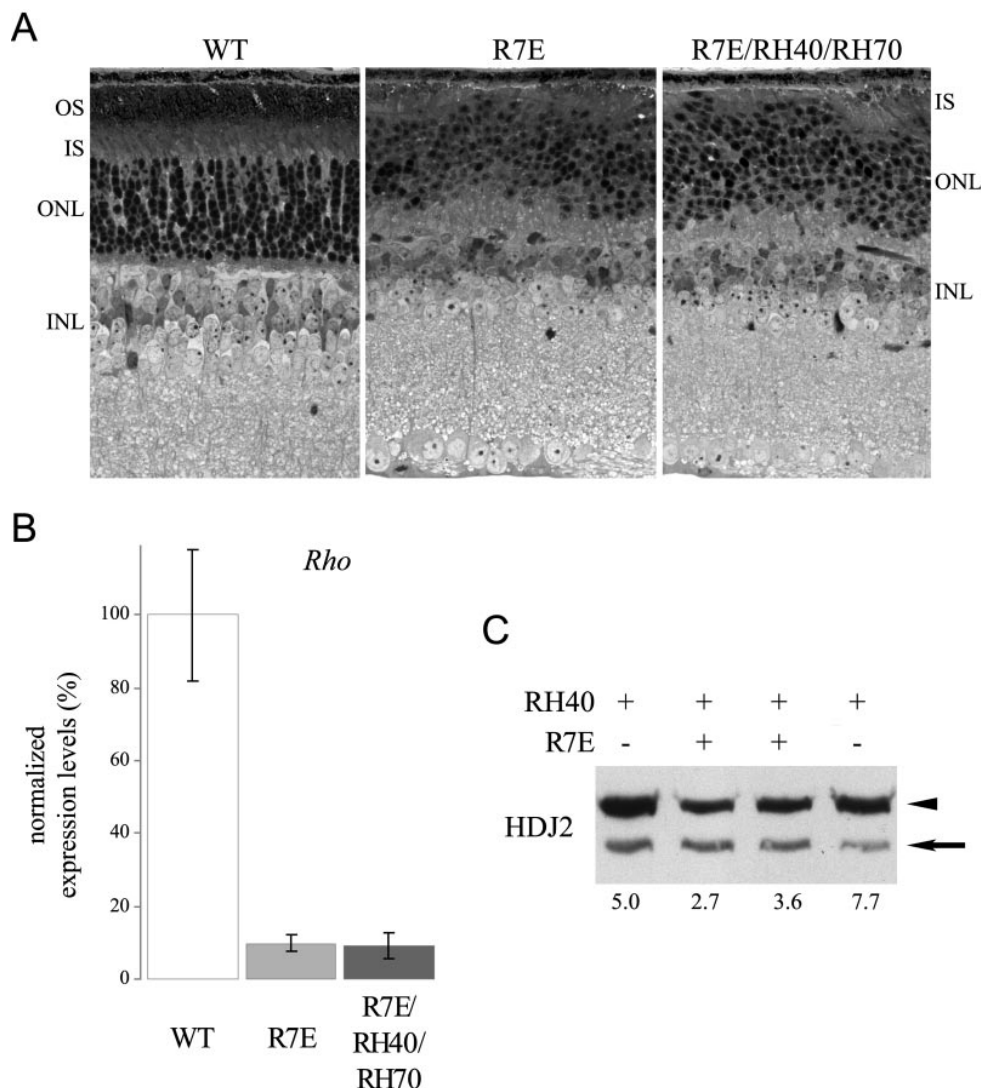
**FIG. 3. Overexpression of HSP70 and HDJ2 does not modulate rod photoreceptors dysfunction in SCA7 mice.** *A*, scotopic ERG responses of WT, R7E, and R7E/RH40/RH70 triple transgenic mice at 9 weeks of age. Each recording corresponds to the average response of one animal at maximal light intensities (from 0 to 1.4 log candela  $\text{s m}^{-2}$ ). Single recordings revealed a similar decrease of rod electrophysiological activity in R7E and R7E/RH40/RH70 mice. *B*, scotopic ERG responses from R7E, R7E/RH40, R7E/RH70, R7E/RH40/RH70 transgenic mice and control littermates (CT) at 9 weeks of age ( $n = 6-12$  for each group). Rod photoreceptors activity is quantified as the mean amplitude of the a-wave recorded at the four maximal light stimulus intensities (from 0 to 1.4 log candela  $\text{s m}^{-2}$ ). Quantification of rod dysfunction revealed a similar decline in electrophysiological activity in all the genotypes studied compared with control mice. Each bar represents the mean value  $\pm$  S.E. of 6–12 animals in each group.

week-old R7E/RH40 double transgenic mice, compared with RH40 littermates (Fig. 4C). Recombinant HDJ2 levels were still higher than endogenous Hdj2 in R7E/RH40 mice (Fig. 4C). Similar observations were made for recombinant HSP70 levels, as a FLAG antibody detected a significant signal in R7E/RH40/RH70 mice (Fig. 5D). Thus, the decrease of recombinant chaperone expression was less severe than SCA7 transgene repression in R7E mice, indicating that the absence of a protective effect is unlikely to be due to HSP70/HDJ2 down-regulation.

