

Université de Montréal

**Caractérisation de la fonction des complexes histone
déacétylases Rpd3S et Set3C.**

par

Simon Drouin

Programmes de Biologie moléculaire

Faculté de Médecine

Thèse présentée à la Faculté de Médecine
en vue de l'obtention du grade de Docteur
en Biologie moléculaire

Mai 2011

© Simon Drouin, 2011

Université de Montréal
Faculté de Médecine

Cette thèse intitulée :

Caractérisation de la fonction des complexes histone déacétylases Rpd3S et Set3C.

Présentée par :
Simon Drouin

a été évaluée par un jury composé des personnes suivantes :

Dre Marie Kmita, président-rapporteur
Dr François Robert, directeur de recherche
Dr Alain Verreault, membre du jury
Dr Jacques Côté, examinateur externe
Dr Dindial Ramotar, représentant du doyen de la FES

Résumé

La chromatine est essentielle au maintien de l'intégrité du génome, mais, ironiquement, constitue l'obstacle principal à la transcription des gènes. Plusieurs mécanismes ont été développés par la cellule pour pallier ce problème, dont l'acétylation des histones composant les nucléosomes. Cette acétylation, catalysée par des histones acétyl transférases (HATs), permet de réduire la force de l'interaction entre les nucléosomes et l'ADN, ce qui permet à la machinerie transcriptionnelle de faire son travail. Toutefois, on ne peut laisser la chromatine dans cet état permissif sans conséquence néfaste. Les histone déacétylases (HDACs) catalysent le clivage du groupement acétyle pour permettre à la chromatine de retrouver une conformation compacte.

Cette thèse se penche sur la caractérisation de la fonction et du mécanisme de recrutement des complexes HDACs Rpd3S et Set3C. Le complexe Rpd3S est recruté aux régions transcrites par une interaction avec le domaine C-terminal hyperphosphorylé de Rpb1, une sous-unité de l'ARN polymérase II. Toutefois, le facteur d'élongation DSIF joue un rôle dans la régulation de cette association en limitant le recrutement de Rpd3S aux régions transcrites. L'activité HDAC de Rpd3S, quant à elle, dépend de la méthylation du résidu H3K36 par l'histone méthyltransférase Set2.

La fonction du complexe Set3C n'est pas clairement définie. Ce complexe est recruté à la plupart de ses cibles par l'interaction entre le domaine PHD de Set3 et le résidu H3K4 di- ou triméthylé. Un mécanisme indépendant de cette méthylation, possiblement le même que pour Rpd3S, régit toutefois l'association de Set3C aux régions codantes des gènes les plus transcrits.

La majorité de ces résultats ont été obtenus par la technique d'immunoprécipitation de la chromatine couplée aux biopuces (ChIP-chip). Le protocole technique et le design expérimental de ce type d'expérience fera aussi l'objet d'une discussion approfondie.

Mots-clés : Immunoprécipitation de la chromatine; ChIP-chip; biopuces; histone acétyltransférase (HAT); histone déacétylase (HDAC); histone méthyltransférase (HMT); méthylation de H3K36; méthylation de H3K4; transcription cryptique; transcription; recrutement de Rpd3S; Set2; DSIF; Spt4; Spt5; Kin28; Ctk1; Bur1; Rpd3; Rco1; Eaf3; recrutement de Set3C; Set1; Hst1; Hos2.

Abstract

Chromatin is essential for the maintenance of genomic integrity but, ironically, is also the main barrier to gene transcription. Many mechanisms, such as histone acetylation, have evolved to overcome this problem. Histone acetylation, catalyzed by histone acetyltransferases (HATs), weakens the internucleosomal and nucleosome-DNA interactions, thus permitting the transcriptional machinery access to its template. However, this permissive chromatin state also allows for opportunistic DNA binding events. Histone deacetylases (HDACs) help restore a compact chromatin structure by catalyzing the removal of acetyl moieties from histones.

This thesis focuses on the characterization of the function and of the recruitment mechanism of HDAC complexes Rpd3S and Set3C. The Rpd3S complex is recruited to actively transcribed coding regions through interactions with the hyperphosphorylated C-terminal domain of Rpb1, a subunit of RNA polymerase II, with the DSIF elongation factor playing a role in limiting this recruitment. However, the HDAC activity of Rpd3S depends on H3K36 methylation, which is catalyzed by the Set2 histone methyltransferase.

The Set3C complex' function is still not clearly defined. It is recruited to most of its targets through the interaction between the Set3 PHD domain and di- or trimethylated H3K4. However, Set3C recruitment to genes displaying high RNA polymerase II occupancy is independent of H3K4 methylation. The mechanism by which Set3C is recruited to this gene subset is under investigation.

These results have mostly been obtained through chromatin immunoprecipitation coupled to tiling microarrays (ChIP-chip). The protocol and experimental design challenges inherent to this technique will also be discussed in depth.

Keywords : Chromatin immunoprecipitation; ChIP-chip; microarrays; histone acetyltransferase (HAT); histone deacetylase (HDAC); histone methyltransferase (HMT);

H3K36 methylation; H3K4 methylation; cryptic transcription; transcription; Rpd3S recruitment; Set2; DSIF; Spt4; Spt5; Kin28; Ctk1; Bur1; Rpd3; Rco1; Eaf3; Set3C recruitment; Set1; Hst1; Hos2.

Table des matières

1	Introduction.....	12
1.1	Les histone acétyltransférases (HATs).....	16
1.2	Les histone déacétylases	25
1.2.1	Les HDACs de classe I	25
1.2.2	Les HDACs de classe II	37
1.2.3	Les HDACs de classe III.....	38
1.3	Les HDACs chez l'humain	44
1.3.1	La classe I.....	45
1.3.2	La classe II	49
1.3.3	La classe III.....	52
1.3.4	La classe IV.....	56
1.4	Les histone lysine méthyltransférases (KMTs).....	58
1.5	Les histone déméthylases (HDMs)	63
1.6	Hypothèses et objectifs	65
2	Le CHIP-Chip : technique et contrôles	68
2.1	Manuscrit : « Genome-wide Location Analysis of Chromatin-associated Proteins by CHIP on CHIP: Controls Matter ».....	70
2.1.1	Abstract	71
2.1.2	Key words	71
2.1.3	Introduction.....	71
2.1.4	Experimental design.....	72
2.1.5	Growing and cross-linking cells.....	73
2.1.6	Breaking cells.....	75
2.1.7	Preparation of soluble chromatin fragments	76
2.1.8	Immunoprecipitation.....	77
2.1.9	DNA Purification	79

2.1.10	Amplification and labeling.....	80
2.1.11	Comparison between experimental designs	85
2.1.12	Conclusions.....	86
3	Caractérisation de l'activité et du mode de recrutement du complexe HDAC Rpd3S aux régions codantes activement transcrites	89
3.1	Manuscrit : « DSIF and RNA polymerase II CTD Phosphorylation Coordinate the Recruitment of Rpd3S to Actively Transcribed Genes ».....	91
3.1.1	Abstract	92
3.1.2	Keywords	92
3.1.3	Introduction.....	92
3.1.4	Results.....	95
3.1.5	Discussion.....	104
3.1.6	Materials and Methods.....	110
3.1.7	Acknowledgments.....	112
4	Détermination génomique des caractéristiques de liaison du complexe Set3C et de sa sous-unité Hst1	127
4.1	Manuscrit : « Histone H3 Lysine 4 Methylation and RNAPII CTD Phosphorylation Recruit Distinct Set3 Complexes to Different Sets of Genes. »	130
4.1.1	Abstract	131
4.1.2	Introduction.....	131
4.1.3	Results.....	133
4.1.4	Discussion.....	137
4.1.5	Materials and Methods.....	143
5	Conclusion	152
6	Perspectives.....	159
7	Bibliographie.....	162

Liste des tableaux

Tableau 1 : Les HDACs humaines et leurs homologues chez la levure. 57

Liste des figures

Figure 1-1 : Structure du nucléosome à 2.4Å, d'après Luger et coll., 1997.	14
Figure 1-2 : Résumé des modifications covalentes post-traductionnelles des histones H2B, H3 et H4.	15
Figure 1-3 : Composition et localisation des complexes Rpd3L et Rpd3S.	33
Figure 1-4 : Le modèle courant de la fonction de Rpd3S.	34
Figure 2-1 : Genome-wide Location Analysis of the transcription factor Sum1 using input DNA (A) or a control IP (B) as the hybridization control.	87
Figure 2-2 : Control Genome-wide Location Analysis experiments performed on yeast strain containing no tagged protein.	88
Figure 3-1 : Rpd3S and Rpd3L target overlapping sets of genes.	113
Figure 3-2 : The recruitment of Rpd3S to transcribed genes does not require Set2-dependent H3K36 methylation in vivo.	115
Figure 3-3 : Set2-dependent H3K36 methylation is required for the function of Rpd3S.	117
Figure 3-4 : Spt4 negatively regulates the recruitment of Rpd3S	118
Figure 3-5 : Phosphorylation of the RNAPII and Spt5 CTD are important for the Rpd3S recruitment on genes.	120
Figure 3-6 : A model summarizing how Rpd3S is recruited to active genes.	121
Figure 3-7 : A complement to Figure 3-1. Average signal of RNAPII, Rxt2 and Sds3 over the genes of the 3 clusters.	122
Figure 3-8 : A complement to Figure 3-2. The recruitment of Rpd3S to transcribed genes does not require Set2-dependent H3K36 methylation in vivo.	123
Figure 3-9 : A complement to Figure 3-3. Deletion of <i>SET2</i> or disruption of Rpd3S causes a general decrease in histone acetylation at promoters.	124

Figure 3-10 : A complement to Figure 3-4. Spt4 negatively regulates the recruitment of Rpd3S.	125
Figure 3-11 : Rpd3S ORF occupancy level is gene length-dependant in a <i>spt4Δ / ctk1Δ</i> mutant.	126
Figure 4-1 : Hst1, Hos2 and Set3 bind functionally distinct regions.	145
Figure 4-2 : Loss of H3K4 methylation affects Set3 but not Hst1 binding.	147
Figure 4-3 Set3 and RNAPII occupancies correlate better in H3K4 methylation-deficient backgrounds than in WT.	148
Figure 4-4 : A complement to Figure 4-1. Set3, but not Hst1, binding is affected by loss of H3K4 di- and trimethylation.	149
Figure 4-5 : H3K4 mono-, di- and trimethylation patterns in WT cells.	150
Figure 4-6 : Rpd3L and Set3C occupancies overlap at a set of ORFs.	151

À Roland, le Jeune et le Vieux.

Remerciements

J'aimerais tout d'abord remercier François pour la chance qu'il m'a donnée de travailler avec lui ainsi que pour son support moral et financier, mais surtout pour la grande confiance qu'il a placée en moi au cours de ces années. Il a su quand encourager mais aussi quand freiner mon enthousiasme et me réorienter vers l'essentiel. J'aimerais aussi remercier mes collègues, particulièrement Alain et Louise, pour toutes les discussions et l'appui moral et technique qu'ils m'ont donnés.

Évidemment, un merci qui ne sera jamais assez gros pour mes parents pour qui il ne m'est resté qu'un an ou deux de travail avant de finir mon doctorat, et ce pendant quatre ans. Ils n'ont jamais (ou rarement) répondu « parce que » à mes questions sur le monde qui m'entourait et ont ainsi planté une graine qui a germé au fil des ans pour culminer en cette thèse.

À ma (bien) meilleure moitié, merci d'avoir dit « Oui, je le veux », pour Ti-Pout, pour toutes les belles années passées ensemble et celles à venir. Merci pour tous les beaux moments, mais aussi pour ceux qui ont été plus difficiles. Je ne me serais pas rendu jusqu'ici sans toi.

1 Introduction

Il y a de cela bientôt 40 ans, Roger Kornberg postulait que la structure de la chromatine était basée sur une répétition d'environ 200 paires de bases d'ADN enroulées autour de nucléosomes composés de huit histones [1]. Les histones sont les protéines les mieux conservées chez les eucaryotes et ont d'abord été confondues avec le matériel génétique lui-même lors de leur découverte en 1894 dû à leur poids total équivalant à celui de l'ADN. Ce n'est qu'en 1980 qu'une première structure du nucléosome fut présentée et que sa configuration interne fut déduite [2]. Des structures à plus haute résolution obtenues en 1984 (7 Å) [3], 1991 (3.6 Å) [4] et 1997 (2.8 Å) [5] ont permis de confirmer que 147 paires de bases d'ADN étaient en effet enroulées autour du nucléosome, lui-même composé d'un tétramère H3-H4 flanqué de deux dimères H2A-H2B (**Figure 1-1**).

Les nucléosomes sont présents sur tout le génome et représentent l'obstacle principal aux mécanismes cellulaires impliquant l'ADN, tels que sa réplication ou la transcription. La cellule eucaryote a donc dû trouver des solutions pour contourner ce problème : une multitude de complexes protéiques dont la fonction est d'ouvrir ou de refermer la matrice chromatinienne pour réguler l'accès à l'ADN aux différentes protéines effectrices.

Ces complexes utilisent divers mécanismes pour parvenir à leurs fins. Certains, tels que les complexes ISWI, utilisent l'énergie de l'hydrolyse de l'ATP pour déplacer les nucléosomes. Les chaperonnes d'histones, quant à elles, sont capables d'enlever et de remettre des dimères ou tétramères H2A/H2B ou H3/H4, ou même d'échanger ceux-ci pour une histone variant, tel que H2A.Z. Ces différents mécanismes travaillent toutefois main dans la main avec une multitude de complexes ayant pour fonction de modifier de manière covalente les histones nucléosomales.

Au moins une centaine de modifications covalentes des différentes histones ont été recensées au fil des ans. On y retrouve, entre autres, l'acétylation, la méthylation, la phosphorylation, l'ubiquitination, la sumoylation et l'isomérisation de résidus prolines. Toutes ces modifications ont des fonctions distinctes, quelquefois redondantes, d'autres

fois opposées. Un exemple typique de redondance est l'acétylation des queues N-terminales des histones H3 et H4, exception faite de la lysine 16 de l'histone H4 : bien que l'acétylation d'un seul de ces résidus soit suffisante pour activer la transcription [6], on observe l'acétylation de plusieurs de ces résidus par différents complexes *in vivo*. Par contre, cette même modification sur la lysine 16 de l'histone H4 est nécessaire pour le bon fonctionnement des mécanismes régulant l'inhibition épigénétique de la transcription (« silencing ») aux télomères et loci HM et d'ADN ribosomal.

Cette thèse traite principalement des complexes histone déacétylases (HDACs) Rpd3S et Set3C. Je passerai donc en revue les différentes protéines ayant une fonction reliée à ces complexes, soit les histone acétyltransférases, déacétylases, méthyltransférases et déméthylases. La **Figure 1-2** illustre les résidus modifiés pertinents à cette discussion.

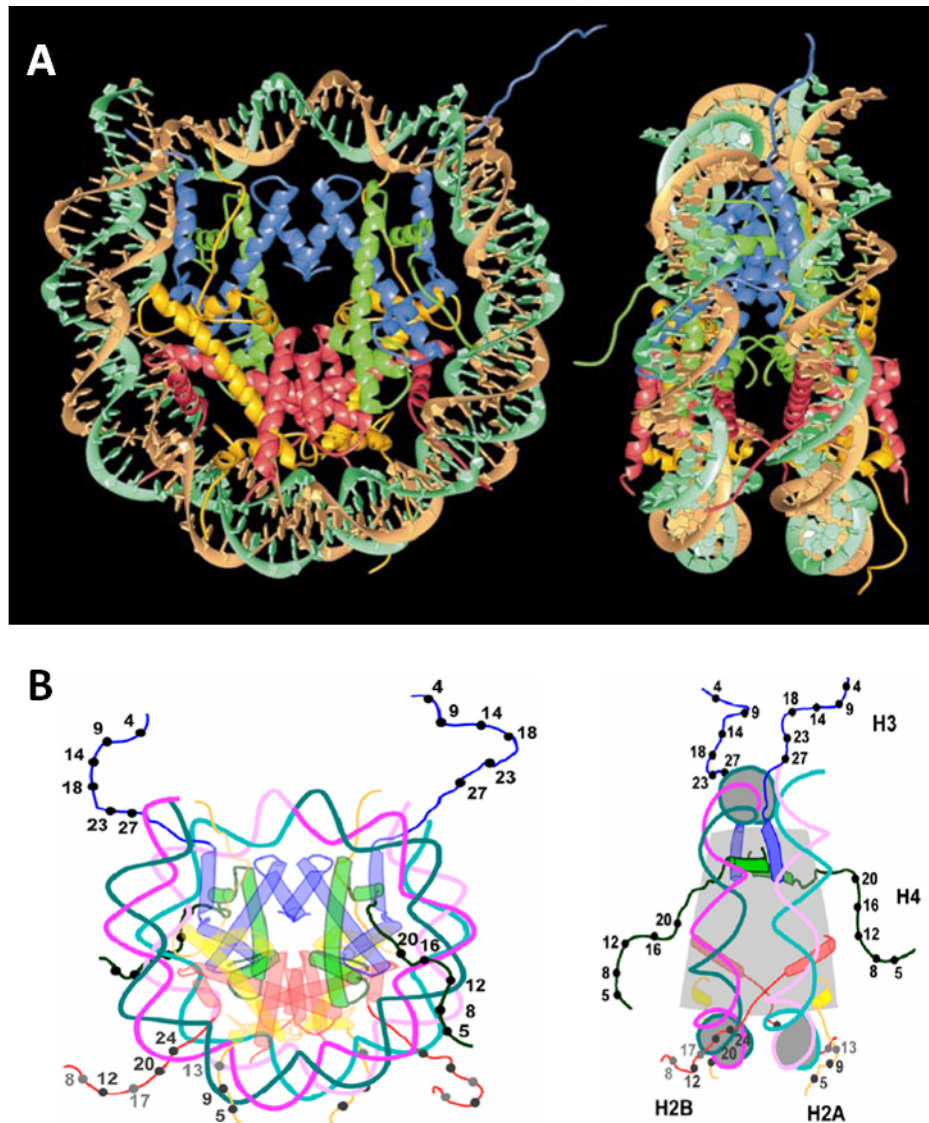


Figure 1-1 : Structure du nucléosome à 2.4Å, d'après Luger et coll., 1997 [5].

(A) Vues dans l'axe et de côté d'un nucléosome et de l'ADN y étant associé. L'histone H2A est en jaune, H2B en rouge, H3 en bleu et H4 en vert.

(B) Schématisation de la structure 3D présentée en (A). Les résidus lysine situés sur les queues N-terminales sont identifiés. Le code de couleur est le même qu'en (A).

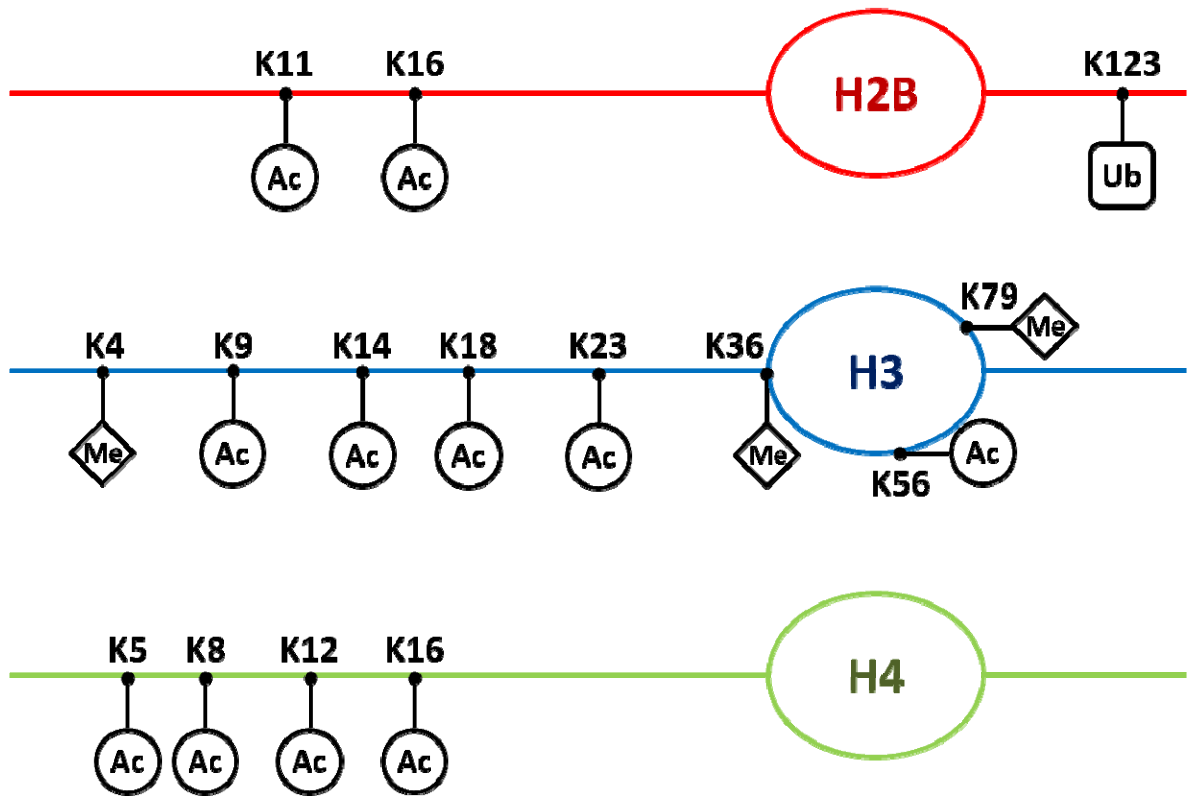


Figure 1-2 : Résumé des modifications covalentes post-traductionnelles des histones H2B, H3 et H4.

On retrouve ici une sélection des marques d'acétylation, de méthylation et de monoubiquitination présentes sur les différentes histones. Cette sélection est restreinte aux modifications pertinentes à cette thèse.

1.1 Les histone acétyltransférases (HATs)

La famille des histone acétyltransférases comprends huit membres chez la levure. Ces enzymes catalysent le transfert d'un groupement acétyle de l'acétyle CoA aux résidus lysine des histones pour former un résidu ϵ -N-acétyle lysine. L'acétylation annule la charge positive associée aux résidus lysine, ce qui affaiblit les interactions électrostatiques internucléosomales et nucléosomes-ADN. Cet affaiblissement donne à la chromatine une structure plus détendue propice aux interactions entre l'ADN et les différents facteurs impliqués entre autres dans le contrôle de la transcription; l'impact de l'activité des HATs sur la transcription est donc généralement positif. Il est important de noter que l'acétylation est l'une des marques covalentes les plus dynamiques : sa demi-vie, déterminée par la compétition entre les acétyltransférases et déacétylases, est de l'ordre de quelques minutes [7, 8]. Ces changements rapides de niveaux d'acétylation permettent donc une régulation précise des mécanismes qui en dépendent. Je passerai maintenant en revue les huit différentes HATs présentes chez la levure.

1.1.1.1 *ELP3*

La HAT Elp3 fait partie du complexe Elongator et est capable d'acétyler les quatre histones principales *in vitro*, mais seulement H3 et H4 *in vivo* [9, 10]. Le complexe Elongator est associé avec la forme active hyperphosphorylée de l'ARN polymérase II et aide à l'élongation. Des délétions de chacune des sous-unités d'Elongator sont viables, sauf *elp5 Δ* qui est létale. Ces délétions confèrent toutes les mêmes phénotypes à la cellule, nommément une croissance lente, un défaut d'activation de la transcription, une sensibilité aux hautes températures ainsi qu'une résistance à la zymocine [9, 11-13].