**HSP70/HDJ2 Overexpression Has No Effect on Mutant Ataxin-7 Aggregation *In Vivo***—Despite the lack of functional improvement in R7E/RH40/RH70 triple transgenic mice, we examined whether mutant ataxin-7 aggregation is modulated by overexpression of these molecular chaperones. We performed immunostaining using an anti-ataxin-7 antibody to detect NIs in retinal cryosections from 10-week-old R7E and

R7E/RH40/RH70 triple transgenic mice. In marked contrast with what was observed in transfected cells, NI content was indistinguishable between R7E (Fig. 5, *A* and *B*) and R7E/RH40/RH70 (Fig. 5, *C* and *D*) retina. At this age, all ataxin-7 immunoreactivity is concentrated into a single, intense nuclear aggregate in most rod nuclei. These NIs were distributed throughout the whole photoreceptor nuclear layer and were of similar size in R7E and R7E/RH40/RH70 retina. Immunostaining using anti-HA or anti-FLAG antibodies indicated that recombinant HDJ2 or HSP70 were not recruited into NIs (Fig. 5, *C* and *D*). These findings indicate that overexpression of HSP70 together with its co-factor HDJ2 is not sufficient to prevent aggregate formation in this SCA7 mouse model. The differential effect of molecular chaperones on mutant ataxin-7 aggregation between cellular and mouse models suggest that the aggregation process is different *in vitro* and *in vivo*.

**Proteolytic Processing and Aggregate Formation of Mutant Ataxin-7**—As NI formation in R7E mice is not suppressed by HSP70/HDJ2 overexpression, we addressed what might differentiate mutant ataxin-7 aggregation in transfected cells *versus* transgenic mice. Yvert *et al.* (9) previously reported that NIs are composed of an N-terminal fragment of mutant ataxin-7 in R7E mice. Proteolytic processing of mutant ataxin-7 was confirmed in another transgenic SCA7 model and in the brain of a SCA7 patient (14, 32). We thus examined whether mutant ataxin-7 is cleaved in transfected HEK293T cells. We performed immunostaining using anti-ataxin-7 antibodies that recognize either N-terminal (2A10) or C-terminal (1598) epitopes 48 h after transfecting vector that express full-length mutant SCA7 cDNA carrying 128 CAGs. Both antibodies stained mutant ataxin-7, whether it displayed homogenous nuclear staining or was aggregated into multiple NIs (Fig. 6B). This result indicates that, in transfected cells, NIs are formed by full-length mutant ataxin-7. Similarly, aggregation of full-length ataxin-7 in NIs was independently reported by Zander *et al.* (14). We next assessed whether mutant ataxin-7 proteolytic processing occurred in R7E/RH40/RH70 triple transgenic mice. To confirm and better characterize this proteolytic event, we generated a new polyclonal antibody (1599) raised against a synthetic peptide corresponding to amino acids 226–245 of human ataxin-7. Western blots of retinal homogenates from R7N mice (R7N.C line), which overexpress normal ataxin-7 (10Q) in rod photoreceptors, were probed with the purified 1599 antibody. Full-length recombinant ataxin-7 was detected in whole cell extracts from both R7N retina and cells transfected with SCA7 cDNA constructs (Fig. 6D). The 1599 antibody also detected a cross-reacting band in both WT and R7N animals. However, immunostaining of retina cryosections showed that this antibody specifically detected recombinant ataxin-7 in rod photoreceptor nuclei in R7N (Fig. 6C) but not in WT mice (data not shown), similarly to what was observed using other specific anti-ataxin-7 antibodies (Fig. 6C, *upper panels*). In both R7E and R7E/RH40/RH70 mice, an N-terminal antibody (1261) stained NIs in most rod photoreceptors, while the two more C-terminal antibodies (1599 and 1598), recognizing different epitopes, failed to detect NIs (Fig. 6C, *middle and lower panels*). These results indicate that NIs are composed of an N-terminal fragment of mutant ataxin-7 and that this processing is not affected by HSP70/HDJ2 overexpression. In transfected cells, the aggregation process is markedly different, as indicated by the fact that NIs are composed of the full-length mutant protein and that chaperone overexpression could suppress full-length or truncated mutant ataxin-7 aggregation. These discrepancies suggest that mutant ataxin-7 processing is markedly different *in vitro* from the *in vivo* situation and might provide one explanation for why chaperone overexpression has



**FIG. 4. Absence of rescue at the histological and molecular levels.** *A*, histological examination of WT, R7E, and R7E/RH40/RH70 triple transgenic mice at 10 weeks of age. Semi-thin toluidin blue-stained sections revealed that the photoreceptor nuclear layer (*ONL*) is disrupted. There are numerous "waves" that correspond to focal loss of the outer segments (*OS*) and thinning of the inner segments (*IS*) in both R7E and R7E/RH40/RH70 mice. *OS*, outer segment; *IS*, inner segment; *ONL*, outer nuclear layer; *INL*, inner nuclear layer. *B*, rhodopsin mRNA levels are severely reduced in both R7E and R7E/RH40/RH70 mice. Quantification of *Rho* expression in control, R7E, and R7E/RH40/RH70 mice at 10 weeks of age ( $n = 4$  for each group) by real-time RT-PCR using primers specific for *Rho* and *Ppia* as an internal control. Down-regulation of *Rho* mRNA is comparable in R7E and R7E/RH40/RH70 mice. *Rho* levels are represented as a percentage of the mean of control (*CT*) littermate mice, after normalization to *Ppia* levels. Each bar represents the mean value  $\pm$  S.E. *C*, analysis of recombinant HDJ2 expression in 10-week-old RH40 and R7E/RH40 mice. The Western blot of retinal homogenates was probed with an anti-HDJ2 antibody and quantified by densitometric analysis. Recombinant HDJ2 (arrowhead) was normalized to endogenous Hdj2 (arrow) levels. Ratios are indicated at the bottom of the gel for each mouse. There is a 2-fold reduction of recombinant HDJ2 levels in the R7E background at this age.

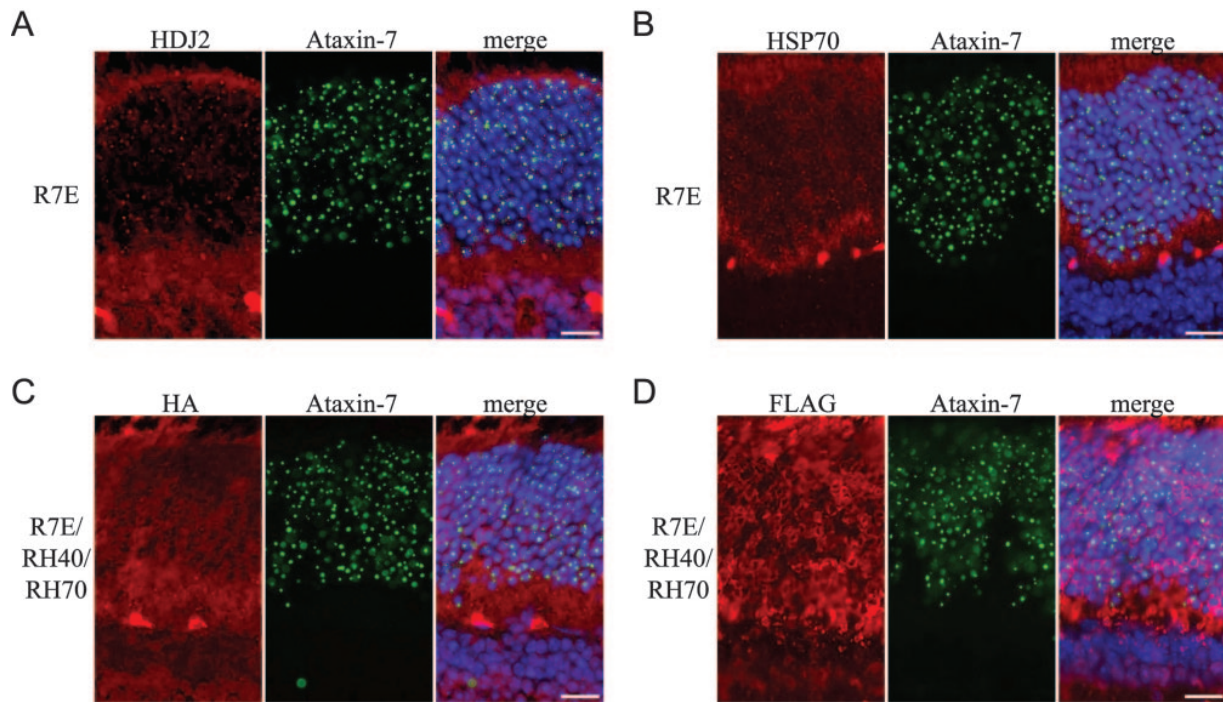
no beneficial effects on aggregate formation and eventually on neuronal dysfunction in this SCA7 model.

#### DISCUSSION

This report demonstrates that high level expression of HSP70 together with its co-factor HDJ2 does not afford protection against poly(Q)-induced toxicity in a mouse model of a poly(Q) disease. The retinopathy in our SCA7 mouse model offers an excellent opportunity to reproducibly quantify the progressive dysfunction of a single cell type and thus to assess the protective effects of chaperone overexpression in mice. We thus generated transgenic lines expressing HSP70 and/or HDJ2 at high levels specifically in rod photoreceptors and backcrossed them with R7E mice. Rod dysfunction was not overcome in R7E mice overexpressing both HSP70 and HDJ2,

as assessed by the similar reduction of ERG responses and the comparable histological and transcriptional abnormalities.

Several control experiments were performed to check that the absence of rescue was not due to flaws in our experimental design. First, we controlled that our tagged *HSP70/HDJ2* transgenes retained a refolding activity. These constructs were indeed able to interfere with mutant ataxin-7 aggregation, as shown by the complete inhibition of NI formation after overexpressing FLAG-tagged HSP70 and HA-tagged HDJ2 in transfected HEK293T cells (Fig. 1). Furthermore, this observation indicates that mutant ataxin-7 misfolding and aggregation can be prevented by molecular chaperones in cells, as observed for other poly(Q)-containing proteins (7, 33–43). Then, we confirmed that rod photoreceptors can simultaneously express multiple transgenes controlled by the same promoter by generating