L'importance de la fonction HAT d'Elp3 a été déduite par ses interactions génétiques. En effet, une délétion *elp3 Δ* est synthétique-malade ou létale dans une souche

où la queue N-terminale de l'histone H3 ou H4 est absente et est synthétique-malade lorsque combinée avec une délétion de la HAT *GCN5*. Ces phénotypes sont supprimés lorsque les HDACs *HDA1* et *HOS2* sont toutes les deux mutées [14], ce qui confirme que le rôle d'Elp3 *in vivo* est opposé à celui de ces HDACs et complémentaire à celui de Gcn5. L'hypoacétylation des nucléosomes observée chez le double mutant *elp3Δ gcn5Δ* a aussi pour effet de permettre un étalement excessif de Sir3, ce qui cause une répression aberrante des gènes situés dans les régions sous-téломériques. Cet effet est supprimé par des délétions des protéines Sir [15].

1.1.1.2 *TAF1*

Taf1 (TATA binding protein-Associated Factor) fait partie du complexe TFIID, qui comprend aussi Spt15/TBP (*TATA-Binding Protein*) ainsi que 12 autres protéines. Ce complexe est impliqué dans l'activation de la transcription par l'ARN polymérase II et peut reconnaître et se lier au promoteur principal (« core promoter ») [16]. Fait intéressant, plusieurs sous-unités de TFIID, à l'exception notable de Taf1, se retrouvent aussi dans le complexe HAT SAGA, ce qui permet à ce dernier de lier les éléments TATA aux promoteurs [17]. Les fonctions HAT de SAGA et de TFIID sont toutefois redondantes selon une étude détaillant les effets sur la transcription de délétions de TFIID et de SAGA [18]. En 2004, le laboratoire de Pugh a publié deux études démontrant que seulement 20 % des gènes de la levure comprennent un site TATA dans leur promoteur, que ceux-ci sont majoritairement associés à la réponse aux stress environnementaux et qu'ils utilisent préférentiellement SAGA pour leur activation. Les 80 % restants font surtout partie du groupe de gènes dit « de ménage » (« housekeeping genes ») dont l'expression est généralement réprimée lors de la plupart de ces stress. L'activation de ces gènes dépend surtout de TFIID [19, 20].

1.1.1.3 GCN5

Le gène *GCN5* (General Control Nonderepressible) a été décrit en 1992 comme un élément de contrôle positif pour l'activation de la transcription des gènes de synthèse d'acides aminés par le coactivateur *GCN4* [21]. On a ensuite découvert que Gcn5, Ada2 et Ada3, deux autres protéines activatrices de la transcription, interagissent physiquement pour former un complexe stable [22-24]. L'activité HAT de Gcn5 a été observée sur des histones libres *in vitro* en 1996 [25, 26]. Par contre, deux études ont démontré que Gcn5 ne peut pas acétyler un substrat nucléosomal seul et doit être associé en complexe pour que cette activité puisse avoir lieu [27, 28]. Ces deux études ont aussi identifié deux complexes contenant Gcn5 possédant une activité HAT dirigée vers l'histone H3 en complexe nucléosomal : SAGA, qui contient aussi des membres du complexe TFIID tel que mentionné ci-dessus, et ADA. Ces deux complexes ciblent l'activité HAT de Gcn5 à des cibles différentes : SAGA cible les quatre lysines présentes sur la queue N-terminale de l'histone H3 (K9, 14, 18 et 23), tandis qu'ADA cible plutôt les lysines 14 et 18 de cette histone [29].

Bien qu'on sache qu'une mutation *gcn5Δ* perturbait la structure de la chromatine au promoteur de *PHO5* [30], on n'a établi une relation claire entre l'activité HAT de Gcn5 et l'activation de la transcription que lorsqu'on a observé que les mutations abolissant l'activité HAT de Gcn5 affectaient aussi l'activation de la transcription *in vivo* [31, 32]. Il a aussi été démontré que les complexes contenant Gcn5 sont recrutés par des coactivateurs et acétylent les nucléosomes au site de recrutement pour activer la transcription [33]. Comme discuté dans la section sur Taf1, SAGA est plus spécifiquement impliqué dans la régulation de gènes induits par les stress environnementaux [19, 20].

En 2002, deux groupes ont identifié un nouveau complexe dont Gcn5 fait partie : SLIK/SALSA [34, 35]. Ce complexe est physiquement et fonctionnellement similaire au complexe SAGA, mais est aussi impliqué dans la réponse rétrograde, un mécanisme

touchant la régulation de gènes nucléaires dans des souches où la mitochondrie est dysfonctionnelle.

La localisation génomique de Gcn5 a été étudiée par le groupe de Robert en 2004 [36]. Ils ont découvert que Gcn5 était présent à la plupart des promoteurs et que l'abondance de Gcn5 à ces promoteurs était positivement corrélée au niveau transcriptionnel des régions codantes associées. Ils ont aussi observé que Gcn5 était recruté au promoteur de gènes inactifs lors de leur activation.

1.1.1.4 *ESAI*

La HAT Esa1 (*Essential Sas2-related Acetyltransferase*) a été découverte par homologie de séquence à la HAT humaine Tip60 en 1998. Il s'agit de la seule HAT essentielle chez la levure. Elle fait partie de la famille MYST (*MOZ, YBF2/SAS3, SAS2, Tip60*) et a comme substrat l'histone H4 [37]. Esa1 est membre du complexe NuA4, qui acétyle les quatre résidus lysines de l'histone H4 (K5, 8, 12, 16) et joue un rôle dans la régulation de la transcription et du cycle cellulaire [38, 39]. Esa1 est aussi importante pour la transcription des gènes encodant les protéines ribosomales [40] et la réparation des cassures double brins [41].

Esa1, comme Gcn5, est recruté au promoteur des gènes transcrits et sa présence est positivement corrélée à leur taux d'activité transcriptionnelle. Elle est aussi recrutée au promoteur de gènes inactifs lorsqu'ils sont activés [36].

Cette protéine fait aussi partie d'un autre complexe composé de quelques membres de NuA4, Piccolo. Ce complexe ne serait pas recruté spécifiquement aux promoteurs, mais serait plutôt responsable de l'acétylation globale des histones H4 chromatiniennes [42, 43]. Il y aurait aussi compétition entre Esa1 et Rpd3L pour l'état d'acétylation de la lysine 12 de

l'histone H4. Cette compétition serait importante pour déterminer le niveau d'expression de plusieurs gènes [44].

Esa1 joue aussi un rôle important dans l'incorporation de la variante H2A.Z en l'acétylant en K14 lors de son intégration, catalysée par le complexe SWR1-C, en remplacement de H2A. Fait intéressant, SWR1-C et NuA4 partagent quatre sous-unités et leurs fonctions se retrouvent réunies chez l'humain dans le complexe Tip60, qui serait né au cours de l'évolution de la fusion de SWR1-C et NuA4 [45]. L'impact de la substitution de H2A pour H2A.Z est encore l'objet de débats, mais comme cette variante est globalement localisée aux promoteurs elle jouerait un rôle dans l'activation de la transcription [46], ce qui correspondrait au rôle connu de NuA4.

1.1.1.5 SAS3

Le gène *SAS3* a été identifié lors d'un criblage visant à trouver des amplificateurs des défauts d'inhibition épigénétique de la transcription causés par une délétion *sir1Δ* [47]. L'analyse de la séquence de ce gène a mis en évidence des domaines ressemblant à ceux d'acétyltransférases connues, fournissant un indice sur sa fonction. Sas3 fait partie de la famille MYST, comme Sas2 et Esa1. L'activité HAT de Sas3 vis-à-vis de l'histone H3 a été confirmée *in vitro* peu après [48].

Le complexe NuA3, dont Sas3 est la sous-unité HAT, acétyle préférentiellement les lysines 14 et 23 de l'histone H3. Fait intéressant, NuA3 comprend aussi Taf14, un membre des complexes TFIID, TFIIF, INO80 et SWI/SNF, et interagit avec le complexe FACT, deux indices qui laissent supposer un rôle d'activation de la transcription pour Sas3 [49]. En 2001, un groupe étudiant les interactions génétiques entre les différentes HATs a observé une perte globale d'acétylation de l'histone H3 ainsi qu'un arrêt en phase G2/M chez un double mutant *gen5Δ / sas3Δ*. Ces phénotypes ressemblent à ce qu'on observe chez

un mutant *esal1Δ*, sont spécifiquement dus à la perte d'activité HAT de ces deux protéines et ne sont pas observés dans d'autres combinaisons de mutants de HATs [50].

Le mécanisme de recrutement de NuA3 à ses sites d'activité a été éclairci lorsque l'importance du domaine PHD de la protéine Yng1, une sous-unité de NuA3, a été révélée [51]. Ce domaine lie la lysine 4 méthylée de l'histone H3, une marque de transcription active, positionnant ainsi NuA3 à la région proximale au site d'initiation de la transcription de ces gènes et permettant à Sas3 d'acétyler les queues N-terminales des histones H3 présentes [52].

Finalement, un groupe a démontré que les HATs Gcn5 et Sas3 sont recrutées à des gènes similaires et que leur liaison est corrélée positivement avec le niveau de transcription desdits gènes [53]. Toutes ces données confirment le rôle de Sas3 dans l'activation de la transcription.

1.1.1.6 HAT1

La protéine Hat1 a été caractérisée en 1995 comme une HAT spécifique aux lysines 5 et 12 de l'histone H4 [54]. Elle requiert un compagnon, Hat2, pour lier et acétyler les histones H4 libres dans le cytoplasme et le noyau, mais pas ceux intégrés dans la matrice chromatinienne [55, 56]. Fait intéressant, la séquence protéique de Hat2 ressemble à celle de Msi1, un membre du complexe CAF-1 responsable de l'assemblage des dimères H3/H4 dans le nucléosome lors de la réplication de l'ADN.

L'acétylation en K5 et K12 de l'histone H4 par Hat1 est importante pour le maintien de l'inhibition épigénétique de la transcription aux télomères. Une délétion *hat1Δ* ou *hat2Δ* combinée à l'ablation de la queue N-terminale de l'histone H3 dé-réprime l'expression des gènes situés dans ces régions. Toutefois, la présence même d'un seul résidu lysine

acétylable sur cette queue N-terminale suffit à supprimer ce phénotype [6]. Ce rôle contraste avec l'activité de stimulation de la transcription typique aux autres HATs.

La découverte en 2004 d'un partenaire additionnel pour le complexe nucléaire Hat1/Hat2 (HAT-B), la chaperonne d'histones H3/H4 Hif1, supporte l'idée d'un rôle de Hat1 dans l'assemblage de la chromatine [57, 58]. Cette hypothèse est renforcée par la découverte d'une interaction entre les complexes HAT-B et Asf1, une chaperonne de dimères H3/H4 aussi impliquée dans la déposition de nucléosomes lors de la réplication et de la transcription [59, 60].

1.1.1.7 *RTT109*

Cette protéine est la HAT caractérisée le plus récemment. Elle est associée à la chaperonne des histones H3/H4, Vps75, et est responsable, avec l'aide d'Asf1, de l'acétylation de la lysine 56 des histones H3 nouvellement synthétisées [61-66]. Tout comme Gcn5, elle est aussi capable d'acétyler la lysine 9 de l'histone H3, ne requérant qu'Asf1 pour ce faire [59].

La lysine 56 de l'histone H3 est acétylée de l'interphase à la phase S [67, 68]. Elle est par la suite déacétylée en phase G2 par Hst3 et Hst4 [69, 70]. La régulation temporelle stricte de cette marque d'acétylation est très importante, comme le démontrent les phénotypes observés chez les cellules ayant des défauts d'acétylation de H3K56 : sensibilité accrue aux agents génotoxiques et instabilité génomique prononcée [67, 68].

Deux groupes ont découvert que l'acétylation de H3K56 par Rtt109 aurait aussi une autre fonction indépendante de la réplication de l'ADN, mais impliquant toujours l'interaction avec les chaperonnes de H3/H4 : l'acétylation de ce résidu serait importante à l'activation de la transcription. Le premier groupe a observé une meilleure activation de la transcription causée par un désassemblage plus efficace du tétramère H3/H4 acétylé en

H3K56 des promoteurs de gènes transcrits [71], tandis que le second groupe démontre que cette acétylation permet la transcription efficace de matrices hétérochromatiniennes [72].

Le fil conducteur liant toutes ces fonctions de Rtt109 et de l'acétylation de H3K56 est la localisation spatiale de ce résidu sur le nucléosome : il est situé à l'interface où l'ADN contacte le nucléosome. Son acétylation déstabiliserait donc l'interaction nucléosome-ADN et faciliterait le travail des chaperons responsables de l'assemblage de la chromatine.

1.1.1.8 *SAS2*

Le gène *SAS2* a été identifié dans le même crible visant à identifier des éléments amplifiant les défauts d'inhibition épigénétique de la transcription que *SAS3* [47]. La délétion de *Sas2* dé-réprime l'expression de gènes situés aux télomères et au locus HML dans une souche *sir1Δ* et supprime la perte d'inhibition due à une mutation des motifs de liaison de Rap1 et d'Abf1 au locus HMR [47, 73, 74].

Sa fonction HAT a été inférée par homologie de séquence. La protéine *Sas2* fait partie de la famille MYST, tout comme *Esa1* et *Sas3*. Elle forme un complexe avec *Sas4* et *Sas5* et une délétion de n'importe laquelle des trois sous-unités de ce complexe donne les mêmes phénotypes qu'une simple délétion de *Sas2* [74]. L'intégrité de ce complexe est aussi requise pour l'activité HAT de *Sas2*, qui cible préférentiellement la lysine 16 de l'histone H4 [75]. Ce complexe est associé à la chaperonne d'histones *Asf1*, ce qui laisse supposer une activité spécifique aux histones en solution [74]. Cette préférence a été démontrée, mais on observe aussi l'acétylation d'histones faisant partie de la matrice chromatinienne, même si cette activité est très faible comparée à celle d'*Esa1* [75, 76].

L'activité de *Sas2* à la lysine 16 de l'histone H4 est requise pour contrer la diffusion de *Sir3*, donc de l'inhibition épigénétique de la transcription, sur la région sous-télomérique

[77, 78]. L'équilibre entre Sas2, qui acétyle H4K16, et Sir2, qui le déacétyle, spécifie les régions silencieuses du génome, telles que les télomères et les loci HM. Ce mécanisme passe par la liaison de Sir3 au résidu H4K16 déacétylé et son oligomérisation subséquente, ce qui étend la région silencieuse au-delà de la normale [77]. L'acétylation de H4K16 serait aussi requise pour l'intégration de la variante H2A.Z, qui agit en synergie avec H4K16 pour prévenir la diffusion de l'hétérochromatine [79].

1.2 Les histone déacétylases

Les histone déacétylases (HDACs) sont une famille d'enzymes présents chez tous les eucaryotes. Ils catalysent le clivage du groupement acétyle des résidus ϵ -N-acétyle lysines, redonnant ainsi leur charge positive aux lysines affectées. Les interactions électrostatiques internucléosomales et nucléosomes-ADN sont renforcées lorsque ces lysines sont positivement chargées, ce qui confère à l'environnement chromatinien une conformation plus compacte pouvant restreindre l'accès aux facteurs de transcription. Les HDACs ont donc une fonction répressive vis-à-vis la transcription.

On dénombre à ce jour chez la levure dix membres de cette famille, regroupés en trois classes, I, II et III, selon leur degré de similarité ainsi que leur mécanisme enzymatique. Je décrirai en détail dans cette prochaine section les caractéristiques des membres des différentes classes en insistant particulièrement sur Rpd3, qui est le sujet de l'article présenté au chapitre 3.

1.2.1 Les HDACs de classe I

La classe I est définie par homologie de séquence à *RPD3* et comprend les gènes *RPD3*, la première HDAC caractérisée, *HOS1* et *HOS2*. La fonction HDAC de ces enzymes peut être inhibée par la trichostatine A.

1.2.1.1 *RPD3*

1.2.1.1.1 *Découverte et caractérisation*

Le gène *RPD3* (*Reduced Potassium Deficiency 3*) a été identifié lors d'un criblage à la recherche de mutants pouvant supprimer les phénotypes associés à la délétion de *TRK1*, un transporteur à haute affinité de potassium [80]. La délétion de *RPD3*, ainsi que celle de *RPD1*, causait une augmentation de l'activité de *TRK2*, un gène encodant un autre transporteur de potassium. Le même groupe a démontré l'année suivante que *RPD3* et *RPD1* (*SIN3*) étaient impliqués dans la régulation négative et positive de plusieurs gènes liés aux réponses aux stress environnementaux, de sporulation et sexuels [81, 82]. Dans la même veine, un autre groupe a rapporté qu'une délétion de *RPD3* ou de *SIN3* pouvait rétablir l'expression du locus *HO* dans un mutant *swi5Δ* [83]. Ce n'est que quelques années plus tard que la fonction HDAC de Rpd3 fut inférée par homologie de séquence à deux protéines humaines ayant cette activité [84]. La fonction HDAC de Rpd3 a été confirmée la même année lorsque deux complexes HDACs distincts, l'un contenant Rpd3 et l'autre Hda1, furent purifiés [85]. Cette étude a aussi démontré que le niveau d'acétylation des histones H3 et H4 augmentait lorsque *RPD3* ou *HDA1* étaient délétées.

Maintenant que la fonction de *RPD3* était plus claire, une question importante demeurait : vu que ni *RPD3* ni *SIN3* ne possèdent de domaine de liaison à l'ADN, qu'est-ce qui les recrute à leurs cibles? Kadosh a démontré en 1997 qu'*UME6*, un répresseur de gènes de méiose pouvant se lier aux séquences URS1, pouvait recruter un complexe stable formé de Rpd3 et de Sin3 et que ce recrutement était nécessaire et suffisant à la répression des gènes ciblés par Ume6 [86]. Il a par la suite démontré que l'activité HDAC de *RPD3* était nécessaire à cette répression [87]. Parallèlement, le groupe de Kasten a purifié la même année un complexe stable d'environ 2 MDa comprenant entre autres Rpd3 et Sin3 [88]. Dans une autre étude publiée en 1998, Kadosh a caractérisé les substrats de *RPD3* comme étant les histones H3 et H4 et plus particulièrement les acétyl-lysines 5 et 12 de l'histone H4. Cette déacétylation ciblée était restreinte aux quelques nucléosomes adjacents au site de recrutement [89]. Ces résultats ont été indépendamment confirmés par Rundlett et son groupe quelques semaines plus tard [90]. Deux études démontrant les effets opposés de

la HDAC *RPD3* et la HAT *GCN5* au locus *HO* [91] et aux gènes régulant la méiose [92] ont renforcé le lien entre la modulation de l'état d'acétylation des nucléosomes par *RPD3* et le niveau de transcription des régions codantes associées aux promoteurs ciblés par cette HDAC. Des données intéressantes fournies par Deckert en 2002 indiquent que la répression observée est due à l'inhibition du recrutement de facteurs de transcription tels que la « TATA-binding protein » (TBP), le complexe Swi/Snf et le complexe HAT SAGA par la déacétylation locale causée par Rpd3 [93].

1.2.1.1.2 Les fonctions de Rpd3

La fonction de Rpd3 a été étudiée en détail depuis sa découverte. Une des premières observations a été faite en 1994 quand le groupe de Stillman a découvert qu'une mutation de *RPD3* rétablissait l'expression du locus *HO* dans un mutant *swi5Δ* [83]. L'implication de Rpd3 dans la régulation de la méiose a aussi été caractérisée rapidement : le complexe Rpd3/Sin3 réprime activement des gènes encodant des protéines régulatrices de la méiose, comme *IME2*, via son corépresseur Ume6 [86], et on observe des problèmes de méiose et de sporulation dans un délétant *rpd3Δ* [94]. L'implication de Rpd3 dans le cycle cellulaire de la levure fut ensuite étudiée plus en profondeur quand on a remarqué des problèmes de synchronisation dans l'activation des origines de réplication tardifs dans un mutant *rpd3Δ* [95, 96]. En 2004, Robert et ses collègues ont caractérisé à l'échelle génomique les promoteurs liés par Rpd3. Ils ont découvert que Rpd3 était lié aux promoteurs de gènes régulant le cycle cellulaire et que cette liaison dépendait de la présence de certains des principaux régulateurs du cycle cellulaire tel que le complexe Swi4/Swi6 (SBF) [36]. Des études plus récentes ont confirmé que Rpd3 est en effet impliqué dans la régulation de ces gènes en étant recruté aux promoteurs par Whi5, un corépresseur homologue à Rb qui inhibe l'activité de SBF et MBF [97, 98].

Rpd3 a aussi été impliqué dans la régulation des réponses aux stress environnementaux tels le choc thermique [99] et le choc osmotique [100]. Un modèle bien étudié est celui de la réponse au manque de nutriments impliquant le chemin métabolique TOR (*Target of Rapamycin*). Rpd3 y est responsable de la répression de gènes encodant des protéines ribosomales [101] ainsi que d'autres gènes ciblés par cette réponse [102]. La structure du locus d'ADN ribosomal est aussi affectée par Rpd3 lors de cette réponse [103].

Finalement, Rpd3, initialement caractérisé comme un répresseur de la transcription, tient un rôle paradoxal dans l'inhibition épigénétique de la transcription. En effet, un mutant *rp3Δ* peut supprimer la perte de cette inhibition au locus HMR [104]. On observe même un rétablissement de la compétence à la conjugaison dans des souches *Mata* chez ces mutants, ce qui confirme cet effet [105]. La suppression de l'inhibition épigénétique est aussi présente aux loci d'ADN ribosomal et sous-télomériques [106]. Une hypothèse expliquant ces observations surprenantes implique la déacétylation par Rpd3 du substrat du complexe Sir2, la lysine 16 de l'histone H4. Ce faisant, Sir2 ne serait pas en mesure d'agir pour inhiber la transcription dans ces régions [107]. Toutefois, d'autres études ont démontré que l'hypoacétylation de ces régions est suffisante pour l'inhibition épigénétique de la transcription modulée par les sirtuines [108, 109]. L'hyperacétylation causée par la perte de Rpd3 serait donc un obstacle à l'établissement de ces régions silencieuses.

1.2.1.1.3 Identification des protéines associées à Rpd3

Depuis que Kadosh et Kasten ont démontré que Rpd3 et Sin3 forment un complexe stable, plusieurs autres groupes ont étudié par des méthodes génétiques et directes les interactions entre Rpd3 et diverses autres protéines : Sap30 [110], Sds3 [111], Pho23 [112], Ume1 [113], Cti6 [114] ainsi que Rxt2 [115]. Par contre, ce n'est qu'avec l'avènement de méthodes protéomiques à haut débit telles que la technologie multidimensionnelle d'identification de protéines (MudPIT) qu'il fût possible de dresser une liste « définitive »

des composantes du complexe Rpd3. Les groupes de Carrozza et de Keogh ont publié en 2005 des résultats très convaincants démontrant que Rpd3, contrairement à ce que l'on croyait, ne formait pas qu'un seul, mais bien deux complexes fonctionnellement indépendants [116-118]. De plus, ces deux complexes, bien que partageant les sous-unités Rpd3, Sin3 et Ume1, posséderaient des fonctions indépendantes. En effet, le plus gros des deux, nommé Rpd3L pour « large », est composé, en plus du sous-complexe commun mentionné ci-dessus, de protéines associées à la fonction répressive de Rpd3, nommément Sds3, Pho23, Dep1, Rxt2, Ash1, Sap30, Cti6, Rxt3 et Ume6.