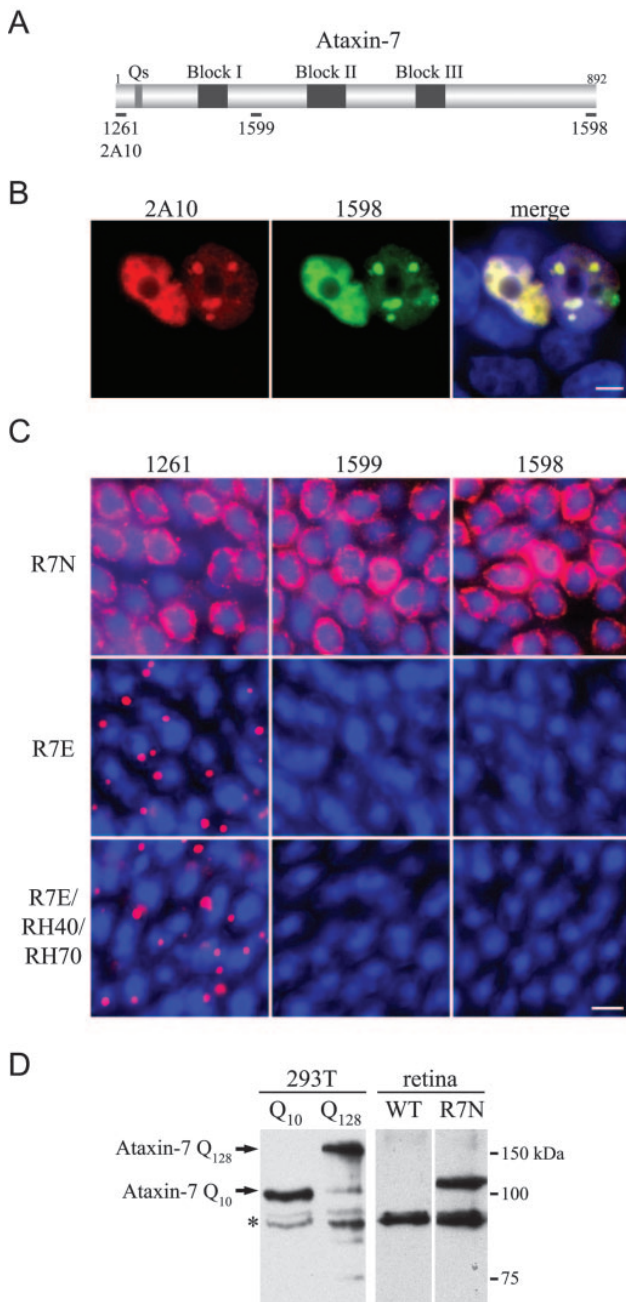
**FIG. 5. Hsp70 and Hdj2 do not suppress aggregates formation in SCA7 mice.** Retinal cryosections (10  $\mu\text{m}$ ) from R7E (A and B) and R7E/RH40/RH70 triple transgenic mice (C and D) at 10 weeks of age were immunostained using anti-ataxin-7 (1261), anti-HDJ2 (A), anti-HSP70 (B), anti-HA (C), and anti-FLAG (D) antibodies. Numerous mutant ataxin-7 nuclear aggregates (green) were found throughout the photoreceptor nuclear layer from both genotypes (compare A and B with C and D). Endogenous Hdj2 was recruited into nuclear inclusions (red, A). Recombinant HDJ2 and HSP70, detected using anti-tag antibodies, did not co-localize with mutant ataxin-7 (red, panels C and D). Nuclei were counterstained with DAPI. The scale bar represents 10  $\mu\text{m}$ .

double transgenic RH70/RH40 or R7N/RH40 mice. In both cases, the expression levels of both recombinant proteins were similar to those observed in single transgenic mice (data not shown). Furthermore, homozygous RH40 and RH70 mice showed 2-fold higher mRNA and protein levels as compared with hemizygotes (data not shown). This suggests that rods from triple transgenic mice can overexpress different transgenes at the expected levels. Finally, Muchowski *et al.* (44) showed that HSP70, in conjunction with HSP40, efficiently inhibited poly(Q) aggregation only when added during the early phase of the process but failed to inhibit polymerization once seeding has occurred. In this study, the respective transgenes were all controlled by the same promoter, leading to an identical onset of expression, in the same cell type and at comparable levels. We previously reported that poly(Q) expansions repress rhodopsin promoter activity, resulting in repression of *SCA7* transgene expression (29). Similarly, *HDJ2* and *HSP70* transgene expression was also affected but to a lesser extent than the *SCA7* transgene. Thus, in R7E/RH40/RH70 triple transgenics, rod photoreceptors expressed recombinant ataxin-7, HSP70, and HDJ2 at comparable levels during the time course of *SCA7* retinopathy.

Polyglutamine aggregates are highly dynamic structures that sequester various proteins either irreversibly or transiently (45–47). In R7E mice, NIs are also dynamic rather than permanent structures, as we previously reported partial clearance of aggregates after loss of mutant *SCA7* transgene expression (29). Expression of both HSP70 and HDJ2 in R7E mice had no effect on the number and size of nuclear inclusions (Fig. 5), indicating that mutant ataxin-7 aggregation was not altered. This observation may not be sufficient to explain why overexpressed chaperones have no beneficial effect on *SCA7* retinopathy. While it is clear that progressive accumulation of poly(Q)-containing proteins is a key pathogenic event, the exact role of

aggregation in neurodegeneration is not fully understood (6, 48). NI formation could be dissociated from neuronal toxicity in one *SCA1* mouse model (49, 50). Furthermore, chaperone modulation of poly(Q) neurodegeneration in flies and of *SCA1* motor impairment in mice was not accompanied by a reduction in aggregate formation (19, 21, 27). Although mature aggregates are the observable outcome of the aggregation process, poly(Q) toxicity might stem from earlier oligomer intermediates, which are difficult to detect and quantify (51–53). Interestingly, some studies have suggested that chaperone overexpression renders poly(Q) aggregates more sensitive to detergent solubilization (17, 25). In R7E mice, HSP70/HSP40 chaperone overexpression had no effect on the SDS-extractability of mutant ataxin-7 aggregates (data not shown). In marked contrast, the same molecular chaperones completely suppressed mutant ataxin-7 aggregation in transfected cells, similarly to numerous previous studies.