L'autre complexe, plus petit donc nommé Rpd3S pour « small », ajoute seulement deux protéines, Rco1 et Eaf3, au sous-complexe Rpd3-Sin3-Ume1. Il est très intéressant de noter qu'Eaf3 fait aussi partie du complexe HAT NuA4 [119] et que cette protéine comprend un chromodomaine, un type de domaine protéique habituellement impliqué dans des interactions avec des lysines méthylées [120]. La **Figure 1-3** illustre la composition et la localisation de Rpd3L et Rpd3S à leurs gènes cibles.

1.2.1.1.4 Le complexe Rpd3S

La découverte de l'existence de deux complexes contenant Rpd3 a suscité beaucoup d'intérêt. Le complexe Rpd3L étant clairement responsable de l'activité de régulation de la transcription au niveau du promoteur de gènes ciblés par ses corépresseurs, quelle est la fonction de Rpd3S? Une piste intéressante que plusieurs groupes ont suivie est celle de la sous-unité de Rpd3S, Eaf3. Cette protéine possède un domaine de liaison aux lysines méthylées, nommé chromodomaine. Le chromodomaine d'Eaf3 pourrait donc se lier à au moins un des trois résidus lysines méthylés présents chez la levure, nommément les lysines 4, 36 et 79 de l'histone H3. Ces lysines sont méthylées par les histone

méthyltransférases (HMT) Set1 (K4), Set2 (K36) et Dot1 (K79), et ce de manière dépendante de la transcription par l'ARN polymérase II [121].

Plusieurs études ont déterminé que le complexe Rpd3S est capable de se lier à des peptides [117, 122] et à des mononucléosomes immunopurifiés [118] ou reconstitués *in vitro* [123] méthylés en H3K36, et que cette liaison dépend de l'intégrité du chromodomaine d'Eaf3. La sous-unité Rco1 possède quant à elle un domaine « plant homeobox » (PHD) qui est lui aussi essentiel à cette liaison [123]. On avait préalablement observé qu'une délétion d'Eaf3 abaissait le niveau d'acétylation des nucléosomes associés aux promoteurs et augmentait celui des régions codantes [124]. Les trois études mentionnées ci-dessus ont déterminé que ces altérations des niveaux d'acétylation étaient similaires pour des délétions d'Eaf3 ou de son chromodomaine, de la HMT Set2, de Rco1 ou dans un mutant où l'histone H3 est non méthylable en K36 (H3K36A) [117, 118, 122]. L'intégrité du système Rpd3S/Set2 est donc essentielle à la déacétylation des régions codantes activement transcrites.

Le complexe Rpd3S se lierait donc à la lysine 36 méthylée de l'histone H3 par le biais du chromodomaine de sa sous-unité Eaf3 pour déacétyler les régions codantes transcrites. L'impact fonctionnel de ce mécanisme est devenu évident lorsque les groupes de Struhl et de Workman ont découvert que des mutants de Rpd3S ou de Set2 démontrent un phénotype d'initiation de la transcription à l'intérieur des régions codantes [117, 122]. Ce phénomène, aussi nommé « transcription cryptique », fut d'abord observé en 2003 dans des souches où le complexe FACT [125], nécessaire pour faciliter la transcription par l'ARN polymérase II sur la matrice chromatinienne, ou la protéine Spt6 [126], associée au facteur d'élongation DSIF [127, 128], étaient mutés. Lors de la transcription, l'ARN polymérase II et ses facteurs associés bouleversent la structure chromatinienne présente sur la région codante pour la rendre plus permissive à l'élongation. La transcription cryptique aurait lieu lorsque cet environnement n'est pas rétabli à son état prétranscriptionnel, ce qui exposerait des sites d'initiation fortuits à l'ARN polymérase II. La présence de ce

phénomène dans les souches mutées en Rpd3S ou Set2 indique que ces protéines jouent probablement un rôle important dans le rétablissement d'une structure chromatinienne non permissive sur la région codante après le passage de l'ARN polymérase II.

Cette hypothèse est supportée par des données génétiques où il est démontré que des mutations de Rpd3S ou de Set2 suppriment le phénotype létal de mutants *bur1Δ* [118]. Le complexe Bur1/Bur2 est un facteur d'élongation positif qui phosphoryle plusieurs cibles, dont le facteur d'élongation DSIF [129, 130]. Le complexe DSIF est impliqué dans la méthylation de l'histone H3 en K4 et K79 car il recrute le complexe PAF, un facteur d'élongation nécessaire à ces méthylations [131] et au recrutement de Set2 sur les régions codantes activement transcrites [132]. Fait intéressant, PAF interagit aussi avec le complexe FACT mentionné ci-dessus. De plus, une étude récente de Quan et Hartzog [133] mentionne que des mutations dans le complexe PAF, une abolition de la méthylation en K4 ou K36 ou un recrutement réduit de Rpd3S suppriment les phénotypes dus à une délétion de *SPT5*, un membre du complexe DSIF. Toutes ces données démontrent clairement que le système Rpd3S/Set2 influe négativement sur la transcription, comme le suppose leur rôle présumé, et que cette activité est importante pour le maintien de l'identité des régions codantes en réprimant la transcription aberrante aux promoteurs cryptiques. L'état des connaissances sur la fonction de Rpd3S est résumé dans la **Figure 1-4**.

Je présente dans le troisième chapitre de cette thèse des résultats remettant en question le mode de recrutement de Rpd3S. En effet, j'ai observé que l'interaction Eaf3-H3K36 triméthylé ne joue qu'un rôle modeste dans la stabilité de l'association Rpd3S-chromatine *in vivo*, mais est par contre requise pour l'activité HDAC de Rpd3S. De plus, le recrutement de Rpd3S aux régions transcrites passe par l'hyperphosphorylation du domaine C-terminal de l'ARN polymérase II. Le complexe DSIF, quant à lui, joue un rôle antagoniste à association. Ces nouveaux résultats concordent avec les données accumulées jusqu'ici et suggèrent un nouveau modèle pour l'activité de Rpd3S aux régions activement transcrites.

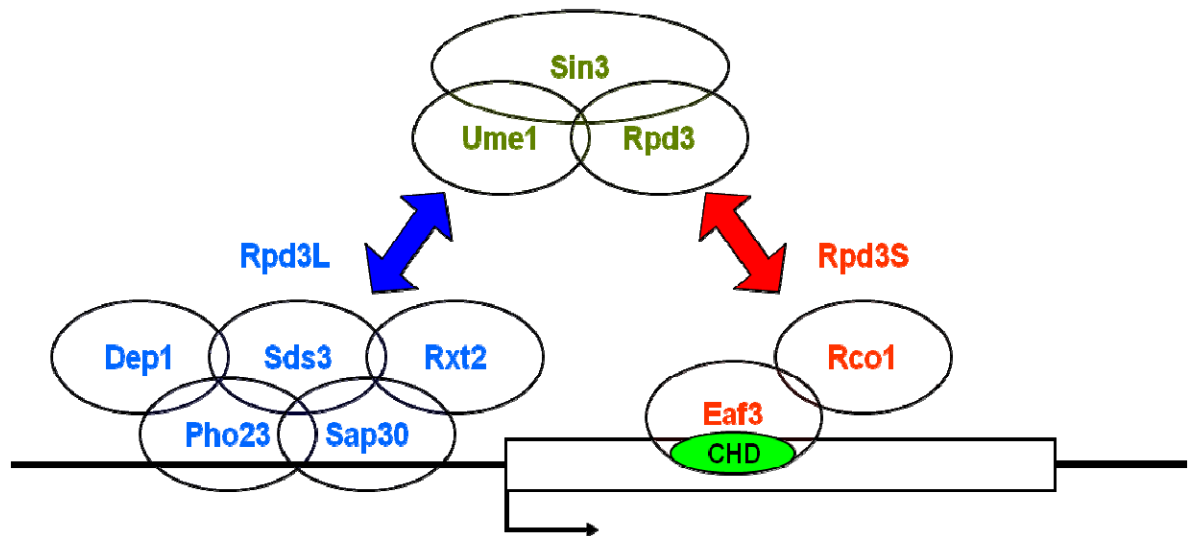


Figure 1-3 : Composition et localisation des complexes Rpd3L et Rpd3S.

Les trois sous-unités communes, en vert, sont associées aux membres de Rpd3L, en bleu, ou Rpd3S, en rouge. Selon les hypothèses courantes, Rpd3L se retrouve principalement au promoteur, où il déacétyle les nucléosomes pour réprimer l'expression du gène associé, tandis que Rpd3S est associé aux régions codantes de gènes activement transcrits via l'interaction entre la lysine H3K36 méthylée et le chromodomaine de sa sous-unité Eaf3 (CHD, en vert lime).

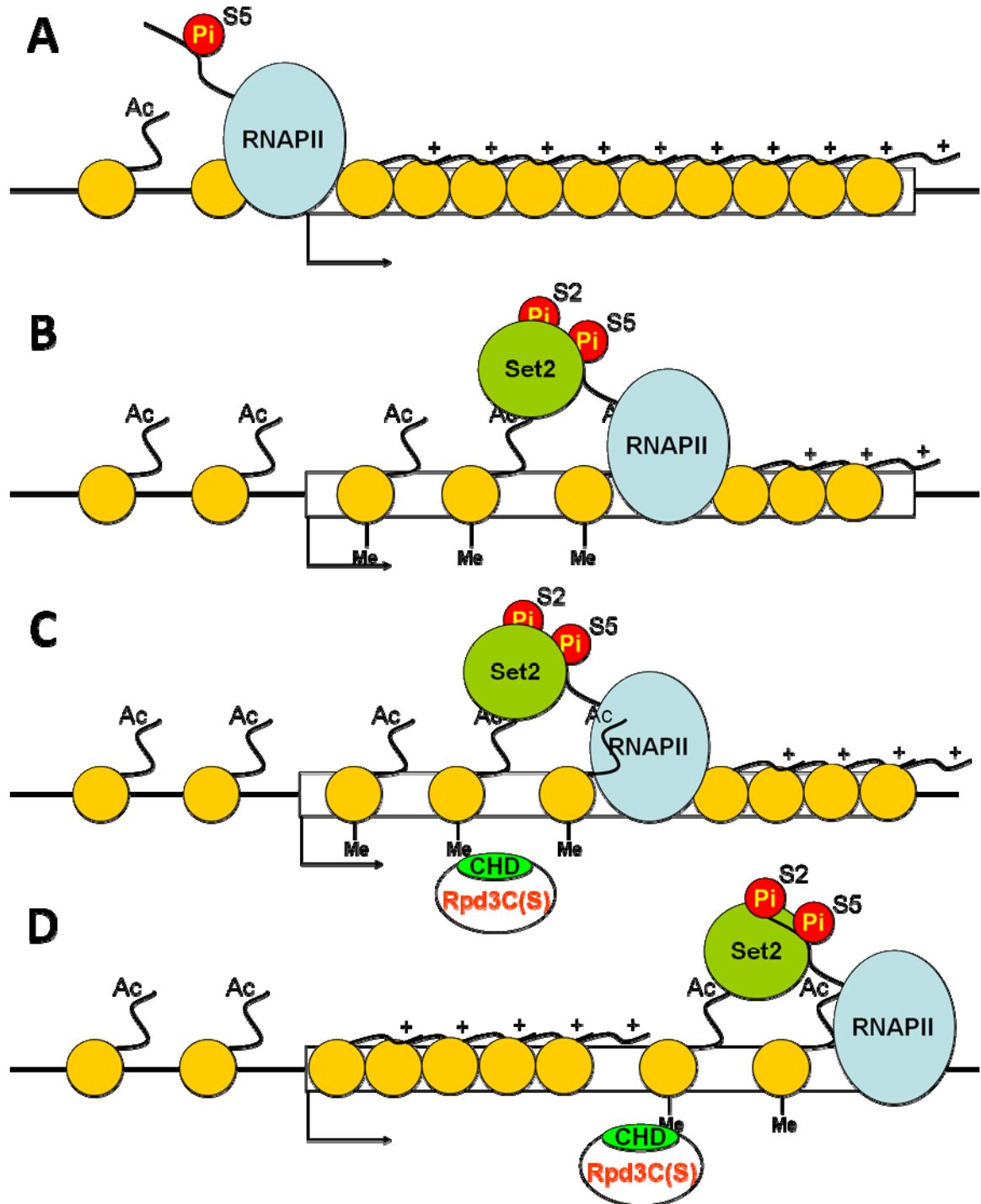


Figure 1-4 : Le modèle courant de la fonction de Rpd3S.

- (A) L'ARN polymérase II initie la transcription et son domaine C-terminal est phosphorylé en sérine 5.
- (B) Les nucléosomes présents sur la région codante sont acétylés par le complexe transcriptionnel. Set2 est recruté au domaine C-terminal de l'ARN polymérase II, maintenant phosphorylé aux sérines 2 et 5. Set2 triméthyle alors H3K36.
- (C) Rpd3S est recruté à la région transcrite via l'interaction entre le chromodomaine de sa sous-unité Eaf3 et le résidu H3K36 triméthylé par Set2.
- (D) Rpd3S déacétyle les nucléosomes sur la région codante, rétablissant ainsi la structure chromatinienne prétranscriptionnelle.

1.2.1.2 *HOS1*

La HDAC *HOS1* (*HDA One Similar 1*) n'est pas bien caractérisée. Elle a été découverte en 1996 par homologie de séquence avec *RPD3* et *HDA1* [85]. Hos1, Hos2 et Rpd3, sont recrutées aux promoteurs de gènes ciblés par le corépresseur Tup1-Ssn6 où elles déacétylent les nucléosomes pour en réprimer l'expression [134, 135]. Hos1 serait aussi ciblé au locus d'ADN ribosomal en compagnie de Hos3 où elle y déacétylerait préférentiellement la lysine 12 de l'histone H4.

1.2.1.3 *HOS2*

La protéine Hos2 a aussi été découverte par homologie de séquence avec *RPD3* et *HDA1* [85]. Comme Hos1, elle interagit directement avec le corépresseur Tup1-Ssn6 pour réprimer l'expression de ses gènes cibles [134, 135]. Hos2 est aussi une sous-unité du complexe Set3C et est impliqué dans la répression de la méiose [136]. On observe de plus une baisse de la sporulation dans un mutant homozygote *hos2Δ* [137].

Des études génomiques ont déterminé que Hos2 déacétylait spécifiquement les promoteurs et régions codantes de gènes encodant les protéines ribosomales [138]. L'activité de cette HDAC et du complexe Set3C est aussi nécessaire à l'induction efficace des gènes du chemin métabolique du catabolisme du galactose. On retrouve ce complexe sur les régions codantes des gènes induits et l'activité HDAC de Hos2 est requise pour leur activation [139].

Le mode de recrutement de Hos2 et du complexe Set3C passe probablement par le domaine PHD de la protéine Set3. Ce domaine serait capable de lier la lysine 4 doublement méthylée sur l'histone H3 et serait donc ciblé à la région 5' des gènes transcrits pour y déacétyler les nucléosomes et donc rétablir une structure de la chromatine réfractaire à la liaison de facteurs de transcription [140].

Nous verrons au chapitre quatre que Hos2 ne co-localise avec Set3 que sur les gènes les plus transcrits liés par Set3. Fait intéressant, on retrouve aussi Hst1 aux régions codantes de ces mêmes gènes. Cela implique forcément que le complexe Set3C n'est pas aussi stable et défini que l'on pensait.

1.2.2 Les HDACs de classe II

Cette classe regroupe les protéines ayant une homologie de séquence significative avec *HDA1*. Chez la levure, elle n'est composée que de *HDA1* et de *HOS3*. Ces protéines, tout comme celles de classe I, sont sensibles à la trichostatine A.

1.2.2.1 *HDA1*

La protéine Hda1 a été découverte et purifiée en 1996 par deux groupes. Il a été immédiatement déterminé qu'elle était distincte de Rpd3 [85] et qu'elle interagissait directement avec Hda3 [141]. Le groupe de Wu et ses collègues a découvert en 2001 que Hda1 était ciblée aux promoteurs par le corépresseur Tup1 pour réprimer entre autres certains gènes régulant la conjugaison, l'utilisation du glucose et de l'oxygène ainsi que la réponse aux stress environnementaux et aux bris d'ADN [142]. Par contre, aucune interaction directe entre Tup1-Ssn6 et Hda1 n'a été détectée *in vivo* [135]. Wu a par la suite découvert que Hda1 était capable d'interagir avec elle-même ainsi qu'avec Hda2 et Hda3 pour former un tétramère. Cette conformation est nécessaire pour l'activité HDAC de Hda1 [143].

Robyr a caractérisé en 2002 les sites où Hda1 déacétyle ses nucléosomes cibles et ce à l'échelle du génome [138]. Il a ainsi découvert que Hda1 déacétyle préférentiellement la lysine 18 de l'histone H3 dans les régions sous-téléométriques et réprime l'expression de

classes de gènes associées à différents stress. Toutefois, Green a démontré en 2004 que la déacétylation et la répression de la transcription ne vont pas nécessairement de pair puisque la déacétylation de K18 catalysée par Hda1 n'était pas suffisante pour réprimer l'expression de gènes où cette HDAC était recrutée par le corépresseur Tup1 [144].

1.2.2.2 *HOS3*

La HDAC *HOS3* (*HDA One Similar 3*) a très peu été étudiée. Nous savons toutefois qu'elle forme un homodimère ayant une activité intrinsèque, ce qui est unique chez les HDACs, et qu'elle est relativement insensible à la trichostatine A [145]. Le niveau d'acétylation au locus d'ADN ribosomal est accru, surtout pour les résidus K16 de l'histone H4 et K12 de l'histone H2B lorsqu'elle est déletée [138]. Hos3 aurait aussi une fonction dans la régulation du signal de mort cellulaire en déacétylant la lysine 11 de l'histone H2B pour permettre la phosphorylation de la lysine 10 de H2B par la kinase Set20 [146]. Cette HDAC serait aussi impliquée dans la régulation du cycle cellulaire, car elle interagit avec Rpd3 et Whi5 pour réprimer les gènes ciblés par SBF.

1.2.3 Les HDACs de classe III

Cette famille, composée de *SIR2* et de ses quatre homologues *HST1*, *HST2*, *HST3* et *HST4* (*Homolog of Sir Two 1-4*), se distingue des deux autres classes de HDACs principalement par leur mécanisme catalytique complètement différent. En effet, là où les HDACs de classe I et II hydrolysent simplement les résidus acétyl-lysine, celles de classe III, les sirtuines, procèdent par un mécanisme associant la déacétylation à l'hydrolyse de nicotinamide adénine dinucléotide (NAD). Cette réaction donne comme produits la lysine

déacétylée, un composé nommé O-acétyl-ADP-ribose, ainsi qu'une molécule de nicotinamide, elle-même un inhibiteur de cette réaction [147].

1.2.3.1 *SIR2*

Cette protéine, la première HDAC de classe III caractérisée, a été identifiée grâce à son rôle dans la répression des loci silencieux HML et HMR [148, 149]. Des mutants des facteurs SIR (*Silent Information Regulator*) permettent d'ailleurs l'entrée en méiose de souches haploïdes causée par l'expression aberrante des facteurs a et α [150]. Ces mêmes facteurs, ainsi que l'histone H4, sont aussi impliqués dans la répression de gènes présents dans les régions télomériques [151]. On associera plus tard l'état répressif de ces régions à leur niveau d'acétylation réduit et on impliquera Sir2 comme étant la protéine responsable du maintien de cette hypoacétylation [152]. Sir2 est aussi impliqué dans la régulation du locus d'ADN ribosomique par l'entremise du complexe RENT (*REGulators of Nucleolar Silencing and Telophase*) composé de Sir2, Net2 et Cdc14 [153-155].

En 1997, deux groupes ont caractérisé les protéines associées à Sir2. On réalise alors que Sir2 et Sir4 forment un complexe stable et que Sir3 pouvait étendre la région silencieuse à partir de ce complexe [156, 157]. On observe aussi que la surexpression de Sir2 ou de Sir3 est toxique pour la cellule car elle mène à la perte de chromosomes [158].

Le complexe Sir2 inhibe la transcription par un mécanisme en plusieurs étapes. Premièrement, le dimère Sir2/Sir4 lie ses sites de recrutement sur l'ADN [159]. L'activité enzymatique de Sir2 n'est pas nécessaire pour cette liaison à l'ADN mais l'est pour l'association et l'oligomérisation de Sir3 sur les régions silencieuses distales. L'étendue de la région couverte par l'effet inhibiteur du complexe Sir2 est régulée par l'état d'acétylation de la lysine 16 de l'histone H4. Cette lysine est acétylée par la HAT Sas2 et déacétylée par Sir2 [77, 78]. La déacétylation en K16 de l'histone H4 est nécessaire et suffisante à la

propagation de la région inhibée [160]. De plus, cette hypoacétylation ne doit pas nécessairement être catalysée par Sir2 [108, 109].

Quelques études démontrant un effet de Sir2 et des sirtuines sur la longévité ont mené à un intérêt considérable pour ce complexe dès 1999. On a d'abord démontré que des mutations de ce complexe menaient à une activité de recombinaison du locus d'ADN ribosomal considérablement accrue, diminuant ainsi la durée de vie de la levure [161, 162]. Sir2 a ensuite été impliqué dans l'extension de la longévité par restriction calorique [163, 164]. On a dit que ce phénomène était expliqué par l'augmentation de la respiration [165], par la déplétion de la nicotinamide [166, 167], ou par la baisse de la concentration en NADH [168]. Toutefois, ces études ont été contredites par plusieurs groupes qui ont tous démontré que l'effet d'extension de la durée de vie était dû à des mécanismes indépendants des sirtuines [169-175]. Ces nouvelles conclusions font maintenant consensus.

1.2.3.2 *HST1*

Hst1 a été identifiée en 1995 par homologie de séquence protéique avec Sir2. On a immédiatement remarqué que la délétion de cette protéine ou des autres gènes *HST* ne donnait pas les phénotypes typiquement associés à une mutation *sir2Δ*, mais que sa surexpression pouvait rétablir l'inhibition aux loci HML et HMR observés chez le mutant *sir2Δ* [176, 177].

La caractérisation des protéines associées à Hst1 a permis de caractériser deux complexes indépendants contenant cette protéine : le complexe Set3C, qui contient aussi la HDAC Hos2, et le complexe Hst1-Sum1. Les protéines Hst1 et Sum1 seraient impliqués dans la répression de gènes régulant la sporulation [178] et dans la régulation au locus HMR, où Hst1 déacétyle les nucléosomes pour permettre la condensation de la chromatine et donc de l'inhibition de la transcription [179]. Robert et son groupe ont observé en 2004

que Hst1 lie les mêmes promoteurs que Sum1 cible, tels que les gènes de sporulation et de synthèse du NAD. Il démontre aussi que Hst1 est capable de déacétyler les promoteurs de ces gènes [36].

Une étude récente de Kim et Buratowski rajoute une fonction à Hst1 [140]. En effet, ils avancent que Set3C serait recruté à la région 5' des régions codantes transcrites par l'interaction entre la lysine 4 diméthylée de l'histone H3 et le domaine PHD de Set3. Les HDACs Hos2 et Hst1, toutes deux membres de Set3C, déacétyleraient les nucléosomes de cette région pour rétablir une conformation chromatinienne neutre à la matrice transcriptionnelle. Toutefois, bien que l'effet de Hos2 dans ce système soit relativement clair, celui de Hst1 reste ambigu car les phénotypes transcriptionnels observés dans un mutant *hst1Δ* ne correspondent pas à ceux de mutants des autres sous-unités du complexe Set3C. Il est probable que cette différence soit due aux autres rôles que joue Hst1 dans la cellule, par exemple son implication avec Sum1.

Dans le chapitre quatre je décortique par CHIP-chip les relations entre Hst1 et le complexe Set3C. J'y confirme que Hst1 est surtout présente au promoteur de gènes ciblés par Sum1. Par contre, on retrouve aussi Hst1 sur la région codante d'environ le quart des ces gènes, qui sont par ailleurs parmi les plus transcriptionnellement actifs du génome. Cette localisation concorde avec celle de Set3 et de Hos2, tous deux membres du complexe Set3C. Toutefois, le chevauchement entre Hst1, Set3 et Hos2 aux régions codantes n'est détecté qu'à une fraction des gènes liés par Set3, ce qui explique probablement les phénotypes ambigus observés par les études susmentionnées. De plus, l'interaction de Set3C sur ces régions codantes ne dépend pas de la méthylation de H3K4, contrairement à ce que Kim *et coll.* avance.

1.2.3.3 *HST2*

La protéine Hst2, comme son homologue Sir2, est impliquée dans l'inhibition épigénétique de la transcription et déacétyle préférentiellement la lysine 16 de l'histone H4 [180]. Les souches portant une délétion *hst2Δ* ne démontrent pas de phénotype évident, mais une surexpression de ce gène a un impact sur la répression de la transcription aux télomères ainsi que sur la fréquence de recombinaison au locus d'ADN ribosomal [181, 182]. Cette protéine joue aussi un rôle dans la répression des gènes sous-télomériques *FLO* [183]. Finalement, Hst2 possède un mode de régulation de son activité unique chez les HDACs de la levure. En effet, la majorité des molécules de Hst2 sont séquestrées dans le cytoplasme, ce qui les empêche évidemment d'agir sur leurs cibles génomiques [181, 182, 184].

1.2.3.4 *HST3* et *HST4*

Ces deux HDACs sont fonctionnellement redondantes et ne montrent de phénotypes que lorsque délétees toutes deux. Comme d'autres sirtuines, elles sont impliquées dans le maintien de l'inhibition épigénétique de la transcription aux régions télomériques. Elles jouent aussi un rôle important dans la progression du cycle cellulaire, le maintien de l'intégrité génomique et la résistance à la radiation et aux agents génotoxiques [70, 176].

Hst3 et Hst4 ont pour cible la lysine K56 acétylée de l'histone H3 [69, 70, 185]. Cette modification, catalysée par la HAT Rtt109, est présente chez les nucléosomes nouvellement assemblés et intégrés dans la chromatine en phase S. La déacétylation rapide de ce résidu par Hst3 et Hst4 est très importante pour l'intégrité du génome : les phénotypes d'instabilité génomique des doubles mutants *hst3Δ / hst4Δ* sont supprimés par une mutation de la lysine 56 en un résidu non acétylable et par la délétion de la chaperonne d'histones Asf1 [70, 186]. L'hyperacétylation en K56 résultant de délétions de Hst3 et Hst4 cause des

problèmes d'inhibition de la transcription aux loci silencieux, malgré la présence ou même la surexpression de Sir2. Ce phénotype n'est supprimé que par la délétion de Rtt109 [186].

1.3 Les HDACs chez l'humain

La découverte d'homologues humains de Rpd3 en 1997 a amorcé une véritable révolution de l'étude de la régulation épigénétique de la transcription chez les mammifères. Auparavant, la méthylation de l'ADN, plus particulièrement des îlots CpG, était considérée comme le mode dominant de régulation épigénétique de la transcription chez les mammifères. On sait maintenant que ce mécanisme n'en est qu'un parmi d'autres utilisés par ces cellules pour identifier les régions où la transcription devrait être réprimée.

Le contrôle du niveau d'acétylation des nucléosomes associés aux promoteurs est, comme chez la levure, étroitement associé à l'activité transcriptionnelle. Ce champ d'études a connu un essor incroyable depuis que plusieurs pathologies ont été associées à une répression transcriptionnelle aberrante causée par l'hypoacétylation des promoteurs affectés. L'inhibition des HDACs responsables est d'ailleurs une voie thérapeutique prometteuse pour le traitement de ces maladies.

Les HDACs humaines ont été découvertes par homologie de séquence avec celles de la levure. On en dénombre toutefois presque deux fois plus chez l'humain (18 contre 10 chez la levure). Ces protéines sont regroupées en classes correspondant à ce que l'on retrouve chez la levure (classes I à III), à l'exception de HDAC11 qui est dans une classe à part (classe IV).

L'activité des HDACs humaines est très strictement régulée par une panoplie de mécanismes allant de la phosphorylation à l'utilisation d'isoformes de sous-unités dans la composition des complexes en passant par la séquestration des HDACs dans le cytoplasme. La multiplication et la complexification de ces modes de contrôle par rapport à la levure sont des effets directs de la spécialisation de ces protéines chez l'humain pour des fonctions qui ne se retrouvent pas chez la levure (régulation spécifique à certains tissus, mécanismes développementaux, etc.). En revanche, les principes et fonctions de base demeurent les mêmes pour les quelques homologues fonctionnels qui se retrouvent à la fois chez ces deux

organismes, par exemple Rpd3 / HDAC1/2 et les sirtuines, ce qui fait de la levure un excellent modèle pour étudier et caractériser ceux-ci.

Je passerai maintenant en revue les différentes classes de HDACs chez l'humain en faisant lorsque possible des liens avec leurs homologues fonctionnels chez la levure. Un résumé des différentes HDAC humaines et de leurs homologues chez la levure est présenté dans le **Tableau 1** à la fin de cette section.

1.3.1 La classe I

Cette classe est définie par la forte homologie de ses membres à la HDAC Rpd3 de la levure. Elle comprend quatre protéines, HDAC1, HDAC2, HDAC4, et HDAC8, toutes exprimées de manière ubiquitaire chez l'humain [187]. Comme chez la levure, elles ne possèdent pas de domaine de liaison à l'ADN; leur fonction est donc déterminée par les complexes desquels elles font partie. Finalement, les HDACs de classe I humaines sont sensibles à la trichostatine A (TSA), un inhibiteur général des HDACs aussi efficace chez la levure.

1.3.1.1 *HDAC1 et HDAC2*

Ces protéines sont les plus proches homologues de Rpd3 chez l'humain et sont presque identiques (82 % d'identité de séquence) [188]. Elles interagissent pour former le cœur catalytique de plusieurs complexes, tels que mSin3a, mSin3b (*Mammalian Sin3a* et *Sin3b*), NuRD (*Nucleosome Remodeling Deacetylase*) et CoREST (*Corepressor of Repressor Element 1 Silencing Transcription factor*) [189]. Ces complexes, tout comme Rpd3L chez la levure, sont recrutés aux promoteurs de gènes ciblés par des corépresseurs pour en réprimer l'expression.

Les complexes mSin3a et mSin3b sont virtuellement identiques à l'exception de la protéine plate-forme Sin3 qui peut s'y retrouver sous deux isoformes, Sin3a et Sin3b. Ils partagent le même cœur de sous-unités que NuRD, soit HDAC1, HDAC2, RbAp48 (*Rb-Associated Protein 48*, capable de lier l'histone H4) et RbAp46. Les complexes mSin3 possèdent aussi plusieurs autres sous-unités dont la protéine ING2 qui, comme le complexe Set3C chez la levure, est capable de se lier au résidu H3K4 méthylé. Cette interaction est fonctionnellement importante, car elle lie la fonction HDAC de mSin3 à ce résidu méthylé. Le parallèle avec Set3C est d'autant plus intéressant que plusieurs sous-unités du complexe Set3C semblent interagir avec Rpd3L [190]. Bien que la composition des complexes soit différente de l'humain à la levure, le complexe mSin3 pourrait être l'homologue fonctionnel du complexe Rpd3L-Expanded décrit au chapitre 4.

Ce complexe partage aussi une autre fonction importante homologue à Rpd3L : le contrôle du cycle cellulaire via le mécanisme Rb/ E2F, lui aussi conservé de la levure à l'humain. La protéine Rb, qui a pour homologue fonctionnel Whi5 chez la levure, interagit avec les facteurs de transcription E2F (SBF et MBF chez la levure) pour en inhiber l'activité et bloquer la cellule en phase G1 [191]. Cette répression nécessite le recrutement des complexes mSin3 chez l'humain et de Rpd3L chez la levure [98, 192]. La phosphorylation de Rb par le complexe cycline D / Cdk4/6 (Cln3/ Cdk1 chez la levure) force le transfert de Rb au cytoplasme, laissant ainsi E2F activer ses gènes cibles pour permettre le passage à la phase S [98]. L'implication de ces HDACs dans ces mécanismes essentiels conservés de la levure à l'humain est un autre argument de poids pour justifier l'étude détaillée des HDACs chez la levure.

Le complexe NuRD est composé du même cœur que les complexes mSin3 (HDAC1, HDAC2, RbAP46 et RbAP48) auquel s'ajoute les protéines MTA (*Metastasis-Associated Protein*), Mi-2, p66 et MBD3 (*Methyl CpG-Binding Domain 3*). Ces protéines, particulièrement MBD3, ciblent ce complexe aux îlots CpG méthylés en interagissant avec MeCP1 et MeCP2 (*Methyl CpG-binding Protein 1/2*) [187]. Ces îlots d'ADN méthylé sont

une marque épigénétique de répression de la transcription, auquel mécanisme NuRD contribue en déacétylant les nucléosomes leur étant associés. La protéine MTA, quant à elle, est essentielle à l'activité HDAC de NuRD. Fait intéressant, MTA, Mi-2 et p66 se déclinent en plusieurs isoformes pouvant former une panoplie de variantes du complexe NuRD [189]. La fonction de ces variantes n'est toutefois pas claire.

Les HDAC1 et 2 font aussi partie du complexe coREST, retrouvé principalement dans le système nerveux où il contrôle l'expression de plusieurs gènes [189]. Il n'est présent que chez les métazoaires, ce qui est compatible avec son activité neuronale. Fait intéressant, ce complexe comprend la première histone déméthylase (HDM) à être caractérisée, LSD1 [193]. Cette protéine déméthyle spécifiquement le résidu K4 de l'histone H3. Cette déméthylation stimule l'activité HDAC de coREST, qui en retour stimule l'activité HDM de LSD1 [194]. Ce mécanisme est à l'opposé du mécanisme de recrutement des complexes HDAC1/2 mSin3, pour qui l'activité HDAC dépend de la liaison d'ING2 aux résidus H3K4 méthylés. Cette dichotomie est probablement due aux fonctions spécifiques de coREST dans le contexte neuronal.

Des études récentes chez l'humain et la drosophile ont mis à jour un complexe homologue à Rpd3S chez la levure. Ce complexe, nommé SHMP chez l'humain [195, 196] ou RLAF chez la mouche [197, 198], est composé de protéines homologues à celles de Rpd3S : HDAC1 et 2 (Rpd3), mSin3B (Sin3), TLE (Ume1), Mrg15 (Eaf3) et Pfl (Rco1) [195, 196]. De plus, Mrg15 fait aussi partie du complexe Tip60, l'homologue de NuA4, ce qui renforce son homologie fonctionnelle avec Eaf3 [195], et son chromodomaine serait aussi responsable de l'interaction de SHMP avec les histones méthylées en H3K36 et H3K4 [195]. Le rôle de SHMP semble aussi très similaire à celui de Rpd3S : en effet, ce complexe est présent aux régions codantes activement transcrites et la répression de l'expression de ses sous-unités par ARN interférant cause une activité transcriptionnelle accrue [196]. Contrairement à la levure où Rpd3S est présent sur toute la longueur des régions transcrites, SHMP est concentré en un point environ 1 kb en aval du site d'initiation

de la transcription [196]. Certaines pistes portent aussi à croire que SHMP et RLAF seraient impliqués dans la déméthylation de H3K4 via une interaction avec l'histone déméthylase LID [195, 197, 198]. Cette fonction n'a pas été encore détectée pour Rpd3S, mais étant donné le degré d'homologie fonctionnelle considérable entre SHDM, RLAF et Rpd3S il ne serait pas surprenant de la retrouver aussi chez la levure.

1.3.1.2 HDAC3

Cette protéine a été identifiée en 1997 par similarité de séquence avec HDAC1 et HDAC2. On retrouve des homologues de HDAC3 de la mouche à l'humain, mais pas chez la levure [189]. Elle est aussi l'une des HDACs qui sont présentes à l'extérieur du noyau : elle possède en effet des signaux d'exportation et d'importation nucléaire, ce qui lui confère des fonctions potentielles dans le cytoplasme et dans le noyau. Ses fonctions nucléaires sont toutefois beaucoup mieux caractérisées [199].

La plupart des molécules de HDAC3 font partie de deux complexes corépresseurs : N-CoR (*Nuclear receptor Co-Repressor*) et SMRT (*Silencing Mediator of Retinoic and Thyroid receptors*). L'activité HDAC de ces complexes joue un rôle important dans les mécanismes de différenciation et de prolifération cellulaire et dans le processus d'apoptose. En effet, la déacétylation réduit l'affinité de coactivateurs ayant des bromodomains pour leurs cibles; ils influent ainsi sur l'activité de plusieurs récepteurs nucléaires [199].

La fonction de corépresseur de HDAC3 est principalement régulée par son exclusion du noyau, stimulée entre autres par l'interleukine IL-1 β . De plus, les caspases activées par le processus d'apoptose peuvent couper le signal de localisation nucléaire de HDAC3 pour la séquestrer dans le cytoplasme tout en conservant son activité HDAC [200]. L'activité catalytique de HDAC3 peut être stimulée par sa phosphorylation en sérine 424 par les kinases CK2 et DNA-PK [201, 202]. Sa déphosphorylation est catalysée par la sérine/thréonine phosphatase 4, qui incidemment fait partie du complexe N-CoR [201].

Curieusement, HDAC3, lorsqu'associé à la membrane cytoplasmique, est aussi une cible de la kinase Src, ce qui suggère un rôle pour HDAC3 dans la transduction de signaux émanant de cette membrane [203]. Ceci dit, l'activité catalytique de HDAC3 dans le cytoplasme n'est pas bien caractérisée.

1.3.1.3 HDAC8

Cette protéine a été découverte en 2000 en cherchant dans une banque d'EST des séquences similaires à HDAC1/2/3. Contrairement aux autres HDACs de classe I, celle-ci n'a toujours pas été associée à un complexe protéique. Quoique ses cibles soient encore inconnues, on sait qu'elle influe sur la régulation de certains mécanismes développementaux [204]. Son expression est spécifique à certains tissus et elle est présente aussi bien dans le noyau que le cytoplasme de plusieurs tissus normaux et lignées cancéreuses [204].

1.3.2 La classe II

La classe II regroupe les HDACs ayant une certaine homologie de séquence avec le domaine HDAC de Hda1 chez la levure. Des six protéines appartenant à cette classe, quatre possèdent une structure bipartite : un domaine HDAC C-terminal et un domaine N-terminal comprenant plusieurs sites d'interaction protéiques conservés [205]. Ces quatre HDACs, HDAC4, 5, 7 et 9, forment la sous-classe IIa. Les deux autres HDACs, HDAC6 et 10, ne possèdent pas ce domaine d'interaction et forment la sous-classe IIb [189]. Comme les fonctions des différentes HDACs de sous-classe IIa sont très similaires, je les traiterai d'un seul bloc. De la classe IIb, seule HDAC6 est relativement bien caractérisée et sera passée en revue.

1.3.2.1 *La sous-classe IIa*

Cette famille, qui comprend les HDAC4, 5, 7 et 9, est relativement homogène en structure et fonction. Bien qu'elles possèdent un domaine HDAC similaire à celui de Hda1 chez la levure, la ressemblance s'arrête là. Les membres de la sous-classe IIa possèdent un domaine N-terminal truffé de sites d'interactions protéiques leur conférant leurs fonctions biologiques. On y retrouve entre autres des signaux d'importation et d'exportation nucléaires et des sites d'interaction avec les protéines de la famille 14-3-3, pour la famille de facteurs de transcription MEF2 (*Myocyte Enhancer Factor 2*), ainsi que pour la protéine HP1 [205].

On a surtout étudié l'impact fonctionnel des interactions entre ces HDACs et les facteurs MEF2 et 14-3-3. Ces protéines apportent un mécanisme intéressant pour la régulation de l'activité corépressive de cette classe. La liaison de protéines 14-3-3 au domaine N-terminal de ces HDACs, stimulé par la phosphorylation de résidus sérine proche des sites de liaison 14-3-3, force leur expulsion du noyau, les empêchant ainsi de remplir leur rôle. La liaison de la protéine 14-3-3 protège aussi ce site phosphorylé contre les phosphatases PP1 β et PP2A, ce qui renforce la stabilité du complexe HDAC/14-3-3 [189]. Plusieurs kinases sont responsables de cet état : les kinases dépendantes du calcium ou de la calmoduline (CaMKs), la kinase de protéines D, les kinases de régulation de l'affinité des microtubules, les kinases induites par le sel et la kinase CTK1, impliquée dans le contrôle du cycle cellulaire [205]. On peut donc voir que ces HDACs sont impliquées dans la régulation d'une multitude de mécanismes différents.

Les facteurs de transcription de la famille MEF2, présents dans plusieurs tissus, repêchent quant à eux ces HDACs dans le cytoplasme pour les importer dans le noyau de manière à ce qu'elles répriment la transcription des gènes ciblés par MEF2. Il est intéressant de noter que cette répression se fait conjointement avec le complexe N-CoR via l'activité catalytique de HDAC3 [205].

La protéine HP1, associée à l'hétérochromatine, est aussi une partenaire importante de cette sous-classe de HDACs. Cette protéine se lie spécifiquement aux résidus H3K9 méthylés par la KMT Suv39H1, une marque des régions chromatiniennes silencieuses [205]. La fonction précise des HDACs recrutées à ces sites est inconnue, mais on sait que la déacétylation de ce résidu lysine est requise pour la méthylation par Suv39H1. On peut donc présumer que les HDACs de classe IIa jouent un rôle prépondérant dans ce mécanisme.

1.3.2.2 *La sous-classe IIb*

Cette sous-classe est composée des HDACs 6 et 10, qui n'ont pas d'homologues fonctionnels chez la levure. La HDAC10, bien qu'identifiée en 2002, a été très peu étudiée. On sait par contre qu'elle est impliquée dans la répression de la mélanogenèse [206] et que son inhibition perturbe l'état redox des cellules de cancer gastrique [207].

HDAC6 est la seule HDAC qui possède deux domaines catalytiques fonctionnels et un doigt de zinc capable de lier l'ubiquitine. Bien qu'elle soit capable de déacétyler des histones *in vitro*, ce qui lui confère son nom, elle est presque exclusivement cytoplasmique et n'a aucun rôle de régulation d'expression génique dans le noyau. Elle a trois cibles principales : la tubuline, la protéine Hsp90 (*Heat Shock Protein 90*) et la cortactine, ce qui lui confère des rôles dans la migration cellulaire (via la tubuline et la cortactine) et le mécanisme de dégradation des protéines aux conformations aberrantes (via Hsp90) [208].

L'activité de HDAC6 sur les microtubules a un impact sur la migration cellulaire, mais aussi sur la formation de synapses immuns lors de la réponse immunitaire. Toutefois, l'impact de son activité catalytique dans ces mécanismes est encore flou et semble dépendre du type cellulaire en question [208].

1.3.3 La classe III

Cette classe, comme chez la levure, regroupe les HDACs ayant recours à une réaction impliquant le NAD^+ pour déacétyler ses cibles. Elle comprend sept membres chez l'humain, nommés SIRT1 à SIRT7, ayant des degrés variables d'homologie avec Sir2. Seulement trois de ces protéines se retrouvent surtout dans le noyau et ont une fonction HDAC (SIRT1, 6 et 7). Du reste, trois sont mitochondriales (SIRT3, 4 et 5) et la dernière, SIRT2, est principalement localisée dans le cytoplasme.

Il y a un intérêt considérable pour ces protéines, particulièrement SIRT1, depuis la publication de travaux chez la levure faisant état d'un lien entre le vieillissement et les sirtuines. Bien que ces conclusions soient maintenant controversées, une multitude de groupes poursuivant cette piste ont néanmoins identifié des liens importants entre les sirtuines et plusieurs fonctions biologiques importantes telles que la régulation du métabolisme énergétique et le contrôle du cycle cellulaire.

Je passerai brièvement en revue les connaissances accumulées sur les sirtuines nucléaires SIRT1, 6 et 7 ainsi que sur SIRT2.

1.3.3.1 *SIRT1*

Cette sirtuine est la première à avoir été découverte chez l'humain [209, 210] et la mieux caractérisée. Elle est impliquée dans certains mécanismes de réparation de l'ADN, le maintien du rythme circadien et la régulation du métabolisme énergétique de plusieurs tissus.

Le rythme circadien est maintenu par l'expression cyclique d'une quinzaine de gènes qui régulent ensuite l'expression d'environ 10 % des gènes exprimés dans divers tissus [211]. SIRT1 est d'abord requise pour la régulation circadienne de la transcription de plusieurs de ces régulateurs. Elle interagit aussi avec le complexe HAT CLOCK/BMAL1,

un facteur de transcription influençant positivement l'expression de la plupart des gènes ayant une expression circadienne, pour déacétyler BMAL1 ainsi que les promoteurs où ce complexe est lié et en réprimer l'activité. De plus, SIRT1 déacétyle PER2, un autre régulateur circadien, ce qui mène à sa dégradation [212]. L'inhibition de SIRT1 mène à des problèmes de régulation des gènes circadiens *in vitro* et *in vivo* [211].

La réparation des lésions simple- et double-brin de l'ADN est partiellement régulée par l'activité de SIRT1. Les cellules où SIRT1 est absente présentent une sensibilité accrue aux agents génotoxiques pouvant causer ces bris ainsi qu'une formation réduite de foci γ H2AX. SIRT1 est recruté aux sites de cassures double-brins pour en favoriser la réparation par recombinaison homologue via son interaction avec la protéine NBS1, qui fait partie du complexe MRN responsable de la détection des dommages à l'ADN [212]. Cette sirtuine est aussi impliquée dans l'activation de la protéine responsable de la réparation par excision de base, APE1.

L'étude du rôle de SIRT1 dans la régulation du métabolisme énergétique découle en grande partie des études chez la levure démontrant l'impact de Sir2 sur la longévité. On sait maintenant que cet effet est très indirect et surtout dû à la fonction de contrôle de ces mécanismes métaboliques par ces sirtuines. SIRT1 joue un rôle prépondérant dans le foie à plusieurs niveaux. Premièrement, elle régule le niveau de glucose en activant la gluconéogenèse et en réprimant la glycolyse en déacétylant divers facteurs impliqués dans leur régulation (PGC-1 α , FOXO1 et CRTC2). De plus, son activité sur PGC-1 α et PPAR α promeut l'oxydation des acides gras dans les situations de jeûne [213]. Finalement, la déacétylation du récepteur nucléaire LXR affecte la régulation du niveau de cholestérol et de l'homéostasie lipidique [214].