Many nonexclusive explanations may account for the differential effects of molecular chaperones on poly(Q) aggregation in cellular and animal models. In transfected cells, NIs become visible by light microscopy as early as 1 day after overexpression of poly(Q) proteins, whereas in most transgenic mouse models, aggregates are detectable a few weeks after the onset of expression (9, 50, 54). In cell-based models, the massive and rapid overload of poly(Q) proteins might hide specific or unspecific molecular events associated with continuous formation and processing of aggregates *in vivo*. We observed that mutant ataxin-7 is proteolytically processed in transgenic mice but not in transfected cells. This discrepancy suggests that the full-length protein is handled differently *in vivo*. However, we could not assess whether mutant ataxin-7 is specifically cleaved or non-specifically trimmed, since the aggregated protein could not be resolubilized using standard procedures. Furthermore, HSP70/HSP40 chaperones were able to handle aggregates



**FIG. 6. Nuclear inclusions contain different forms of ataxin-7 in transfected cells and SCA7 mice.** *A*, schematic representation of human ataxin-7 indicating the positions of the epitopes recognized by the antibodies used in this study. *B*, HEK293T cells were transfected with a full-length mutant SCA7 cDNA construct (pc7E2FL). Immunolocalization of transiently expressed mutant ataxin-7 (128 Gln), using an N-terminal monoclonal antibody (2A10, red) and a C-terminal polyclonal antibody (1598, green) was performed 24 h after transfection. The merged image shows that nuclear inclusions are labeled by both antibodies, indicating that NIs are composed of full-length mutant ataxin-7. *C*, retinal cryosections (10  $\mu$ m) from R7N, R7E, and R7E/RH40/RH70 mice at 10 weeks of age were immunostained using anti-ataxin-7 antibodies (1261, 1599, and 1598) covering the full-length protein. Normal ataxin-7 was stained with all antibodies (upper panels) and displayed a characteristic nuclear staining in rod photoreceptors. In contrast, mutant ataxin-7 accumulated in NIs that were detected using N-terminal antibodies (1261) but not using more C-terminal antibodies (1599 and 1598) in R7E mice (middle panels). Similar results were obtained in R7E/RH40/RH70 mice (lower panels) indicating that C-terminal sequences are lost in both genotypes. Nuclei were counterstained with DAPI. The scale bar represents 5  $\mu$ m. *D*, 1599 polyclonal antibody

formed in cells transfected with constructs encoding full-length or truncated mutant ataxin-7 (Fig. 1C). Thus, proteolytic processing of ataxin-7 in transgenic mice might be one out of a number of nonexclusive molecular mechanisms, which altogether account for the differential sensitivity of poly(Q) aggregates to molecular chaperones in different model systems.

So far, the protective effects of chaperones have been investigated by crossing SCA1, SBMA, and HD transgenic models with mice that overexpress HSP70 ubiquitously. HSP70 overexpression in SBMA mice improved their motor phenotype and survival rate, probably by reducing nuclear translocation of the mutant androgen receptor (55). Similar experiments in SCA1 mice led to a milder amelioration of the motor dysfunction (27), while in R6/2 mice, two independent studies reported no effect on disease progression and survival (17, 28). We show here that even high and specific overexpression of both HSP70 and its co-factor HSP40/HDJ2 did not ameliorate the SCA7 mouse phenotype. Altogether, these observations indicate that chaperone overexpression in mammals might have different effects, depending on the specific features of each poly(Q) protein. In addition, there is a striking discrepancy between these findings and the identification of molecular chaperones as powerful suppressors of poly(Q) toxicity in all fly models generated so far. HSP70 overexpression has been shown to protect flies against other pathological conditions, including  $\alpha$ -synuclein toxicity, fragile X premutation, and normal aging (56–58). Molecular chaperones might have beneficial effects by preventing misfolded proteins from adopting toxic conformations or through a variety of other pathways which could be differentially regulated and implicated between model systems. For instance, chaperones might inhibit apoptosis that is triggered by misfolded proteins. Indeed, HSP70 has been shown to counteract the activity of the stress-induced kinase JNK and to negatively regulate the apoptosome (59). At variance with mammalian models, poly(Q) flies are characterized by massive neuronal loss. This phenotype can be fully rescued by overexpression of chaperones, which could exert their beneficial effect via anti-apoptotic functions. Furthermore, the HSP70/HSP40 system helps ubiquitin ligases to recognize and target misfolded proteins for proteasomal degradation (60). Many mechanisms at the crossroads of protein misfolding, protein degradation, and apoptosis could be influenced by chaperones and could contribute to the differential effect of molecular chaperones in invertebrate and mouse models.