Le rôle de régulation énergétique de SIRT1 s'étend aussi aux muscles squelettiques, au tissu adipeux blanc et au pancréas. SIRT1 active l'oxydation des acides gras par les mitochondries en réponse aux situations de jeûne via la déacétylation de PGC-1 α [215]. Le

tissu adipeux blanc, quant à lui, sera porté à relâcher plus d'acides gras dans ces situations grâce à la répression de PPAR γ par SIRT1 [216]. SIRT1 stimule aussi la production d'insuline dans le pancréas en réponse à un taux de glucose sanguin élevé [213].

Il est assez remarquable que la majorité de la fonction régulatrice de SIRT1 passe par la déacétylation de protéines effectrices et non pas d'histones. Ce mécanisme tranche nettement avec ses homologues chez la levure où l'activité HDAC est le mode d'action principal. Il est toutefois possible que l'on se retrouve simplement devant un biais observationnel où les cibles non nucléosomales des sirtuines chez la levure ne seraient tout simplement pas encore étudiées et décrites adéquatement.

1.3.3.2 *SIRT2*

Cette protéine est la seule sirtuine à n'être présente presque exclusivement que dans le cytoplasme, où elle interagit avec les microtubules pour déacétyler l' α -tubuline de concert avec HDAC6 [212], ce qui réduit la stabilité des microtubules. Il est intéressant de noter que la sur- ou sous-expression de SIRT2 est impliquée dans plusieurs maladies neurodégénératives de par son impact sur le cytosquelette [212].

SIRT2 joue aussi un rôle prépondérant dans le cycle cellulaire, particulièrement au point de contrôle G2/M. Sa fonction exacte est encore mal caractérisée, mais on la retrouve dans le noyau lors de la mitose, où l'on croit qu'elle régule l'expression de certains gènes [217]. De plus, elle peut bloquer la condensation des chromosomes lorsque le réseau de microtubules est perturbé.

Finalement, SIRT2 déacétyle aussi quelques facteurs de transcription pour en réguler l'activité. Par exemple, la déacétylation de FOXO1 en cause l'importation dans le noyau, lui permettant ainsi de réprimer PPAR γ , ce qui résulte en l'inhibition de la différenciation des adipocytes [218].

L'activité de SIRT2 est, comme pour la plupart des HDACs humaines, hautement régulée. Sa phosphorylation par CDK1 est requise pour sa fonction de point de contrôle G2/M [217], tandis que la phosphorylation par les complexes cycline-kinase cycline E/Cdk2, cycline A/Cdk2 et p35/Cdk5 réduisent son activité catalytique [219]. Ironiquement, la fonction de SIRT2 peut aussi être inhibée par son acétylation par la HAT p300 [220].

1.3.3.3 SIRT6

SIRT6, comme ses homologues fonctionnels chez la levure, à un rôle important dans le maintien des régions hétérochromatiniennes. Elle déacétyle en effet le résidu H3K9 pour en permettre la méthylation par Suv39H1 [221]. Curieusement, c'est aussi une fonction présumée des HDACs de classe IIa, décrites ci-dessus. La déplétion de SIRT6 entraîne entre autres des fusions chromosomiques et divers problèmes aux télomères [221].

Elle possède aussi une spécificité pour le résidu H3K56, comme Hst3/4 chez la levure [222, 223]. Cette fonction est probablement responsable des phénotypes d'instabilité génomique et de réparation de lésion à l'ADN déficiente observés lorsque SIRT6 est déplété [223].

1.3.3.4 SIRT7

Cette sirtuine n'est pas bien caractérisée. On sait toutefois qu'elle est surtout localisée au nucléole et est impliquée dans le contrôle de la transcription de l'ADN ribosomal par l'ARN polymérase I en fin de mitose [224].

Cette sirtuine gagne à être mieux étudiée car sa déplétion réduit de moitié la durée de vie chez souris et cause de l'hypertrophie cardiaque dégénérative. Les auteurs de l'étude

rapportant ces résultats croient que ces effets sont dus principalement à l'hyperacétylation de p53 et à une sensibilité accrue au stress oxydatif [225].

1.3.4 La classe IV

Cette classe ne contient qu'un membre, HDAC11. Elle présente des similarités de séquences avec les HDACs de classe I et II chez la levure, mais des études phylogéniques ont clairement démontré qu'elle fait classe à part [189, 226]. Bien que sa fonction est très mal caractérisée on sait qu'elle est impliquée dans la tolérance immunitaire en réprimant l'expression de l'interleukine 10 dans les cellules présentatrices d'antigènes [227] ainsi que dans le développement des oligodendrocytes [228].

Tableau 1 : Les HDACs humaines et leurs homologues chez la levure.

Classe	<i>H. sapiens</i>			Homologue
	HDAC	Localisation	Complexe	(<i>S. cerevisiae</i>)
I	HDAC1,	Noyau	mSin3,	Rpd3L
	HDAC2		NuRD,	N/A
			CoREST	N/A
			SHMP	Rpd3S
	HDAC3	Noyau	N-CoR /SMRT	N/A
	HDAC 8	Noyau / Cytoplasme	N/A	
IIa	HDAC4	Noyau / Cytoplasme		Hda1
	HDAC5		(domaine HDAC	
	HDAC7		seulement)	
	HDAC9			
IIb	HDAC6	Cytoplasme / Noyau		Hda1
	HDAC10	Noyau (?)		(domaine HDAC seulement)
III	SIRT1	Noyau / Cytoplasme		Sir2, Hst1
	SIRT2	Cytoplasme / Noyau		Hst2
	SIRT3	Mitochondrie		N/A
	SIRT4	Mitochondrie		N/A
	SIRT5	Mitochondrie		N/A
	SIRT6	Noyau		Sir2, Hst3/Hst4
	SIRT7	Nucléole		N/A
IV	HDAC11	Noyau (?)		N/A

1.4 Les histone lysine méthyltransférases (KMTs)

La méthylation est une marque covalente intéressante et polyvalente. En effet, elle n'altère pas la charge positive des résidus affectés, contrairement à l'acétylation, et existe sous trois formes (lysine mono-, di- et triméthylée) qui peuvent avoir des impacts fonctionnels différents selon la situation. De plus, l'acétylation et la méthylation sont mutuellement exclusives sur un résidu donné, ce qui peut être utilisé aux fins de signalisation par la cellule. Finalement, la méthylation est une marque covalente qui a une longue demi-vie, contrairement à l'acétylation, ce qui en fait une marque « mémoire » pour la cellule.

La levure compte trois KMTs : *SET1*, *SET2* et *DOT1*. Les deux premières, *SET1* et *SET2*, possèdent un domaine SET mais pas *DOT1*. Le domaine SET est nommé d'après trois protéines de la famille trithorax chez la Drosophile : Su(var), E(z) et Trithorax [229]. Il a été conservé au cours de l'évolution et est généralement associé aux histone méthyltransférases (HMTs) [230]. Toutefois, la seule présence d'un domaine SET n'est pas garante d'une fonction HMT; la levure possède dix protéines ayant un domaine SET dont seulement cinq ont une activité méthyltransférase détectable : *SET1*, *SET2*, *RKM1*, *RKM2*, *RKM3* et *RKM4*. De ces dernières, seulement Set1 et Set2 ciblent les histones; Rkm1-4 ciblent des protéines ribosomales [231-233].

Je passerai maintenant en revue les trois histone KMTs présentes chez la levure.

1.4.1.1 *SET1*

Le gène *SET1* a été identifié par homologie de séquence avec les gènes de la famille trithorax chez la Drosophile, particulièrement par la présence d'un domaine SET bien conservé. Un mutant *set1Δ*, bien que viable, montre des défauts d'inhibition épigénétique

de la transcription, de croissance, de sporulation ainsi qu'une morphologie aberrante [230]. Set1 fait partie d'un complexe nommé COMPASS (*Complex Proteins Associated with Set1*) composé de huit sous-unités [234-236] et capable de méthylater la lysine 4 de l'histone H3 [235-237].

Bien que la méthylation catalysée par COMPASS serait impliquée dans l'inhibition de la transcription aux télomères [238, 239], elle est surtout une marque importante de transcription active [240]. Le complexe COMPASS serait recruté par le complexe Paf1 au domaine C-terminal phosphorylé en sérine 5 de l'ARN polymérase II en phase d'élongation [131, 241, 242]. Le complexe Paf1 est aussi impliqué dans la monoubiquitination de la lysine 123 de l'histone H2B par Rad6 en recrutant le complexe Rad6/Bre1 [242-244]. Cette monoubiquitination est essentielle à la triméthylation de H3K4 par COMPASS et de H3K79 par Dot1 [242, 245, 246], mais n'est pas requis pour la monométhylation de H3K4 par COMPASS [247]. Par contre, l'activité de la kinase du domaine C-terminal de l'ARN polymérase II, Ctk1, est essentielle pour cette monométhylation [248].

Le rôle de la méthylation de H3K4 est resté éluif jusqu'à la découverte de protéines pouvant interagir avec ce résidu. Deux domaines protéiques connus sont capables d'interagir avec les lysines méthylées : les domaines CHD (chromodomaine) [249, 250] et PHD (*Plant HomeoDomain*) [251, 252]. La sous-unité des complexes SAGA et SLIK, Chd1, possède un domaine CHD et est capable de lier H3K4 triméthylé. Cette liaison est importante à l'activité de SLIK *in vivo* [253].

De plus, la moitié des protéines PHD de la levure peuvent se lier aux lysines 4 ou 36 méthylées [254]. Celles-ci sont toutes impliquées de près dans la régulation de la transcription. On y retrouve des sous-unités de complexes HDACs (Cti6, Pho23, Set3), HATs (Yng1, Yng2, Nto1) et COMPASS (Spp1).

Une étude récente a apporté de nouvelles observations confirmant le rôle de « mémoire transcriptionnelle » de la méthylation en H3K4. Dans cette étude il est démontré

que la fonction KMT de Set1 ainsi que le résidu H3K4 sont essentiels à la transmission des patrons d'expression géniques de la cellule-mère à la cellule-fille chez *Dictyostelium discoideum* [255]. Cette transmission est particulièrement importante, par exemple pour le maintien de l'expression de gènes impliqués dans la résistance aux stress environnementaux.

1.4.1.2 SET2

La protéine Set2 a été caractérisée en 2002 comme une KMT spécifique à la lysine 36 de l'histone H3 ayant une fonction répressive vis-à-vis la transcription [256]. L'année suivante quatre groupes ont démontré que Set2 interagit avec la forme hyperphosphorylée de l'ARN polymérase II et que cette interaction dépendait plus spécifiquement de la phosphorylation en sérine 2 du domaine C-terminal de Rpb1, catalysée par Ctk1 [132, 257-259]. Cette observation conférait à Set2 une fonction liée à la phase d'élongation de la transcription.

2005 fut une autre année faste pour l'élucidation de la fonction de Set2. En effet, trois groupes ont observé que le complexe Rpd3S était capable de lier des peptides ou des nucléosomes méthylés en H3K36 *in vitro* et que cette interaction était importante pour la déacétylation de gènes activement transcrits [117, 118, 122]. Cette interaction entre Rpd3S et le résidu H3K36 méthylé dépend du domaine CHD de la sous-unité Eaf3 de Rpd3S.

Le rôle répressif de Set2 sur la transcription se confirme encore par quelques observations. Premièrement, des délétions de Set2 ou de sous-unités de Rpd3S suppriment partiellement les problèmes d'élongation de mutants du complexe d'élongation Bur1/Bur2 [118]. De plus, l'activité de Set2 et la déacétylation catalysée par Rpd3S sur la région codante est essentielle à la répression de l'initiation de la transcription à partir de sites d'initiation cryptiques contenus dans l'ORF, tel qu'observée dans les mutants *spt6Δ* et du

complexe d'élongation FACT [117, 122]. Finalement, la méthylation catalysée par Set2 a une activité opposée à celle du complexe d'élongation Spt4/Spt5 (homologue de DSIF chez l'humain). Toutes ces données supportent un rôle répressif de Set2 et de Rpd3S sur la transcription. Ils agiraient en tandem, probablement pour restaurer la matrice chromatinienne à son état initial non permissif après le passage de l'ARN polymérase II.

L'activité KMT de Set2 est régulée par différents facteurs. L'isomérisation de la proline 38 de l'histone H3 par Fpr4 inhibe l'activité de Set2 *in vitro* et un mutant Fpr4 catalytiquement inactif démontre un niveau de H3K36 méthylé plus élevé *in vivo* [260]. Fait intéressant, la diméthylation de H3K36 ne requiert que le domaine SET de Set2, est indépendant de la phosphorylation de l'ARN polymérase II catalysée par Ctk1 et est suffisant pour la déacétylation et la répression de la transcription cryptique [261]. Par contre, la triméthylation de H3K36 requiert Spt6 et Ctk1. La présence d'Asf1 aux régions codantes peut aussi stimuler cette triméthylation [262].

1.4.1.3 *DOT1*

Le gène *DOT1* (*Disruptor Of Telomeric silencing 1*) a été identifié dans un crible visant à découvrir les gènes qui, lorsque surexprimés, pouvaient perturber l'inhibition épigénétique de la transcription aux télomères [263]. La surexpression de Dot1 affecte aussi l'inhibition aux loci HM et d'ADN ribosomal, ce qui explique l'initiation aberrante de la méiose dans un mutant *dot1Δ* [264].

Dot1 est la seule KMT ne comprenant pas de domaine SET chez la levure et cible la lysine 79 de l'histone H3 [265-269]. Cette triméthylation, comme celle de H3K4 catalysée par Set1, dépend de la monoubiquitination de la lysine 123 de l'histone H2B par Rad6 [131, 243, 244]. Cette monoubiquitination est liée à l'activité transcriptionnelle de l'ARN

polymérase II via le complexe Paf1, ce qui impliquerait Dot1 et la méthylation de H3K79 dans le contrôle de la transcription.

Le résidu H3K79 est situé dans le corps globuleux de l'histone sur les surfaces contactant l'ADN. La méthylation de H3K79 par Dot1 est importante pour restreindre l'étendue des régions transcriptionnellement silencieuses. Ces régions, riches en protéines Sir, sont hypométhylées en H3K79. La distribution des protéines Sir ainsi que l'identité des régions silencieuses changent dramatiquement lorsque Dot1 est délété ou le résidu H3K79 muté [267-270]. Toutefois, une étude récente postule que l'effet de Dot1 sur l'inhibition de la transcription aux régions silencieuses serait dû à une exclusion des protéines Sir des régions méthylées en H3K79 et non pas d'un effet direct de Dot1 : l'impact de Dot1 sur l'inhibition de la transcription dépendrait aussi de la méthylation en H3K4 par Set1 ainsi que de l'activité des HATs Gcn5, Elp3 et Sas2. De plus, le recrutement direct des protéines Sir supprime le phénotype d'un mutant *dot1Δ* [271].

Dot1 joue aussi un rôle important dans la réparation des bris de l'ADN. En effet, une délétion *dot1Δ* confère une sensibilité accrue aux rayons X [272] et empêche l'activation de Rad9, une protéine impliquée dans la réparation de ces bris et dans le contrôle du cycle cellulaire [273, 274]. Les mutants *dot1Δ* démontrent aussi une déficience des points de contrôle G1 et intra-S, mais conservent celui de G2/M [273].

Dot1 possède donc trois fonctions distinctes, mais possiblement reliées : le contrôle de la transcription via sa dépendance en H2B123 monoubiquitiné, le contrôle de l'étendue des régions silencieuses par l'exclusion des protéines Sir des zones méthylées en H3K79 et le contrôle du cycle cellulaire via l'activation de Rad9 pour la réparation des cassures double-brins.

1.5 Les histone déméthylases (HDMs)

On a longtemps cru que la méthylation des histones était une marque stable et irréversible. Toutefois, quelques études avaient déjà observé des fluctuations des niveaux d'histones méthylés, ce qui implique un contrôle dynamique de la méthylation [275, 276]. La découverte en 2004 chez l'humain de la première histone déméthylase, LSD1, confirme l'existence de mécanismes servant à réguler les niveaux d'histone méthylées. Cette protéine, retrouvée aussi bien chez l'humain que chez *S. pombe* (mais pas chez *S. cerevisiae*), est une amine oxydase utilisant le cofacteur FAD qui catalyse la déméthylation en générant une molécule de formaldéhyde. Elle a pour cibles les résidus lysine mono- et diméthylés 4 et 9 de l'histone H3 [193].

Deux HDMs contenant un domaine JmjC, JHDM1A et JHDM1B, ont été identifiés chez l'humain en 2006. JHDM1A utilise un atome de fer et une molécule d'alpha-kétoglutarate comme cofacteurs pour catalyser la déméthylation de la lysine H3K36 diméthylée mais pas triméthylée. Des protéines ayant un domaine JmjC sont présentes chez la plupart des eucaryotes, dont *S. cerevisiae* [277]. Les protéines appartenant à cette famille sont donc de bonnes candidates pour le rôle de HDM chez la levure, ce qui est confirmé en 2007 lorsqu'on en caractérise cinq, dont quatre possédant une activité HDM : Jhd1, Jhd2, Rph1 et Gis1 [278-281]. L'activité de Jhd1 et de Gis1 est spécifique au résidu H3K36 mono- et diméthylé [278, 281], celle de Jhd2 au résidu H3K4 triméthylé [280, 281] et finalement Rph1 déméthyle la lysine H3K36 triméthylée [279, 281].

Bien que l'activité biochimique des différentes HDM présentes chez la levure soit maintenant caractérisée, leur régulation, mode de recrutement et fonction biologique ne le sont presque pas. On sait toutefois que l'activité de Jhd2 est contrôlée par sa polyubiquitination par la ligase E3 Not4 et sa dégradation subséquente par le protéasome [282]. De plus, son activité déméthylase dilue le signal épigénétique de transcription active

transmis de cellule-mère à cellule-fille [283]. Il reste donc beaucoup de travail pour déchiffrer les mécanismes et fonction de cette nouvelle famille d'enzymes.

1.6 Hypothèses et objectifs

La structure de la chromatine est un obstacle de taille pour tous les mécanismes se rapportant à l'ADN. Comme on vient de le voir, la cellule a non seulement mis au point de multiples mécanismes visant à contourner cet obstacle mais l'exploite pour réguler l'activité des complexes protéiques devant avoir accès à l'ADN.

Les fonctions répressives des complexes HDAC Rpd3S et Set3C sur la transcription par l'ARN polymérase II ont été jusqu'ici assez mal caractérisées. Les principales conclusions sur les fonctions *in vivo* de ces complexes sont basées sur des observations indirectes de nature biochimiques, génétiques ou protéomiques. Il est d'ailleurs curieux que ces hypothèses n'aient pas été remises en question plus tôt, car il est très risqué d'extrapoler une fonction biologique catégorique à partir de données de cette nature.

Au cours de mes travaux de thèse je me suis efforcé de remédier à cette situation en caractérisant *in vivo* les fonctions de ces complexes. Pour y parvenir, j'ai principalement utilisé une technique de localisation génomique des protéines associées à la chromatine, le ChIP-chip. Cette technique, dont je discute en profondeur dans le chapitre 2 de cet ouvrage, permet d'identifier les sites d'interactions protéine-ADN ou encore la distribution d'histones possédant une marque spécifique (p. ex. méthylation, acétylation, etc.) à l'échelle génomique.

Les résultats que j'ai obtenus, présentés au chapitre 3, démontrent clairement les failles dans les données et conclusions publiées sur Rpd3S jusqu'ici. En effet, on avait observé que Rpd3S était capable de lier le résidu H3K36 méthylé par Set2 et que les mutants où cette interaction est troublée démontrent des phénotypes d'initiation cryptique de la transcription et d'acétylation nucléosomale aberrante. On en a déduit que cette interaction était essentielle au recrutement de Rpd3S sur les régions codantes activement transcrites. Je démontre toutefois que bien que l'activité catalytique de Rpd3S dépend de la méthylation en H3K36, comme l'indiquent les perturbations des niveaux d'acétylation aux

ORFs, son recrutement, contrairement à ce que l'on pensait jusqu'ici, ne requiert pas cette marque de méthylation. Je me suis alors mis à la recherche du facteur responsable du recrutement de Rpd3S et ce faisant j'ai découvert un mécanisme très intéressant : en effet, le recrutement de ce complexe aux régions codantes dépend plutôt de l'état de phosphorylation du domaine C-terminal de Rpb1, plus spécifiquement de la double phosphorylation en sérine 2 et sérine 5 par les kinases Ctk1 et Kin28, respectivement. J'ai aussi découvert que le complexe Spt4/5, l'homologue chez la levure du facteur d'élongation DSIF, influait négativement sur l'association Rpd3S-Rpb1 phosphorylé, mais que la phosphorylation de sa sous-unité Spt5 par le complexe Bur1/2 réprimait cette activité négative. Ce mécanisme est beaucoup plus complexe que ce que le modèle courant proposait et implique plusieurs niveaux de régulation à l'activité de Rpd3S.

La composition du complexe Set3C fait consensus depuis plusieurs années déjà. On sait qu'on y retrouve deux HDACs, Hst1 et Hos2, ainsi que plusieurs autres protéines toutes importantes à sa fonction. Des expériences génétiques et de caractérisation d'interactions protéiques ont démontré que toutes les sous-unités de Set3C, à l'exception de Hst1, étaient essentielles à sa fonction. On en a conclu que Hst1 n'était probablement pas un membre stable du complexe, mais que le reste agissait comme un tout. Les données publiées récemment par le groupe de Kim *et coll.* vont dans le même sens et ils ont proposé un modèle de recrutement impliquant la liaison du domaine PHD de Set3 aux résidus H3K4 diméthylés. Je me suis vite intéressé à ce complexe dont le mode de recrutement semble être étrangement similaire à celui postulé pour Rpd3S. De plus, une étude protéomique publiée en 2008 laisse croire que Rpd3L et Set3C formeraient en fait un seul supercomplexe nommé « Rpd3L-Expanded ». J'ai confirmé que la méthylation en H3K4 est en effet requise pour le recrutement de Set3C à la plupart de ses cibles et que Hst1 liait principalement les promoteurs de gènes où son partenaire Sum1 est aussi lié indépendamment de Set3C. J'ai toutefois découvert que le recrutement de Set3C sur les régions codantes les plus transcrites ne dépend pas de la méthylation en H3K4 et que Hst1,

qui ne lie habituellement que les promoteurs, était présent sur ces mêmes régions. De plus, Hos2 n'est présent que sur les régions codantes des gènes où Hst1 et Set3 étaient aussi présents, mais pas aux gènes où la liaison de Set3 dépend de la méthylation en H3K4. Ces résultats étonnants sont détaillés au chapitre 4.

2 Le ChIP-Chip : technique et contrôles

L'étude de la chromatine et des facteurs interagissant avec l'ADN a été grandement facilitée par la mise au point de l'immunoprécipitation de la chromatine, ou ChIP. Brièvement, cette technique consiste à créer des liens covalents entre l'ADN et les protéines se trouvant à proximité à l'aide d'agents chimiques ou physiques. L'immunoprécipitation de ces protéines permet ainsi de coprécipiter et d'analyser les fragments d'ADN auxquels ils sont liés. Cette technique ainsi que plusieurs de ses applications relatives à l'étude de l'environnement chromatinien ont été très bien décrites par Kuo et Allis en 1999 [284].