In mammals, the high redundancy of the chaperone and ubiquitin-proteasome systems suggest that there are additional protein quality control components which might not be present or required in invertebrates. This redundancy is probably critical for mammals to protect themselves against specific stress associated with an extended neuronal life span or environmental conditions. If many of these additional factors are altered in poly(Q) mouse models, correction of a single pathway would not be sufficient to suppress the neurological phenotype. In agreement with that hypothesis, Hay *et al.* (17) recently reported that there is a decrease in the levels of several heat shock proteins, which correlates with phenotype progression in R6/2 mice, and that pharmacological induction of chaperones

recognizes normal and mutant ataxin-7 by Western blot. Whole cell extracts of HEK293T cells transfected with full-length normal (pc7NFL) or mutant (pc7E2FL) SCA7 cDNA constructs were analyzed by Western blot using 1599 antibody. Bands corresponding to normal (10 Gln) and mutant (128 Gln) ataxin-7 were specifically revealed by 1599 at the expected sizes (110 and 160 kDa, respectively). Western blot of retinal homogenates from WT or R7N mice confirmed that 1599 recognizes normal recombinant ataxin-7 at the expected size (110 kDa). A nonspecific cross-reacting band of lower molecular weight (asterisk) was detected in both WT and R7N retina.



significantly delayed aggregates formation in an organotypic slice culture assay. Thus, induction of a global heat shock response, not exclusively HSP40 and HSP70, might be a more useful therapeutic approach that is worthy of further investigation in mouse models of poly(Q) diseases. Our results also indicate that a better understanding of the molecular mechanisms underlying the discrepancies between cellular, invertebrate, and mouse models would be required before further validation of chaperone induction as a therapeutic strategy.

**Acknowledgments**—We thank Bohdan Wasyluk for critical reading of the manuscript and K. Merienne for helpful discussion. We thank J. Bennett for gift of the rhodopsin promoter, R. Morimoto for the HSP70 cDNA, and T. Mohanakumar for the HDJ2 cDNA and D. Metzger for the pGSp2 plasmid. We are indebted to C. Weber, S. Rousseau, J. Herguex, and A. Gansmüller for technical assistance and to G. Yvert and G. Duval for 1599 antibody production and to the staff at the Mouse Clinical Institute for microinjections and mouse care.

## REFERENCES

- Michalik, A., Martin, J. J., and Van Broeckhoven, C. (2004) *Eur. J. Hum. Genet.* **12**, 2–15
- David, G., Abbas, N., Stevanin, G., Durr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., Drabkin, H., Gemmill, R., Giunti, P., Benomar, A., Wood, N., Ruberg, M., Agid, Y., Mandel, J. L., and Brice, A. (1997) *Nat. Genet.* **17**, 65–70
- Helminger, D., Hardy, S., Sasorith, S., Klein, F., Robert, F., Weber, C., Miguët, L., Potier, N., Van-Dorselaer, A., Wurtz, J. M., Mandel, J. L., Tora, L., and Devys, D. (2004) *Hum. Mol. Genet.* **13**, 1257–1265
- Zoghbi, H. Y., and Orr, H. T. (2000) *Annu. Rev. Neurosci.* **23**, 217–247
- Ross, C. A. (2002) *Neuron* **35**, 819–822
- Bates, G. (2003) *Lancet* **361**, 1642–1644
- Cummings, C. J., Mancini, M. A., Antalfy, B., DeFranco, D. B., Orr, H. T., and Zoghbi, H. Y. (1998) *Nat. Genet.* **19**, 148–154
- Opal, P., and Zoghbi, H. Y. (2002) *Trends Mol. Med.* **8**, 232–236
- Yvert, G., Lindenberg, K. S., Picaud, S., Landwehrmeyer, G. B., Sahel, J. A., and Mandel, J. L. (2000) *Hum. Mol. Genet.* **9**, 2491–2506
- Yvert, G., Lindenberg, K. S., Devys, D., Helminger, D., Landwehrmeyer, G. B., and Mandel, J. L. (2001) *Hum. Mol. Genet.* **10**, 1679–1692
- La Spada, A. R., Fu, Y. H., Sopher, B. L., Libby, R. T., Wang, X., Li, L. Y., Einum, D. D., Huang, J., Possin, D. E., Smith, A. C., Martinez, R. A., Koszidin, K. L., Treuting, P. M., Ware, C. B., Hurley, J. B., Ptacek, L. J., and Chen, S. (2001) *Neuron* **31**, 913–927
- Yoo, S. Y., Pennesi, M. E., Weber, E. J., Xu, B., Atkinson, R., Chen, S., Armstrong, D. L., Wu, S. M., Sweatt, J. D., and Zoghbi, H. Y. (2003) *Neuron* **37**, 383–401
- Mauger, C., Del-Favero, J., Ceuterick, C., Lubke, U., van Broeckhoven, C., and Martin, J. (1999) *Brain Res. Mol. Brain Res.* **74**, 35–43
- Zander, C., Takahashi, J., El Hachimi, K. H., Fujigasaki, H., Albanese, V., Lebre, A. S., Stevanin, G., Duyckaerts, C., and Brice, A. (2001) *Hum. Mol. Genet.* **10**, 2569–2579
- Takahashi, J., Fujigasaki, H., Zander, C., El Hachimi, K. H., Stevanin, G., Durr, A., Lebre, A. S., Yvert, G., Trottier, Y., The, H., Hauw, J. J., Duyckaerts, C., and Brice, A. (2002) *Brain* **125**, 1534–1543
- Helminger, D., Yvert, G., Picaud, S., Merienne, K., Sahel, J., Mandel, J. L., and Devys, D. (2002) *Hum. Mol. Genet.* **11**, 3351–3359
- Hay, D. G., Sathasivam, K., Tobaben, S., Stahl, B., Marber, M., Mestril, R., Mahal, A., Smith, D. L., Woodman, B., and Bates, G. P. (2004) *Hum. Mol. Genet.* **13**, 1389–1405
- Sakahira, H., Breuer, P., Hayer-Hartl, M. K., and Hartl, F. U. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, Suppl. 4, 16412–16418
- Kazemi-Esfarjani, P., and Benzer, S. (2000) *Science* **287**, 1837–1840
- Fernandez-Funez, P., Nino-Rosales, M. L., de Gouyon, B., She, W. C., Luchak, J. M., Martinez, P., Turiegano, E., Benito, J., Capovilla, M., Skinner, P. J., McCall, A., Canal, I., Orr, H. T., Zoghbi, H. Y., and Botas, J. (2000) *Nature* **408**, 101–106
- Warrick, J. M., Chan, H. Y. E., Gray-Board, G. L., Chai, Y., Paulson, H. L., and Bonini, N. M. (1999) *Nat. Genet.* **23**, 425–428
- Chan, H. Y., Warrick, J. M., Andriola, I., Merry, D., and Bonini, N. M. (2002) *Hum. Mol. Genet.* **11**, 2895–2904
- Takeyama, K., Ito, S., Yamamoto, A., Tanimoto, H., Furutani, T., Kanuka, H., Miura, M., Tabata, T., and Kato, S. (2002) *Neuron* **35**, 855–864
- Ghosh, S., and Feany, M. B. (2004) *Hum. Mol. Genet.* **13**, 2011–2018
- Chan, H. Y., Warrick, J. M., Gray-Board, G. L., Paulson, H. L., and Bonini, N. M. (2000) *Hum. Mol. Genet.* **9**, 2811–2820
- Hartl, F. U., and Hayer-Hartl, M. (2002) *Science* **295**, 1852–1858
- Cummings, C. J., Sun, Y., Opal, P., Antalfy, B., Mestril, R., Orr, H. T., Dillmann, W. H., and Zoghbi, H. Y. (2001) *Hum. Mol. Genet.* **10**, 1511–1518
- Hansson, O., Nylandsted, J., Castilho, R. F., Leist, M., Jaattela, M., and Brundin, P. (2003) *Brain Res.* **970**, 47–57
- Helminger, D., Abou-Sleymane, G., Yvert, G., Rousseau, S., Weber, C., Trottier, Y., Mandel, J. L., and Devys, D. (2004) *J. Neurosci.* **24**, 1881–1887
- Demand, J., Luders, J., and Hohfeld, J. (1998) *Mol. Cell. Biol.* **18**, 2023–2028
- Freeman, B. C., Myers, M. P., Schumacher, R., and Morimoto, R. I. (1995) *EMBO J.* **14**, 2281–2292
- Garden, G. A., Libby, R. T., Fu, Y. H., Kinoshita, Y., Huang, J., Possin, D. E., Smith, A. C., Martinez, R. A., Fine, G. C., Grote, S. K., Ware, C. B., Einum, D. D., Morrison, R. S., Ptacek, L. J., Sopher, B. L., and La Spada, A. R. (2002) *J. Neurosci.* **22**, 4897–4905
- Chai, Y., Koppenhafer, S. L., Bonini, N. M., and Paulson, H. L. (1999) *J. Neurosci.* **19**, 10338–10347
- Jana, N. R., Tanaka, M., Wang, G., and Nukina, N. (2000) *Hum. Mol. Genet.* **9**, 2009–2018
- Mitsui, K., Nakayama, H., Akagi, T., Nekooki, M., Ohtawa, K., Takio, K., Hashikawa, T., and Nukina, N. (2002) *J. Neurosci.* **22**, 9267–9277
- Ishihara, K., Yamagishi, N., Saito, Y., Adachi, H., Kobayashi, Y., Sobue, G., Ohtsuka, K., and Hatayama, T. (2003) *J. Biol. Chem.* **278**, 25143–25150
- Stenoien, D. L., Cummings, C. J., Adams, H. P., Mancini, M. G., Patel, K., DeMartino, G. N., Marcellini, M., Weigel, N. L., and Mancini, M. A. (1999) *Hum. Mol. Genet.* **8**, 731–741
- Kobayashi, Y., Kume, A., Li, M., Doyu, M., Hata, M., Ohtsuka, K., and Sobue, G. (2000) *J. Biol. Chem.* **275**, 8772–8778
- Sittler, A., Lurz, R., Lueder, G., Priller, J., Lehrach, H., Hayer-Hartl, M. K., Hartl, F. U., and Wanker, E. E. (2001) *Hum. Mol. Genet.* **10**, 1307–1315
- Zhou, H., Li, S. H., and Li, X. J. (2001) *J. Biol. Chem.* **276**, 48417–48424
- Bailey, C. K., Andriola, I. F., Kampinga, H. H., and Merry, D. E. (2002) *Hum. Mol. Genet.* **11**, 515–523
- Chuang, J. Z., Zhou, H., Zhu, M., Li, S. H., Li, X. J., and Sung, C. H. (2002) *J. Biol. Chem.* **277**, 19831–19838
- Carmichael, J., Chatellier, J., Woolfson, A., Milstein, C., Fersht, A. R., and Rubinsztein, D. C. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9701–9705
- Muchowski, P. J., Schaffar, G., Sittler, A., Wanker, E. E., Hayer-Hartl, M. K., and Hartl, F. U. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7841–7846
- Chai, Y., Shao, J., Miller, V. M., Williams, A., and Paulson, H. L. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9310–9315
- Kim, S., Nollen, E. A., Kitagawa, K., Bindokas, V. P., and Morimoto, R. I. (2002) *Nat. Cell Biol.* **4**, 826–831
- Stenoien, D. L., Mielke, M., and Mancini, M. A. (2002) *Nat. Cell Biol.* **4**, 806–810
- Michalik, A., and Van Broeckhoven, C. (2003) *Hum. Mol. Genet.* **12**, R173–R186
- Klement, I. A., Skinner, P. J., Kaytor, M. D., Yi, H., Hersch, S. M., Clark, H. B., Zoghbi, H. Y., and Orr, H. T. (1998) *Cell* **95**, 41–53
- Cummings, C. J., Reinstein, E., Sun, Y., Antalfy, B., Jiang, Y., Ciechanover, A., Orr, H. T., Beaudet, A. L., and Zoghbi, H. Y. (1999) *Neuron* **24**, 879–892
- Yang, W., Dunlap, J. R., Andrews, R. B., and Wetzel, R. (2002) *Hum. Mol. Genet.* **11**, 2905–2917
- Chen, S., Berthelot, V., Yang, W., and Wetzel, R. (2001) *J. Mol. Biol.* **311**, 173–182
- Sanchez, I., Mahlke, C., and Yuan, J. (2003) *Nature* **421**, 373–379
- Davies, S. W., Turmaine, M., Cozens, B. A., Dfiglia, M., Sharp, A. H., Ross, C. A., Scherzinger, E., Wanker, E. E., Mangiarini, L., and Bates, G. P. (1997) *Cell* **90**, 537–548
- Adachi, H., Katsuno, M., Minamiyama, M., Sang, C., Pagoulatos, G., Angelidis, C., Kusakabe, M., Yoshiki, A., Kobayashi, Y., Doyu, M., and Sobue, G. (2003) *J. Neurosci.* **23**, 2203–2211
- Tatar, M., Khazaali, A. A., and Curtsinger, J. W. (1997) *Nature* **390**, 30
- Auluck, P. K., Chan, H. Y., Trojanowski, J. Q., Lee, V. M., and Bonini, N. M. (2002) *Science* **295**, 865–868
- Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K., and Warren, S. T. (2003) *Neuron* **39**, 739–747
- Takayama, S., Reed, J. C., and Homma, S. (2003) *Oncogene* **22**, 9041–9047
- Berke, S. J., and Paulson, H. L. (2003) *Curr. Opin. Genet. Dev.* **13**, 253–261