Ce n'est qu'en 2000 que ce champ d'études prend véritablement son envol avec l'élaboration par deux groupes d'une technique couplant le ChIP aux biopuces (« ChIP-chip ») [285, 286]. Le ChIP-chip, tout comme son ancêtre le ChIP, permet d'identifier précisément les sites d'interaction entre une protéine d'intérêt et l'ADN. Cependant, là où le ChIP classique ne permettait l'étude que d'un nombre restreint de loci d'interactions potentielles devant être choisies au préalable, le ChIP-chip l'étend à l'échelle génomique en amplifiant tous les fragments d'ADN coprécipités et en sondant ces derniers à l'aide de biopuces où le génome entier est représenté de manière impartiale.

Le ChIP-chip est maintenant une méthode robuste largement utilisée dans l'étude des mécanismes affectant la chromatine. Malheureusement, peu d'efforts ont été déployés pour l'élaboration de contrôles adéquats, et ce malgré le caractère essentiel de ceux-ci pour toute expérience. En effet, le ChIP-chip comporte trois sources principales de bruit expérimental pouvant générer des artefacts d'enrichissement : la variation biologique, la variation technique et la spécificité imparfaite de tout anticorps utilisé. Les variations techniques et biologiques sont minimisées simplement par la réplication des expériences. Par contre, il est assez difficile de compenser pour le signal non spécifique dû à l'anticorps

utilisé. Comme le ChIP-chip est la technique principale utilisée pour mes travaux de thèse, l'élaboration de contrôles robustes et appropriés pour corriger ces artefacts fut ma priorité.

Dans ce chapitre, je présenterai d'abord mon protocole de ChIP-chip optimisé, suivi d'une discussion sur les méthodes de contrôle du bruit expérimental recommandées.

2.1 Manuscrit :

« Genome-wide Location Analysis of Chromatin-associated Proteins by ChIP on CHIP: Controls Matter »

Simon Drouin and François Robert

Contribution des coauteurs :

S. Drouin: Design, exécution et analyse expérimentale. Rédaction conjointe avec F. Robert.

F. Robert: Design expérimental, rédaction conjointe avec S. Drouin.

2.1.1 Abstract

ChIP on CHIP is becoming more widely used to determine the genome-wide location of chromatin-associated protein such as transcription factors and chromatin modifying enzymes. Many versions of this technique can now be found in the literature. As for any microarray-based experiment, the design of the experiment has a great impact on the final result. In ChIP on CHIP experiments, the control for hybridization can either be the input DNA from the immunoprecipitation or a control immunoprecipitation. Here we provide a comparison of these two different experimental designs and provide a detailed protocol for the latter, which we found to be the optimal method.

2.1.2 Key words

Chromatin Immunoprecipitation; Location Analysis; ChIP on CHIP; DNA microarrays; Chromatin.

2.1.3 Introduction

Protein-DNA interactions are key events in the control of cellular processes such as DNA replication, repair, recombination, transcription and others. It is therefore no surprise that many efforts have been put into the development of techniques that allows for the identification and quantization of these interactions both *in vitro* and *in vivo*. The immunoprecipitation of cross-linked chromatin fragments (or Chromatin Immunoprecipitation, or “ChIP”) is now well accepted as one of the best methods for the study of protein-DNA interactions *in vivo* [284]. In the standard ChIP method, sequence-specific PCR amplifications are used to monitor the enrichment of specific loci after the precipitation of the chromatin fragments. Despite all its advantages, this strategy is limited

by the fact that one needs to know (or expect) what particular loci are occupied by the protein to be assayed. To circumvent this limitation and get a reading of the genome-wide association of a given protein to genomic DNA, we and others have coupled ChIP to hybridization to DNA microarrays, a process often referred to as “ChIP on CHIP” [285, 286]. Many versions of this technique have been developed since.

We found that the choice of the control used for the hybridization does have a great influence on the quality of the results. Here we provide a complete protocol for what we find is the optimal protocol for ChIP on CHIP in yeast, and provide a quick comparison of the different hybridization controls that can be used.

2.1.4 Experimental design

Many variables exist between the protocols used by the different groups doing ChIP on CHIP. If most of these differences probably do not have much impact on the end result, we found that the choice of the control used for the hybridization can greatly influence the quality of the results. Two different approaches have been used. In the first approach, originally developed by Ren et al.[285], the immunoprecipitated DNA is hybridized along with DNA made from the non-immunoprecipitated extract (input DNA). This method has the advantage of giving a reading for the actual enrichment generated by the immunoprecipitation. On a more practical level, it has the advantage that only one ChIP needs to be performed for every array to be hybridized (the input serving as the control). The disadvantage of this method is that any protein that would cross-reacts with the antibody used is susceptible to generate DNA enrichment, resulting in false positive targets.

The second approach, originally developed by Iyer et al. [286] consists of using a control immunoprecipitation as the control for hybridization. Because ChIP is often performed on epitope-tagged proteins in yeast, an immunoprecipitation performed with the

same antibody but on the parental strain containing no tagged protein is used as the control. Because the same antibody is used in both channels, the enrichment it generates by cross-reactivity is cancelled out. Only the differential enrichment generated by the tagged strain versus the non-tagged strain remains. However, since this design involves the use of two different strains used to prepare two extracts, two immunoprecipitations and so on, there are many steps where noise could be introduced.

Since both approaches are susceptible to different sources of noise, we performed a side-by-side comparison and found that using a control immunoprecipitation from a non-tagged strain generate better results (less noise) than using input DNA as control. Our optimized protocol is described below and the last section provides an example of the comparison between the two designs.

2.1.5 Growing and cross-linking cells

The power of ChIP comes in part from the fact that formaldehyde can be applied directly to living cells and acts very quickly. Virtually any growth condition can therefore be used depending on experimental needs. We typically use 50 mL of exponentially growing cells (i.e. OD600 0.6-1.0), which represents around $3 \times 10^8 - 5 \times 10^8$ cells. We had success doing ChIP with fewer cells, but this parameter was never investigated thoroughly since yeast cells are easy to obtain in large amounts. The cross-linking time is one of the first factors that should be optimized if initial attempts to ChIP a protein fail. It has been reported that excessive exposure to formaldehyde can make the subsequent sonication steps more difficult. Strains need to be grown in tandem; a culture from an isogenic non-tagged strain for every culture of tagged strain. Growing cells in triplicate is also recommended.

2.1.5.1 Material required

- 50 mL of fresh yeast culture (tagged strain) in appropriate conditions
- 50 mL of fresh yeast culture (non-tagged strain) in appropriate conditions
- Formaldehyde
- 2.5 M Glycine (facultative)
- ice-cold TBS

2.1.5.2 Cross-linking

1. For each strain, grow 50 mL of yeast in appropriate liquid medium until OD₆₀₀ reaches 0.6-1.0.
2. Add 1.4 mL 37% Formaldehyde directly to the flask or transfer the content of the flask into a 50 mL conical tube containing 1.4 mL formaldehyde.
3. Incubate 30 minutes at room temperature with agitation. The incubation time may need to be optimized depending on the protein to be immunoprecipitated.
4. Add 2.5 mL of 2.5 M glycine to quench the formaldehyde (facultative). We usually omit this step since the subsequent washing step quickly removes the excess of formaldehyde anyway. However, if all the samples are not to be immediately processed (when doing a time course, for example), glycine should be used to stop the cross-linking reaction and therefore ensure that all samples are treated alike.
5. Centrifuge into 50 mL conical tubes for 5 minutes at 3,600 rpm in a refrigerated (4 °C) table top centrifuge.
6. Pour off the supernatant and wash the cells by resuspending into 40 mL of ice-cold TBS (vigorous shaking is required) and centrifuge again as in step 5.
7. Repeat step 6 for a total of 2 washes.
8. Using the remaining liquid, resuspend the cell pellet and transfer into a 1.5 mL screw cap tube.
9. Centrifuge 1 minute at maximum speed, remove the supernatant with a pipette and snap freeze the cell pellet by dropping the tube in liquid nitrogen.
10. Samples can either be used directly or stored at -80°C.

2.1.6 Breaking cells

Yeast cells are notoriously difficult to break due to their cell wall. Various cell breaking methods have been developed, each optimized for a specific purpose. The method we have elected consists of shaking the cells in the presence of glass beads. We do not recommend using regular vortexers since they tend to break rapidly due to overuse. We use a bead beater from Biospec Products sold as an 8 tube- or 48-tube format.

2.1.6.1 Material required

- Lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 0.1 % Na-deoxycholate, 1 mM PMSF, 1 mM Benzamidine, 10 ug/mL Aprotinin, 1 ug/mL Leupeptin, 1 ug/mL Pepstatin)
- Acid washed glass beads (400-600um) (Sigma)
- Bead beater

2.1.6.2 Cell lysis

1. Thaw cell pellet on ice.
2. Resuspend in 700 uL Lysis buffer.
3. Add the equivalent of a 0.5 mL PCR tube of glass beads.
4. Screw cap tightly and secure tubes to the bead beater.
5. Beat for 5 minutes and put into an ice-water bath for 5 minutes.
6. Repeat step 5 three more times. The amount of beating that is required may vary depending on the equipment, the yeast strain, the age of the culture and the sample/beads ratio used. The first thing that should be done before doing any ChIP experiment is to optimize the cell breaking step. Cell breakage can be easily monitored by looking at cells under a light microscope.

7. Punch a hole at the bottom using an 18G1 ½ needle (do not go all the way through. Stop as soon as liquid or air is going through) and set up over a 2 mL tube.
8. Loosen the cap and spin 3-4 seconds. This should allow the material to transfer into the 2 mL tube while the beads stay trapped into the top tube. This operation may not work with all types of centrifuge. We use an Eppendorf 5417C with a 30 slot rotor.
9. Resuspend the chromatin pellet using its own supernatant and transfer into a new 1.5 mL tube. Alternatively, the supernatant can be discarded and the chromatin pellet resuspended into 700 uL of fresh Lysis buffer. For proteins with a prominent soluble form (not associated with chromatin), this may prevent unbound proteins from competing for antibody during the subsequent immunoprecipitation step.

2.1.7 Preparation of soluble chromatin fragments

Just like for cell lysis, there is a lot of variation in the efficiency of the sonication due to the type of sonicator used, but also the type and age of the probe. This step therefore needs to be optimized, which can be done by looking at the size of the sheared DNA on an agarose gel. If formaldehyde-treated cells are used to optimize sonication (which is recommended since the crosslinking may have an influence on the shearing efficiency) the crosslinking should be reversed before the DNA is purified and loaded on agarose gel. The protocol described here was optimized for a Branson Sonifier 250 equipped with a micro probe.

2.1.7.1 Material required

- Sonicator (Branson Sonifier 250 or your favorite type)
- Refrigerated 1.5 mL tubes centrifuge

2.1.7.2 Sonication

1. Using the “Hold” and “Constant Power” settings, sonicate the sample 4 times for 20 seconds at power 4. Put the tube on ice for at least a minute between every round of sonication.
 - i. Note : When doing multiple samples, sonicate each sample once and go back to the first one. Rinse the probe with distilled water between every sample in order to avoid cross-contamination. Avoid contact between the probe and the side of the tube. This would reduce the efficiency of the sonication.
2. Centrifuge 5 minutes at maximum speed and transfer the supernant (which contains solubilized chromatin fragments) to a new 1.5 mL tube. From now on, this will be referred to as the whole cell extract (WCE).

2.1.8 Immunoprecipitation

Immunoprecipitation can be done using a great variety of systems. Both monoclonal and polyclonal antibodies have been shown to work for ChIP. All sorts of matrices such as agarose, sepharose or magnetic beads can be used. The protocol described here uses magnetic beads coupled to an anti-myc antibody. The test strain expresses a myc-tagged version of the protein to be IP'd and the control strain is isogenic to the test strain but has no tagged gene. Because of the ease at which proteins can be tagged and expressed from their chromosomal loci in yeast, the use of epitope-tagged strains is much more convenient and cost effective than the use of a battery of antibodies.

2.1.8.1 Material required

- Dynabeads Pan Mouse IgG (Invitrogen)
- Anti-Myc antibody (9E10 from SantaCruz Biotech.)
- Hematology Chemistry Mixer 346 (Fisher), or equivalent

- Magnet system MPC-S (Invitrogen)
- Lysis buffer (Same as for breaking the cells)
- PBS+BSA (PBS supplemented with 0.5% Bovine Serum Albumin)
- Wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 0.5 % NP40, 0.5 % Na-deoxycholate, 1 mM EDTA)
- TE (10 mM Tris pH 8.0, 1 mM EDTA)

2.1.8.2 Immunoprecipitation

1. Set up a tube with 5 uL of WCE and store at -20°C. This will not be used for the microarray, but is useful for sequence-specific PCR you may want to do to validate your ChIP.
2. Set up a tube with 500 uL of WCE and 30 uL of anti-Myc-coupled-magnetic beads.
 - i. Note : The beads should be prepared the day prior use the following way. Take enough beads to get 50 uL per immunoprecipitation. Put into a conical plastic tube, centrifuge 3 minutes at low speed to pellet the beads. Wash twice with PBS+BSA and incubate overnight at 4°C with agitation in the presence of anti-Myc antibody (the quantity of antibody to use should be determined empirically) and enough PBS+BSA so ensure a good agitation. The next day (prior to use) wash twice in PBS/BSA to remove non-bound antibodies and resuspend beads in 30 uL of PBS/BSA per immunoprecipitation.
3. Incubate overnight at 4°C with agitation (we use a Hematology Chemistry Mixer 346 from Fisher).
4. In a cold room, wash the beads twice with 1 mL cold Lysis buffer, twice with 1mL Lysis buffer supplemented with 360 mM NaCl, twice with 1 mL cold Wash buffer and once with 1 mL cold TE.
 - i. Note : The washes can be done quickly using a magnetic system (MCP-S from Dynal). Put the tubes to the magnetic system and wait a few

seconds for the beads to get captured by the magnet. By keeping the tubes in the system, remove the liquid by aspiration or simply by inverting the tubes. Add the next washing solution, close the tubes, remove the magnet and agitate to resuspend the beads. Put the magnet back on and remove the liquid as above.

5. After the last wash (TE), centrifuge briefly at 4°C and remove the last bit of liquid using a pipette.

2.1.9 DNA Purification

The cross-link between DNA and the proteins needs to be reversed. This is done here by the action of Tris which, with heat, attacks the DNA-protein bonds generated by the formaldehyde and liberates the DNA. The DNA is purified by a succession of RNase A and protease K treatments, followed by phenol extractions and ethanol precipitation.

2.1.9.1 Material required

- TE/SDS (10 mM Tris pH 8.0, 1 mM EDTA, 1 % SDS)
- TE (10 mM Tris pH 8.0, 1 mM EDTA)
- DNase-free RNaseA solution (10 mg/mL in 10mM Tris-Cl pH 8.0 /50 % glycerol)
- 10 % SDS
- Proteinase K solution (20 mg/mL from Gibco)
- Glycogen (20 mg/mL from Roche)
- Phenol/chloroform/isoamyl alcohol (25:24:1 from Sigma)
- 5M NaCl
- 100 % freezer-cold Ethanol
- 70 % freezer-cold Ethanol

2.1.9.2 DNA purification

1. Add 50 uL of TE/SDS to the bead pellets, vortex and incubate overnight at 65°C (use an oven instead of a water bath in order to minimize evaporation and condensation in the cap of the tubes).
 - i. Note : When using the input DNA (WCE) as control, add 50uL of TE/SDS to 5uL of WCE and process like the immunoprecipitated sample.
2. Vortex, centrifuge briefly and transfer the supernatant to a new tube
3. Add 350 uL of a mixture of (345 uL TE, 3 uL RNase A, 2 uL glycogen), mix by vortexing and incubate 2 hours at 37°C.
4. Add 15 uL of 10% SDS and 7.5 uL of proteinase K solution, mix by vortexing and incubate 2 more hours at 37°C.
5. Extract twice with 400 uL phenol/chloroform/isoamyl alcohol.
6. Add 12-16 uL 5M NaCl (200mM final concentration), vortex; add 2 volumes of 100 % ethanol, vortex and centrifuge 20 minutes at maximum speed.
7. Pour off liquid and wash with 500uL 70% ethanol.
8. Centrifuge 10 minutes and pour off liquid.
9. Centrifuge briefly, remove the last bit of liquid with a pipette and let dry 1 minute (Do not overdry the pellet).
10. Resuspend the pellet (which should be very small but still visible) in 50 uL TE.
11. Store at -20 °C.

2.1.10 Amplification and labeling

The labeling protocol described here is based on the one published by Ren et al. [285] with modification inspired by Hegde et al [287]. In this protocol, the immunoprecipitated DNA fragments are amplified and labeled by ligation-mediated PCR (LM-PCR) where aminoallyl-modified dUTP is incorporated in the product. The labeling

is done post-PCR using monoreactive Cy-dye NHS esters that will react specifically with the aminoallyl-modified dUTP. This method is less expensive and yields better results than the direct incorporation of Cy-dye-modified dNTPs during PCR. We routinely use Cy5 to label the test ChIP sample (tagged) and Cy3 to label the control ChIP sample (non-tagged). We found that dye swap experiments are not necessary with this protocol. The source of reagents are indicated here as a suggestion. They represent combinations that give good results in our laboratory. While it probably does not matter for most of other reagent, we found that using the 5X ligase buffer from Invitrogen (which contains polyethylene glycol) and the high concentration ligase from Invitrogen makes an improvement compared to other brands).

2.1.10.1 Material required

- BSA (10 mg/mL from New England Biolabs)
- dNTP mix (20 mM each dATP, dCTP, dGTP and dTTP)
- 3 M NaOAc pH 5.2 (Sigma)
- 0.1M NaOAc pH 5.2
- Glycogen (20 mg/mL from Roche)
- Phenol/chloroform/isoamyl alcohol (25:24:1 from Sigma)
- Freezer-cold 100 % ethanol
- Freezer-cold 70 % ethanol
- Annealed linker (see protocol below)
- T4 DNA polymerase (3 U/uL) (New England Biolabs)
- NEB #2 buffer
- T4 DNA ligase High concentration (5 U/uL) (Invitrogen)
- 5X ligase buffer (Invotrogen)
- 10X ThermoPol buffer (New England Biolabs)
- 50mM 5-(3-aminoallyl)-2'-deoxyuridine-5'-triphosphate (aa-dUTP) (Ambion)
- aa-dUTP dNTP mix (5 mM each dATP, dCTP, dGTP, 3 mM dTTP, 2mM aminoallyl-dUTP)
- Monoreactive Cy-5 NHS Ester (GE Healthcare)
- Monoreactive Cy-3 NHS Ester (GE Healthcare)
- Oligo 1 (GCGGTGACCCGGGAGATCTGAATTC) (40 uM stock)
- Oligo 2 (GAATTCAGATC) (40 uM stock)
- Taq DNA polymerase (Perkin Elmer)

- PFU DNA polymerase (Fermentas)
- Qiaquick PCR purification kit (Qiagen)
- 80 % Ethanol/5mM KPO₄ pH 8.5 buffer (Phosphate wash buffer)
- 4mM KPO₄ pH 8.5 buffer (Phosphate elution buffer)
- 0.1M Na₂CO₃ pH 9.0 (make fresh every few weeks to a month)

2.1.10.2 Amplification and labeling

1. Transfer 40 uL of immunoprecipitated DNA to a 0.5 mL PCR tube. Keep the sample on ice.
2. Add 70 ul of blunting mix (11 uL 10x NEB #2 buffer, 0.5 uL BSA, 0.5 uL dNTPs, 0.2 uL T4 DNA polymerase and 57.8 uL ice-cold water).
3. Mix by pipetting and incubate 20 minutes at 12°C (we use a thermocycler block).
4. Place on ice and add 12 ul of (11.5 uL 3M NaOAc and 0.5 uL glycogen).
5. Vortex and add 120 uL of phenol/chloroform/isoamyl alcohol.
6. Vortex and centrifuge 5 minutes at maximum speed.
7. Transfer 110 uL to a new 1.5 mL tube and add 230 uL 100% ethanol.
8. Vortex and centrifuge 20 minutes at 4 °C.
9. Pour off supernatant and add 500 uL 70% ethanol.
10. Vortex and centrifuge 10 minutes at 4 °C.
11. Pour off supernatant, centrifuge briefly and remove the remaining liquid with a pipette.
12. Let air dry 1 minute, add 25 uL ice-cold water and put on ice for about 30 minutes.
13. Add 25 uL of ligase mix (8 uL water, 10 uL 5X ligase buffer, 6.7 uL annealed linkers, 0.5 uL T4 DNA ligase).
 - i. Note : The annealed linker is prepared by annealing oligos 1 and 2: Mix 250 uL of 1M Tris pH 7.9, 375 uL of oligo 1 and 375 ul of oligo 2. Make 50 uL aliquots in 1.5 mL tubes. Place in a 95°C heat block for 5 minutes. Transfer tubes to a 70°C water-filled heating block. Place the

block at room temperature and let it cool down to 25°C. Transfer the block to 4°C and allow it to cool down overnight. Store annealed linkers at -20 °C.

- ii. Note : The T_m of the linker is very low. Always keep the linkers on ice and use ice-cold buffers when manipulating to avoid de-annealing.
14. Mix by pipetting and incubate overnight at 16°C (a waterbath is fine for this step).
 15. Add 6 uL of 3M NaOAc, vortex and add 130 uL 100% ethanol.
 16. Vortex to mix and centrifuge 20 minutes at 4 °C, maximum speed.
 17. Pour off supernatant and add 500 uL 70% ethanol.
 18. Vortex and centrifuge 5 minutes at 4 °C, maximum speed.
 19. Pour off supernatant, centrifuge briefly and remove remaining liquid with a pipette.
 20. Let air dry for a minute (do not overdry the pellet), and resuspend in 25 uL ice cold water by pipetting up and down and rinsing the side of the tube (the pellet at this stage is usually diffused on the side of the tube).
 21. Leave on ice for about 30 minutes, vortex, centrifuge briefly and put on ice.
 22. Add 15 uL of labeling mix (4 uL 10X ThermoPol buffer, 7.75 uL water, 2 uL aa-dUTP dNTP mix, 1.25 uL oligo 1).
 23. Mix by pipetting and transfer to a new 0.5 mL PCR tube.
 24. Put in thermocycler and start the program (see below).
 25. When temperature reaches 55°C, add 10 uL of enzyme mix (8 uL water, 1 uL 10X ThermoPol buffer, 1 uL Taq polymerase, 0.01 uL PFU polymerase) and mix by pipetting.
 - i. Note : Use the following PCR program for amplification : Step1 : 4 minutes at 55 °C, Step2 : 5 minutes at 72 °C, Step3 : 2 minutes at 95 °C, Step4 : 30 seconds at 95 °C, Step5 : 30 seconds at 55 °C, Step6 : 1 minute at 72 °C, Step7 : Go back to Step4 31 more times, Step8 : 4 minutes at 72 °C, Step9 : 4 °C forever.