*« C'est dangereux de vouloir tout savoir, il vaut bien mieux,  
s'entourer de mystère, on est plus sûr de s'endormir le soir... »*

Ray Ventura

## Analyses structure fonction du module de déubiquitination du complexe SAGA

### Résumé

Pour faciliter l'initiation de la transcription par l'ARN Polymérase II, le complexe co-activateur de la transcription SAGA possède une activité d'acétylation des histones H3 et une activité de déubiquitination des histones H2B, catalysée chez l'homme par l'enzyme USP22.

Mon travail de thèse a porté sur l'étude de la régulation de cette activité de déubiquitination. Au sein de SAGA, USP22 interagit fortement avec trois protéines pour former un module structural appelé module de déubiquitination (DUBm). Nous avons montré que la formation d'un tel module était requise pour activer USP22.

D'autre part, deux sous-unités du DUBm humain, ATXN7 et ATXN7L3, contiennent un domaine SCA7. Nos résultats montrent que le repliement structural adopté par ces deux doigts de zinc n'avait pas encore été décrit. Nous avons démontré que le domaine SCA7 de ATXN7 peut interagir avec un nucléosome *in vitro* et que cette interaction participe à la régulation fine de l'activité de déubiquitination de SAGA.

Chromatine - Transcription - SAGA - Ubiquitine - Déubiquitination - Doigt de zinc

### Summary

To facilitate transcription by RNA Polymerase II, the transcriptional co-activator SAGA harbours two enzymatic activities. In human cells, these two enzymes are called GCN5 and USP22 and they can respectively acetylate histones H3 and deubiquitinate histones H2B.

During my PhD thesis, I have worked on the regulation of SAGA deubiquitination activity. In the SAGA complex, USP22 interacts strongly with three proteins to form a structural module, named deubiquitination module (DUBm). We have shown that the formation of a stable DUBm increases USP22 catalytic activity.

Secondly, two subunits of the human DUBm, ATXN7 and ATXN7L3, contain a SCA7 domain. Our results show that the new structural fold adopted by these two domains is specific to these zinc-fingers. We have shown that the ATXN7 SCA7 domain can interact *in vitro* with a nucleosome which fine tunes SAGA deubiquitination activity, possibly by an optimal positioning of the module to its substrate.

Chromatin - Transcription - SAGA - Ubiquitin - Deubiquitination - Zinc-finger