- ii. Note : At the end of the PCR reaction, the amplified DNA can be visualized by running a 5 uL aliquot on a 1.5% agarose gel stained with ethidium bromide.
26. Purify PCR reactions on Qiaquick PCR purification kit. Use manufacturer's protocol, but replace buffer PE with the Phosphate wash buffer and buffer EB with the Phosphate elution buffer. Do the wash step once. Elute twice with 30 uL phosphate elution buffer (60 uL total volume).
27. Dry the eluates in a speed-vac.
28. Resuspend the pellets in 4.5 uL fresh 0.1M Na₂CO₃ pH 9.0 buffer.
29. Add 4.5 uL of the appropriate Cy-dye (we use Cy5 for the IP and Cy3 for the control).
 - i. Note : Resuspend new tubes of the Cy3 and Cy5 dyes in 73 uL of DMSO (kept desiccated under vacuum). Store at -80°C.
30. Incubate 1 hour in the dark at room temperature.
31. Add 35 uL 0.1M NaOAC pH 5.2.
32. Add 250 uL buffer PB (Qiagen).
33. Purify on Qiaquick PCR purification columns, using the manufacturer's protocol and using the manufacturer's buffers. Do the wash step once. Elute twice with 30 uL buffer EB (60 uL total volume).
 - i. Note : The color should be visible on the filter and the eluate should have a faint but visible taint of blue/pink.
34. Dry the eluates in a speed-vac.
35. Store the dry colored pellets at -20°C in the dark.
 - i. Note : The samples are now ready to be hybridized on DNA microarray. Use standard hybridization techniques.

2.1.11 Comparison between experimental designs

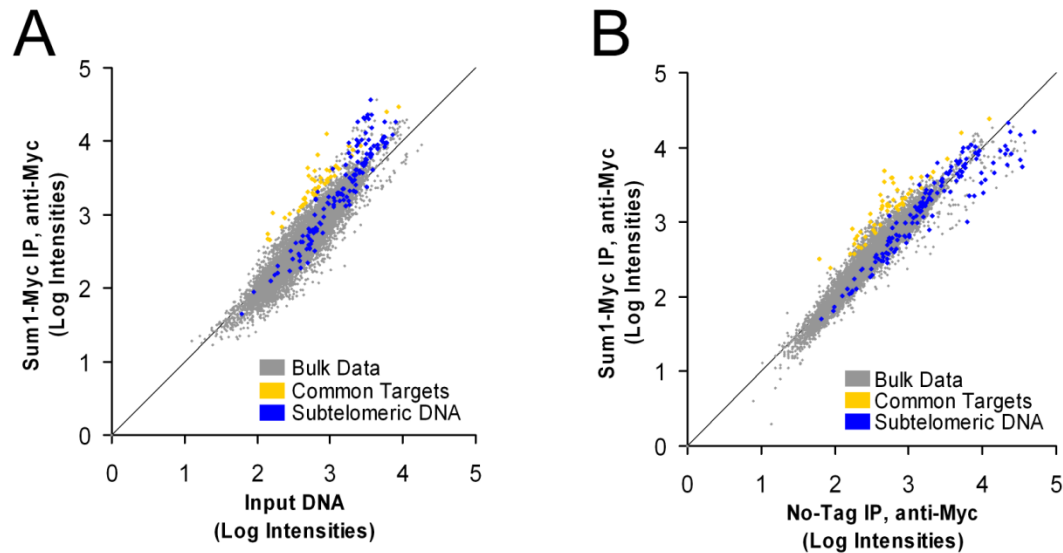
As discussed above, two approaches have been used in the past in ChIP on CHIP experiments. One uses input DNA as the control for the microarray, the other uses a control immunoprecipitation. Both these approaches are susceptible to noise, but the sources of the noise are different. In order to determine which one generates the cleanest datasets, we performed a side-by-side comparison. We used a yeast strain containing a myc-tagged version of the Sum1 transcriptional repressor to generate ChIP samples. These samples were labeled with Cy5 as described above and hybridized either with a Cy3-labeled DNA preparation coming from non-immunoprecipitated extract or from Cy3-labeled ChIP material coming from an isogenic non-tagged strain. As expected, the two datasets contain quite a significant overlap. In addition, however, the dataset generated from the first method (using input DNA as the control) contains subtelomeric DNA regions while these regions are not found when a control ChIP from a non-tagged strain is used as the control (**Figure 2-1**). Also, when the promoter sequences from the genes identified in those two experiments were analyzed, a larger proportion (twice as many) contained the Sum1 DNA-binding site (DNCRCAAAW) when a non-tagged immunoprecipitation was used as the control (not shown).

The presence of subtelomeric DNA in ChIP generated from the first method as shown in **Figure 2-1** is routinely observed (not shown), suggesting it represents noise. Indeed, these DNA regions are enriched even when anti-myc antibodies are used to ChIP strains containing no myc-tagged proteins (**Figure 2-2A**). Interestingly, we found that the enrichment generated by the anti-myc antibody is SIR2-dependent, suggesting that the anti-myc antibody cross-reacts with Sir2p, or with a SIR2-dependent structure present at telomeres (data not shown). This effect is specific to the anti-myc antibody since no subtelomeric DNA is enriched when an anti-HA antibody is used. However, anti-HA antibodies tend to enrich for tRNA genes (**Figure 2-2B**). Taken together, these data suggest that even highly specific and well characterized antibodies can have enough cross-reactivity

to bring down sufficient amount of DNA to generate noise in ChIP-CHIP experiment when input DNA is used as the control.

2.1.12 Conclusions

In conclusion, we found that using ChIP samples using the same antibody but on an isogenic strain containing no tagged protein represent a better design since the noise generated by antibody cross-reactivity is cancelled out. The enrichment generated by this cross-reactivity is small; therefore it usually does not create too many problems since most people are interested in the topmost enriched targets, which are usually greatly outranking the noise described here. However, these false positives can become a source of non-negligible noise when the protein of interest generates small enrichment or when more refined analyses need to be performed. When a sufficient amount of datasets are generated from the same platform, some mathematical models can be applied to eliminate the noise specific to said platform, as demonstrated by the Young group [288]. This method is very powerful and has the advantage of eliminating not only the noise coming from the phenomenon described here, but also the noise that may be incorporated by the labelling and hybridization processes. However, since this mathematical method requires many independent datasets generated on the same platform, we recommend the use of the method described here which has the advantage of eliminating the noise emanating from the antibody at the source.



1.

Figure 2-1 : Genome-wide Location Analysis of the transcription factor Sum1 using input DNA (A) or a control IP (B) as the hybridization control.

(A) Scatter plot from a genome-wide location analysis of the transcription factor Sum1 using input DNA as the hybridization control. The log intensity of the DNA coming from the immunoprecipitated sample is plotted against the log intensity of the DNA coming from a non-immunoprecipitated extract (input DNA) for all the spots on the microarray.

(B) Same as (A) except that the hybridization control (x axis) is DNA coming from an immunoprecipitation performed on an isogenic strain containing no tagged protein.

In both (A) and (B), the bulk data is shown in grey, the genes bound ($p < 0.001$) in both experiments is shown in yellow, and the subtelomeric DNA (sequences within 5KB from a chromosome end) are shown in blue.

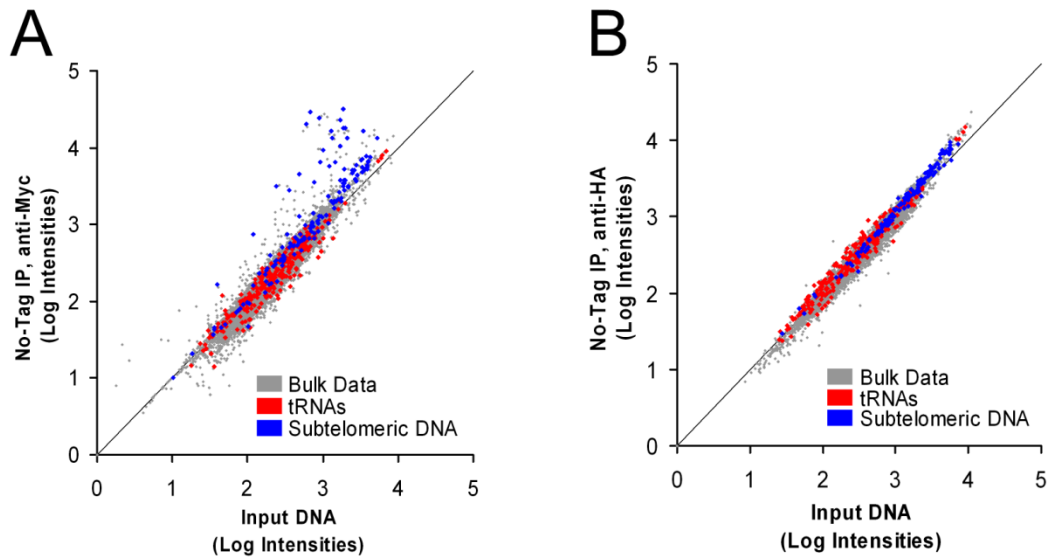


Figure 2-2 : Control Genome-wide Location Analysis experiments performed on yeast strain containing no tagged protein.

The immunoprecipitation was performed using an anti-Myc antibody (A), and an anti-HA antibody (B).

(A) Scatter plot of a Genome-wide Location Analysis from a yeast strain containing no tagged protein. The log intensity of the DNA coming from the Myc-immunoprecipitated sample is plotted against the log intensity of DNA coming from a non-immunoprecipitated extract (input DNA) for all the spots on the microarray.

(B) Same as (A) except that an anti-HA antibody was used.

In both (A) and (B), the bulk data is shown in grey, the subtelomeric DNA (sequences within 5KB from a chromosome end) are shown in blue, and tRNA genes are shown in red.

3 Caractérisation de l'activité et du mode de recrutement du complexe HDAC Rpd3S aux régions codantes activement transcrites

La découverte de l'existence de deux complexes indépendants comprenant la HDAC Rpd3 [117, 118] a provoqué un grand intérêt au sein de la communauté scientifique. On croyait jusqu'alors que Rpd3 n'était impliquée que dans la répression de l'expression des gènes auxquels elle était recrutée, un rôle maintenant attribué au complexe Rpd3L.

La fonction de Rpd3S a été inférée via la structure de sa sous-unité Eaf3. On savait déjà qu'une délétion d'*EAF3* avait un effet sur le niveau d'acétylation des nucléosomes aux promoteurs et aux régions transcrites de certains gènes [124] et démontre un phénotype de transcription cryptique [117, 122]. Eaf3 possède également un chromodomaine, reconnu pour sa capacité à lier les lysines méthylées. Cette affinité a été confirmée *in vitro* à l'aide de mononucléosomes immobilisés [118] et de peptides [117, 122]. Il n'en fallut pas plus pour postuler que Rpd3S déacétyle probablement les nucléosomes des régions codantes de gènes activement transcrits pour remettre la chromatine de ces régions dans l'état où elle était avant le passage de la machinerie transcriptionnelle.

Toutefois, cette hypothèse ne repose que sur des données générées *in vitro* qui n'ont jamais été confirmées *in vivo*. L'interaction spécifique entre Eaf3 et le résidu H3K36 méthylé est probablement importante à la fonction de Rpd3S ou au recrutement de ce complexe aux régions transcrites, mais encore faut-il le confirmer expérimentalement. Cette confirmation est à l'origine du projet menant aux résultats surprenants présentés dans le manuscrit ci-joint. En effet, j'ai découvert que la méthylation du résidu H3K36 n'était pas suffisante pour recruter le complexe Rpd3S, mais par contre nécessaire à l'activité HDAC de ce complexe. De plus, j'ai déterminé que Rpd3S était très probablement recruté par une interaction avec le domaine C-terminal de la sous-unité Rpb1 de l'ARN polymérase II et que la phosphorylation de ce domaine en sérine 2 et 5 était nécessaire pour ce recrutement. Finalement, j'ai déterminé que le complexe d'élongation DSIF, composé des protéines Spt4 et Spt5, a une influence négative sur le recrutement de Rpd3S à l'ARN polymérase II.

Mes résultats redéfinissent le modèle couramment accepté du recrutement et de l'activité de Rpd3S et fournissent une hypothèse alternative intéressante fondée sur l'étude *in vivo* de la dynamique de ces systèmes à l'échelle génomique. Cette nouvelle hypothèse permettra de nouvelles avancées dans l'étude des multiples mécanismes régissant la transcription par l'ARN polymérase II.

Ces résultats ont été publiés dans PLoS Genetics en 2010 [289].

3.1 Manuscrit :

« DSIF and RNA polymerase II CTD Phosphorylation Coordinate the Recruitment of Rpd3S to Actively Transcribed Genes »

Simon Drouin, Louise Laramée[§], Pierre-Étienne Jacques[§], Audrey Forest, Maxime Bergeron and François Robert

[§] Ces auteurs ont contribué de manière égale à ce travail.

Contribution des coauteurs:

S. Drouin: Design, exécution et analyse expérimentale. Rédaction conjointe avec F. Robert.

L. Laramée: Construction de souches et exécuté quelques expériences.

P.-É. Jacques: Contribution à l'analyse et à l'élaboration des figures.

A. Forest: Exécution de quelques expériences.

M. Bergeron: Exécution de quelques expériences.

F. Robert: Contribution au design et à l'analyse. Rédaction conjointe avec S. Drouin.

Ce manuscrit a été publié dans PLoS Genetics en 2010 [289].

3.1.1 Abstract

The yeast histone deacetylase Rpd3 is part of two distinct complexes: the large (Rpd3L) and small (Rpd3S) complexes. While Rpd3L targets specific promoters for gene repression, Rpd3S is recruited to ORFs to deacetylate histones in the wake of RNA polymerase II, to prevent cryptic initiation within genes. Methylation of histone H3 at lysine 36 by the Set2 methyltransferase is thought to mediate the recruitment of Rpd3S. Here, we confirm by ChIP-Chip that Rpd3S binds active ORFs. Surprisingly, however, Rpd3S is not recruited to all active genes, and its recruitment is Set2-independent. However, Rpd3S complexes recruited in the absence of H3K36 methylation appear to be inactive. Finally, we present evidence implicating the yeast DSIF complex (Spt4/5) and RNA polymerase II phosphorylation by Kin28 and Ctk1 in the recruitment of Rpd3S to active genes. Taken together, our data support a model where Set2-dependent histone H3 methylation is required for the activation of Rpd3S following its recruitment to the RNA polymerase II C-terminal domain.

3.1.2 Keywords

Histone deacetylase (HDAC); Histone methyltransferase (HMT); Histone H3K36 methylation; Cryptic transcription; Rpd3S recruitment; ChIP-chip; Set2; DSIF; Spt4; Spt5; Kin28; Ctk1; Bur1; Rpd3; Rco1; Eaf3.

3.1.3 Introduction

Histone acetylation was the first covalent histone modification shown to be involved in transcription regulation. Indeed, histones at the promoter of active genes tend to be hyper-acetylated while repressed genes have promoters with hypo-acetylated nucleosomes. It is now well established that this is due to the recruitment of histone acetyltransferases

(HATs) and histone deacetylases (HDACs) by transcriptional activators and repressors, respectively [290]. The first described and best characterized HDAC is Rpd3. The repressive effect of yeast Rpd3 on transcription has been well studied over the last 15 years, paving the way for the characterization of its mammalian orthologs [291]. Yeast Rpd3 is recruited to the promoters of specific genes by DNA-binding repressors, leading to the repression of many important pathways such as stress response, meiosis, the cell cycle and others [36, 86, 87, 89, 90, 97, 98, 100, 113, 292-299]. Moreover, Rpd3 also plays roles in silencing [85, 104-106, 112, 300-302], DNA replication [95, 96, 303, 304] and recombination [94, 305].

Recent proteomic studies have determined that Rpd3 is found in two distinct complexes: the large (Rpd3L) and the small (Rpd3S) complex [117, 118]. Both complexes share a core composed of Rpd3, Sin3 and Ume1. The large complex is composed of 11 additional proteins whereas the small complex contains only two additional subunits, namely Rco1 and Eaf3. While Rpd3L is likely responsible for the repressive function of Rpd3, the function of Rpd3S remains far less understood. Recent work by several groups has shown that Rpd3S is involved in the suppression of cryptic transcription [117, 306, 307] and that its activity is linked to the Set2 histone methyltransferase (HMT). Furthermore, *in vitro* studies have shown that Rpd3S is recruited to H3K36-methylated nucleosomes and that its Rco1 and Eaf3 subunits are essential for this recruitment to happen [117, 118, 122, 307]. Rco1 mediates interactions with histones in a modification-independent manner through a PHD zinc finger domain, while Eaf3 contains a methyl-lysine binding chromodomain (CHD) that is essential for recognition of H3K36-trimethylated (H3K36me3) nucleosomes *in vitro*. These studies also indicate that genome-wide histone acetylation levels on promoters and coding regions are altered when either H3K36 methylation or Rpd3S is disrupted.

Taken together, these data lead to a model where Rpd3S is recruited to coding regions through the interaction of Eaf3 with H3K36me3 in order to deacetylate

nucleosomes after the passage of the transcriptional machinery [308]. Deacetylation would allow chromatin disrupted by elongating RNA polymerase II (RNAPII) to return to a more ordered and compact structure, thereby restoring an environment hostile to cryptic transcription initiation within the coding region. However, this model relies heavily on the *in vitro* observation that the Eaf3 CHD binds preferentially to H3K36me3 peptides or nucleosomes [117, 118, 122, 307]. It was never formally demonstrated that this interaction is required for the targeting of Rpd3S to coding regions *in vivo*. In addition, whether Rpd3S is recruited to all transcribed genes or to only a subset of them has never been assessed. In order to address these questions, we performed genome-wide ChIP-chip experiments looking at both Rpd3S- and Rpd3L-specific subunits in wild type cells and in various mutants, including *set2* Δ and H3K36A.

Quite interestingly, our data show that Rpd3S specifically binds to the coding region of actively transcribed genes whose promoters are also bound by Rpd3L. Surprisingly, the binding of Rpd3S to active genes does not depend on Set2-mediated H3K36 methylation. However, methylation by Set2 is required for the activity of Rpd3S, as assayed by histone acetylation and RNAPII levels. We also provide *in vivo* evidence that Rpd3S is recruited to active genes via the phosphorylation of the RNAPII C-terminal domain (CTD). Finally, we show that the yeast DSIF transcription elongation factor negatively regulates Rpd3S recruitment. Based on these results, we propose that the recruitment and activity of Rpd3S on ORFs depend on a two steps mechanism: an initial recruitment to the elongation complex - coordinated by DSIF and RNAPII phosphorylation - followed by the H3K36me3-dependant modulation of Rpd3S activity through the Eaf3 chromodomain.

3.1.4 Results

3.1.4.1 Rpd3S is not recruited to all transcribed genes

While it is generally accepted that Rpd3L negatively regulates specific sets of genes, the specific function of Rpd3S remains largely unaddressed. However, it is expected to be ubiquitously recruited to active genes since it interacts with methylated H3K36. In order to address the specificity of Rpd3S and Rpd3L in a systematic manner, we performed ChIP-chip experiments on myc-tagged Rpd3, Rco1 and Sds3, the last two being specific subunits of Rpd3S and Rpd3L, respectively. As shown in **Figure 3-1A**, Rpd3 binds to a large subset of promoters (**Figure 3-1A**, clusters 1 and 2). In addition, a subset of these genes also exhibit Rpd3 binding on their coding regions (**Figure 3-1A**, cluster 2). **Figures 3-1B** and **3-1C** show the average signal of Rpd3, Rco1 and Sds3 over the genes from cluster 1 and cluster 2, respectively. The data for cluster 3, representing the genes not bound by Rpd3, are also shown. Cluster 1, which is enriched for genes previously demonstrated to be repressed by Rpd3 (genes involved in M phase (p-value 10^{-8}), cell cycle (p-value 10^{-9}), etc.), shows binding of both Rpd3 and Sds3 to the promoter. The presence of Rco1 is not detectable on these genes. This cluster therefore represents genes repressed by Rpd3L, which is consistent with the fact that these genes have low level of RNAPII (**Figure 3-7A**). Cluster 2, however, shows evidence for the presence of both Rpd3L and Rpd3S since all three subunits tested for are detected. Rco1 is restricted to the coding region of these genes, consistent with the fact that they are actively transcribed (**Figure 3-7A**). Sds3 is present at the promoter, which is expected since Rpd3 binds to these promoters. More surprisingly, however, some level of Sds3 is also detected on the coding region of these genes. These data - also observed with another subunit of Rpd3L (Rxt2; **Figure 3-7B**) - suggest that the large complex may play some role during transcriptional elongation, perhaps in conjunction with the small complex (see Discussion). Nevertheless, the data presented here clearly show that Rpd3S binds to the coding region of active genes.

Strikingly, these experiments also show that Rpd3S preferentially associates with genes that are also bound by Rpd3L. In fact, we found no clusters of genes where Rpd3 binds in the coding region but not in the promoter. This suggests that Rpd3S, contrary to what has been expected, does not ubiquitously bind to active genes but rather targets some of them, namely a subset of those that are bound by Rpd3L at their promoter. In order to test if Rpd3S ubiquitously binds active genes, we performed ChIP-chip experiments of RNAPII and H3K36me3, two proxies for active gene expression. **Figure 3-1D** (cluster 4) clearly shows that many genes, despite having strong enrichment for RNAPII and H3K36me3, show no evidence of Rpd3S binding. Three conclusions can be drawn from these results; firstly, they confirm that the recruitment of Rpd3S is not a general phenomenon occurring on all transcribed genes; secondly, they suggest that the methylation of H3K36 by Set2 is not providing the specificity for the recruitment of Rpd3S; and finally they suggest that the large complex may play a role in the recruitment of the small complex as Rpd3S appears to target primarily genes also bound by Rpd3L.

3.1.4.2 Set2 is not required for Rpd3S binding to ORFs *in vivo*

The idea that H3K36me3 recruits Rpd3S through the Eaf3 subunit is well established in the literature [308]. Work done *in vitro* by several groups has clearly shown this using peptides and nucleosomal substrates [117, 118, 122, 307]. Our ChIP-chip experiments, however, clearly demonstrate that many genes harboring high levels of H3K36me3 are free of Rpd3S. We have therefore endeavored to examine whether the well characterized interaction between the Eaf3 CHD and H3K36me3 is responsible for the targeting of Rpd3S *in vivo*.

First, we looked at Rco1 binding in a *set2* Δ mutant. Since this mutant cannot methylate H3K36, the current model predicts that Rpd3S should not bind to ORFs under these conditions. To our considerable surprise, Rco1 enrichment on ORFs in this mutant is

not significantly altered for about two-thirds of Rpd3S target genes (**Figure 3-2A**, cluster 5). For other genes, Rpd3S occupancy is decreased significantly, although not completely abolished (**Figure 3-2A**, cluster 6). We next repeated these experiments in a H3K36A mutant (where lysine 36 is mutated into an alanine) with similar results (**Figure 3-2A**). The deletion of the Rco1 PHD domain had no effect on Rpd3S occupancy (despite the fact that it destabilizes the Rco1 protein; **Figure 3-2E**), while deletion of the Eaf3 CHD domain phenocopied *set2* Δ and H3K36A. Interestingly, a *set1* Δ /*set2* Δ /*dot1* Δ triple mutant (abolishing all histone methylation activity in yeast) was similar to wild type, suggesting that deletion of *SET1* and/or *DOT1* can partially suppress the *set2* Δ phenotype. Taken together, these data demonstrate that H3K36 methylation is not required for the recruitment of Rpd3S to most genes.

Since Rpd3S occupancy seems to be more dependent on H3 methylation at some genes than others, we looked more closely at clusters 5 and 6. As shown in **Figures 3-2B** and **3-2C**, clusters 5 and 6 are markedly different with regards to transcription levels. While cluster 5 is highly transcribed (as shown by the presence of high levels of both RNAPII and H3K36me3), cluster 6 is less so. Next we looked at the distribution of Rco1 on genes contained within these clusters in all strains shown in **Figure 3-2A**. As expected from data shown in **Figure 3-2A**, Rco1 occupancy on ORFs is not (or only slightly) affected in these mutants for the cluster 5 genes, but it is reduced for the genes from cluster 6 (**Figure 3-2D** and **Figure 3-8**). In addition, for all Rpd3S-bound genes, a redistribution of Rco1 to the promoter region was observed in all mutants tested. This redistribution towards the promoter remains unexplained but correlates with our observation that histone acetylation is decreased at promoters in these same mutants (**Figure 3-9**). Collectively, these data clearly demonstrate that H3K36 methylation has no impact on Rpd3S occupancy at highly transcribed genes (genes from cluster 5). However, the methyl mark, or the ability to recognize it through the Eaf3 chromodomain, is important for optimal Rpd3S association to ORFs with lower levels of RNAPII (genes from cluster 6).

3.1.4.3 Set2 and Rpd3S regulate the distribution of RNAPII and histone acetylation

It is known that *set2Δ* mutants, along with null mutants of Rpd3S subunits Eaf3 or Rco1, exhibit a cryptic transcription phenotype [117, 306, 307]. This phenotype is thought to be due to the improper deacetylation of transcribed ORFs by Rpd3S after each round of transcription [308] because of a lack of Rpd3S recruitment. Other groups that have characterized acetylation levels on coding regions in Set2 and Rpd3S mutants either looked at bulk chromatin by western blotting [307], or at specific genes by ChIP [117, 118, 122, 309] and have come to the conclusion that acetylation levels increase on ORFs when Set2 or Rpd3S is disrupted.

Since we –quite surprisingly- observed, however, that Rpd3S binding to genes is mostly independent of histone methylation by Set2, we decided to test whether the activity of Rpd3S requires methylation of H3K36 by Set2. To do so, we looked at H4K5 acetylation (H4K5ac) by ChIP-chip. We used H4K5 acetylation to score for Rpd3S activity because it was shown previously to be a robust readout for Rpd3 activity in ChIP-chip assays [138]. Similar to other groups [117, 118, 122, 309], we observed decreased acetylation on promoters in *set2Δ*, H3K36A or Rpd3S mutants (**Figure 3-9**). Histone acetylation is also dramatically affected across ORFs in these mutants. As shown in **Figure 3-3A**, we observed a loss of acetylation for normally highly acetylated ORFs, and a gain in acetylation for ORFs that exhibit low levels in the wild type. Because Set2 and Rpd3S are both known to prevent cryptic initiation within ORFs, we repeated the same analyses on RNAPII ChIP-chip results and found a similar pattern to that observed for H4K5ac, namely that ORFs with high RNAPII enrichment show decreased RNAPII levels in the absence of Set2 or Rpd3S, whereas ORFs with low RNAPII tend to display higher levels of polymerase (**Figure 3-3B**). These results clearly show that the activity of the Rpd3S complex requires methylation of H3K36 by Set2.

The effect of the loss of Rpd3S activity on histone acetylation and RNAPII distribution on ORFs is more complex than previously described. Our data indeed suggest that both histone acetylation and RNAPII occupancy are redistributed in a more even manner across the genome than expected. A plausible explanation of the genome-wide averaging of RNAPII and histone acetylation levels in Set2 and Rpd3S mutants would entail aberrant recruitment of the transcriptional apparatus to low-expression genes whose coding regions were not reset properly by a, now inactive, Rpd3S complex. Assuming a limited pool of transcription machinery in a cell, this would result in a lower abundance of RNAPII at the more active ORFs, and aberrant genomic acetylation levels.

These results, combined with the Rpd3S occupancy profiles in *rcol1Δ*, *rcol1-PHDΔ* and *eaf3-CHDΔ* mutants, suggest that the loss of histone H3 methylation at lysine 36 affects ORF identity through a modulation of Rpd3S deacetylase activity rather than through altered recruitment of the Rpd3S complex on coding regions as previously thought. This hypothesis is further supported by the results obtained in the catalytically inactive Rpd3 mutant (*rpd3-H188A*), which recapitulate both the H4K5ac and RNAPII results (**Figure 3-3**).

3.1.4.4 Spt4 is required for the proper localization of Rpd3S

Since the interaction between the Eaf3 CHD and H3K36me3 does not account for the initial recruitment of Rpd3S to active genes, we decided to look for a factor that would fulfill that role. Quan and Hartzog have shown genetic interactions between H3K36 methylation and Rpd3S with Spt5 [133]. Their data suggest that Rpd3S opposes the function of the elongation factor Spt4/5, which is the yeast ortholog of the human elongation factor DSIF [310]. DSIF negatively regulates elongation in its non-phosphorylated form, but is turned into a positive elongation factor upon phosphorylation by P-TEFb (Bur1 in yeast) [121, 311, 312]. The genetic interaction between Rpd3S and

Spt5 led us to test whether Spt4/Spt5 is involved in the recruitment of Rpd3S. We therefore performed ChIP-chip experiments of Rco1 in *spt4Δ* cells (the deletion of *SPT5* is lethal). As shown in **Figure 3-4A**, deletion of *SPT4* leads to massive changes in Rco1 binding across the genome. Notably, the effect is far more dramatic compared to the deletion of *SET2* (**Figure 3-4A**). Importantly, the level of RNAPII observed on these genes is not significantly affected in the mutant, ruling out the possibility that the effect is solely due to a reshuffling of the transcriptome (Pearson correlation = 0.94, **Figure 3-10**). The deletion of *SPT4* causes a decrease of Rco1 binding at some transcribed genes normally strongly associated with Rco1 (**Figure 3-4A**, cluster 7), as well as an increase at others where Rco1 is otherwise only found at low levels (cluster 8). Deletion of *SPT4* even causes a slight increase of Rco1 occupancy at genes where it is normally undetectable (cluster 9). Overall, these effects lead to a distribution of Rpd3S that correlates better with RNAPII occupancy than in wild type cells (**Figure 3-4B**). To test the possibility that, in the absence of Spt4, Rpd3S is recruited via H3K36 methylation, we profiled Rco1 binding in a *spt4Δ/set2Δ* double mutant. As shown in **Figure 3-4A**, deleting both *SPT4* and *SET2* leads to a similar Rpd3S localization phenotype compared to the single *spt4Δ* mutant, giving further evidence that Set2 does not play a large role in Rpd3S recruitment, even in the absence of Spt4.

3.1.4.5 Spt4 negatively regulates the association of Rco1 with transcribed genes.

To distinguish the direct effect of the loss of *SPT4* from eventual indirect effects of the mutation, we localized Spt4 in wild type cells by ChIP-chip using a strain carrying a myc-tagged *SPT4* gene. Spt4 associates with genes in a manner that correlates with levels of RNAPII (Pearson correlation = 0.84) suggesting that DSIF acts as a general elongation factor. Moreover, it is present across the whole ORF, indicating that it travels with RNAPII, but is further enriched in the 3' end of genes (**Figure 3-10B**), suggesting that it may also regulate the elongation-termination transition, as shown by others [313-315]. This is also consistent with the fact that Spt5 interacts with components of the capping and

termination machineries [128]. Even more interestingly, the more a gene is occupied by Spt4 in wild type cells, the more Rco1 we detect in the *spt4* Δ mutant (**Figure 3-4C**), suggesting that the direct effect of the loss of Spt4 is an increase in Rpd3S binding (as observed for cluster 8). Consequently, the decrease in occupancy observed in cluster 7 is most likely indirect since Spt4 is barely detectable at these genes (**Figure 3-4A**). Similarly, the level of Rco1 increases dramatically in the *spt4* Δ mutant on cluster 4 genes (from **Figure 3-1D**), representing transcribed genes highly occupied by Spt4 where Rco1 is absent in wild type cells (**Figure 3-10C**). In general, genes with a higher Spt4 / RNAPII ratio tend to have less Rco1 than genes with lower Spt4 / RNAPII ratios (Pearson correlation = -0.42, **Figure 3-10D**). Taken together, these data suggest that Spt4 acts as a negative regulator of Rpd3S binding and that its presence prevents the HDAC from freely associating with transcribed genes. This model is in agreement with genetic data showing that Rco1 opposes the function of Spt4/5 [133].

We then looked at the distribution of Rpd3S on transcribed genes where Spt4 is also bound (cluster 8) and found strong differences between the wild type and *spt4* Δ mutants. While Rco1 occupies the whole ORF at a constant level in wild type cells, it accumulates to abnormally high levels towards the 3' end of the gene in *spt4* Δ cells (**Figure 3-4D**, dashed line). This binding pattern is also found in the *spt4* Δ / *set2* Δ double mutant (compare dashed and dotted lines in **Figure 3-4D**). The occupancy profile of Rco1 in a *spt4* Δ mutant shows a strong similarity to the occupancy profile of RNAPII with a CTD phosphorylated at Ser2 (**Figure 3-4D**, red solid line). This led us to hypothesize that CTD phosphorylation by Ctk1 (the major serine 2 kinase) might be implicated in Rpd3S recruitment in the absence of Spt4/5. To test this hypothesis, we profiled Rco1 occupancy in a *spt4* Δ /*ctk1* Δ background. Surprisingly, *ctk1* Δ partially suppressed the *spt4* Δ Rpd3S binding pattern phenotype (**Figure 3-4A**) leading to an intermediate binding profile between wild type and *spt4* Δ . Looking at it more closely, we observed that short genes fail to accumulate Rco1 in *spt4* Δ / *ctk1* Δ cells (**Figure 3-4E**), whereas Rco1 is observed at genes irrespective of their length in wild type (**Figure 3-11**) or *spt4* Δ cells (**Figure 3-4F**), suggesting that Rpd3S

accumulates more slowly in *spt4Δ / ctk1Δ* cells. We therefore conclude that in the absence of Spt4, Rpd3S binds to the RNAPII CTD and that the phosphorylation of the CTD at serine 2 contributes to that phenomenon.

3.1.4.6 Phosphorylation of the RNAPII CTD triggers the association of Rpd3S to active genes

The data presented above demonstrate that Spt4 negatively regulates the association of Rpd3S to highly transcribed genes. The data also suggest that phosphorylation of the RNAPII CTD, notably at serine 2, plays some role in the recruitment of Rpd3S to active genes in the absence of Spt4. We next tested whether the phosphorylation of the CTD is implicated in the association of Rpd3S with transcribed genes in wild type cells and performed ChIP-chip experiments of Rco1 in mutants for the known CTD kinases. As shown in **Figure 3-5A**, deletion of *CTK1*, the major serine 2 kinase, has a clear effect on Rco1 occupancy (compare solid line with dashed line). To test the effect of serine 5 and 7 phosphorylation, we used a strain carrying ATP-analog-sensitive alleles of *KIN28* since the deletion of the *KIN28* gene is lethal. As shown in **Figure 3-5A** (dotted line), inhibition of Kin28 has a dramatic effect on Rco1 occupancy. Indeed, no Rco1 can be detected on ORFs in that mutant. This clearly demonstrates that phosphorylation of serine 5 and/or 7 by Kin28 is a major element in the recruitment of Rpd3S to active genes. These results are supported by data from the Hinnebusch lab who have shown that Rpd3S interacts with the phosphorylated form of RNAPII and has high affinity for doubly phosphorylated (serine 2/5) CTD peptides *in vitro* [316]. Interestingly, CTD peptides carrying a single phosphate group on serine 5 or serine 2 respectively have a much weaker or no affinity to Rpd3S compared to doubly phosphorylated CTD peptides. Also noteworthy is the fact that Rpd3S is redistributed to promoter regions when Kin28 is inactive (**Figure 3-5A**). As we will be discuss below, this might be due to a decrease in H3K36 methylation in that mutant, most likely due to a defective Bur1/2 recruitment, as suggested by [317] and [318] - leading to a

defect in Rad6 phosphorylation. Our genome-wide data, together with these *in vitro* experiments, suggest that phosphorylation of the RNAPII CTD stimulates the recruitment of Rpd3S to transcribed genes while the elongation factor DSIF counteracts this recruitment.

3.1.4.7 Phosphorylation of Spt5 by Bur1 negatively regulates the activity of DSIF on Rpd3S recruitment

In mammalian cells, DSIF is phosphorylated by Cdk9, a cyclin-dependent kinase associated with the elongation factor P-TEFb [319]. Cdk9 also phosphorylates the RNAPII CTD on serine 2 as well as other proteins including NELF. In yeast, the function of Cdk9 is fulfilled by two distinct kinases. Ctk1 mainly phosphorylates the RNAPII CTD and Bur1 targets Spt5 and Rad6 [129, 130, 320]. Inactivation of Bur1 has very modest effects on phosphorylation of the RNAPII CTD, as shown by western blot and by ChIP-chip experiments (data not shown; see also [130, 321, 322]). Because Bur1 phosphorylates Spt5, the partner of Spt4, we tested the effect of inactivating Bur1 activity on Rco1 recruitment. Not surprisingly, inhibiting Bur1 using an ATP-analog-sensitive strategy (*bur1AS*) has a profound effect on Rco1 occupancy. In the absence of a functional Bur1, Rco1 is depleted from the coding region and redistributed to promoter regions (**Figure 3-5B**, dashed line) as was observed in the Kin28 mutant while the RNAPII level is mostly unchanged (data not shown). Deleting the CTD of Spt5 (*spt5 Δ C*), the region phosphorylated by Bur1, caused a similar, although milder, phenotype (**Figure 3-5B**, dotted line). The stronger effect observed in the Bur1 inactivation experiment, compared to truncation of Spt5, is most likely due to the fact that Bur1 has additional targets. For example, Bur1 phosphorylates Rad6, an event that is required for the methylation of H3K4 and K79 by Set1 and Dot1 respectively [244, 245, 323]. In a triple mutant for Set1, Set2 and Dot1, we observed a shift of Rpd3S toward intergenic regions (**Figure 3-9B**), suggesting that the redistribution of

Rpd3S to promoter regions in Bur1-impaired cells is at least in part a consequence of Bur1's activity on Rad6. Nevertheless, on the coding region, where our *spt4Δ* data showed that DSIF negatively regulates Rpd3S binding, we observed a clear decrease in Rco1 occupancy in both *bur1AS* and *spt5ΔC*. These data strongly suggest that phosphorylation of Spt5 by Bur1 negatively regulates the activity of DSIF on Rpd3S recruitment. While we cannot completely rule out the possibility that some of the effect observed in the *bur1AS* strain is due to an effect of Bur1 on the RNAPII CTD, our data on *spt5ΔC* rather suggest that Bur1 regulates Rpd3S recruitment by regulating DSIF. Interestingly, phosphorylation of Spt5 by Bur1 was previously shown to stimulate its activity as an elongation factor while our data suggest that it inhibits its activity as a negative regulator of Rpd3S recruitment. It will therefore be interesting to see whether these activities are linked.

3.1.5 Discussion

The Rpd3S complex is recruited to the coding region of transcribed genes where it represses cryptic transcription by deacetylating histones in the wake of the elongating polymerase, therefore resetting chromatin to its pre-transcriptional state. The recruitment of Rpd3S to transcribed regions was thought to be mediated by the interaction between the Eaf3 CHD and H3K36me. However, the interaction of Rpd3S on the coding regions of actively transcribed genes has never been directly demonstrated *in vivo*. Here, we build on that model and show that: 1) Rpd3S targets a subset of transcribed genes; 2) Rpd3L is present at the promoter of genes where Rpd3S is bound; 3) The activity of Rpd3S requires methylation of H3K36 by Set2 but its association with active genes does not; 4) The recruitment of Rpd3S on ORFs requires the phosphorylation of the RNAPII CTD by Kin28 and Ctk1; and 5) The DSIF elongation complex counteracts the recruitment of Rpd3S to transcribed genes, a phenomenon that is regulated by the phosphorylation of Spt5 by Bur1. We therefore propose a model where the opposing effects of CTD phosphorylation and

DSIF on Rpd3S recruitment together allow for the complex occupancy profile of that HDAC *in vivo* (**Figure 3-6**).

3.1.5.1 The distribution of Rpd3S along the genome is more complex than expected

Because H3K36 methylation correlates with transcription, it is generally accepted to be present at all transcribed genes; Rpd3S was therefore expected to follow the same pattern. Our finding that many transcribed genes do not show any signs of Rpd3S binding despite the presence of H3K36me3 was therefore a considerable surprise. This has important implications since it suggests that not all genes are equally protected against cryptic initiation. Other mechanisms exist that prevent cryptic transcription [324], so it will be interesting to investigate how these various activities share the labor of protecting the genome from aberrant transcription.

3.1.5.2 Does Rpd3L play a role in the targeting of Rpd3S?

Another peculiarity of Rpd3S is that it is preferentially enriched on the coding regions of genes where Rpd3L is found on the promoter: we rarely find Rpd3S on genes where Rpd3L is not present. The presence of Rpd3L on the promoter of active genes is counterintuitive given its role as a co-repressor but it has nevertheless been observed before [298]. The co-occurrence of Rpd3L at promoter and Rpd3S on the ORFs of the same genes suggests that Rpd3L may play a role in targeting Rpd3S. One possibility would be that Rpd3S emerges from Rpd3L during the transition from initiation to elongation. The transition to elongation may trigger the exchange of subunits to transform Rpd3L into Rpd3S. Our finding that DSIF and RNAPII CTD phosphorylation are involved in the proper recruitment of Rpd3S to chromatin suggests that they may play a role in that process.

Unexpectedly, we also observe Rpd3L on the coding regions of the genes where Rpd3S is also present. This last result was surprising in that it does not agree with the generally accepted function of Rpd3L as a promoter-recruited co-repressor. These data argue for a new model where subunits of both Rpd3L and Rpd3S coexist on actively transcribed coding regions. According to the model described above, the presence of Rpd3L subunits may reflect an imperfect exchange of subunits during the transition from initiation to elongation. Alternatively, it may reflect the presence of an Rpd3 “super large” complex that contains subunits of both Rpd3S and Rpd3L. In such a model, Rpd3S (as we normally see it) would only exist in solution and would represent a module that joins Rpd3L after initiation. This hypothesis is supported by data provided by Collins et al. [62 /d] who combined and re-analyzed previously published mass spectrometry data [325, 326] to generate a high-accuracy yeast protein interaction dataset. They found that several subunits of Rpd3L can be co-purified using Rco1 as bait. Reciprocally, subunits of Rpd3S can be co-purified with Rpd3L subunits used as baits. These data argue that some interactions between Rpd3S and Rpd3L do exist in the cell. Since the complexes analyzed in these studies were purified from the soluble cellular fraction, the relative low abundance of these inter-complex interactions may be explained by the fact that these complexes normally exist together only on chromatin. The recent development of techniques to purify protein complexes from chromatin [327, 328] may help testing these intriguing possibilities.

3.1.5.3 The role of the Eaf3 CHD in Rpd3S function

We show that the recruitment of Rpd3S to active genes does not require H3K36 methylation, Set2 or the Eaf3 chromodomain. Interestingly, Rco1 binding is not abolished in a triple deletion mutant for all the known yeast HMTs (Set1, Set2 and Dot1), ruling out the possibility that the previously described interaction between H3K4me3 and Eaf3 [122] (or an eventual interaction with methylated H3K79) is compensating for the loss of H3K36

methylation in *set2Δ* and H3K36A mutants. Moreover, although the methylation state of H3K36 does not affect Rpd3S binding at highly transcribed genes, it has an effect on the occupancy of Rpd3S at ORFs showing lower transcription levels. This suggests that, despite not providing the main recruitment signal, H3K36 methylation provides some stabilizing effect on the association of Rpd3S with chromatin. While Rpd3S recruitment is largely independent on H3K36 methylation, the activity of Rpd3S, however, depends on the integrity of the Set2 / Rpd3S pathway, namely Set2-dependant H3K36 methylation, the Rco1 PHD and the Eaf3 CHD. We therefore propose that the main function of H3K36 methylation is to “activate” Rpd3S after it has been recruited by virtue of its association with the RNAPII CTD. This may involve the anchoring of Rpd3S on chromatin via the Eaf3-H3K36me interaction.

3.1.5.4 RNAPII phosphorylation by Kin28 and Ctk1 is required for the recruitment of Rpd3S

Our data clearly show that phosphorylation by Kin28 is absolutely required for the association of Rpd3S with the coding region of transcribed genes. This suggests that phosphorylation of serine 5 provides the signal for the association of Rpd3S with early elongating RNA polymerases II molecules. Since Kin28 was recently shown to phosphorylate serine 7 in addition to serine 5, we cannot rule out the possibility that serine 7 phosphorylation also contributes to the binding of Rpd3S to RNAPII. Using a *ctk1Δ* strain, we also show that phosphorylation of serine 2 also contributes -although to a lesser extent than serine 5- to the association of Rpd3S with transcribed genes. These results are in perfect agreement with a recent paper from the Hinnebusch group who showed that RNAPII CTD peptides harboring a phosphate group on both serines 5 and 2 have a high affinity for Rpd3S *in vitro* while a single phosphate on serine 5 has a reduced affinity [316]. More importantly, they were not able to detect any interaction of Rpd3S with non-phosphorylated CTD peptides, and peptides carrying a single phosphate on serine 2 barely

show any affinity with Rpd3S [316]. We therefore propose that Rpd3S is recruited to active genes via interaction with the phosphorylated RNAPII CTD. However, this complex remains inactive until it has been “anchored” on chromatin via H3K36 methylation. As discussed above, this phenomenon does not appear to occur equally on all genes. Negative regulation of these interactions by the DSIF elongation complex modulates association of Rpd3S with genes, therefore creating situations where active genes with high level of Spt4 carry much less Rpd3S than expected from their transcriptional level.

3.1.5.5 DSIF is a negative regulator of Rpd3S recruitment

Interestingly, we found that deletion of *SPT4*, a subunit of DSIF, has a profound effect on Rpd3S occupancy *in vivo*. In the absence of Spt4, Rpd3S associates with the genome in a manner that correlates better with transcription than in wild type cells. This suggests that DSIF is involved in regulating the amount of Rpd3S on transcribed genes. DSIF appears to prevent the association of Rpd3S to a subset of transcribed genes, and even at genes occupied by Rpd3S, DSIF also plays a role since it prevents the hyper-accumulation of the HDAC in the 3' end of the gene. How exactly an elongation factor can regulate the association of a HDAC with elongating RNAPII remains obscure but we envision several mechanisms by which it may operate: 1) Rpd3S may directly or indirectly interact with DSIF; 2) The association of DSIF with the elongation complex may prevent the association of the HDAC; 3) DSIF may modulate the speed of the elongation complex in a way that makes it less favorable for Rpd3S binding; 4) DSIF may also impinge on the phosphorylation of the RNAPII CTD. More complex mechanisms may also be envisioned. For instance, Pin1, a proline isomerase that may modify the RNAPII CTD, binds to both, phosphorylated DSIF [329] and Rpd3/Sin3 [330]. CTD isomerisation may therefore be involved in the regulation of the recruitment of Rpd3S.

3.1.5.6 Bur1 downregulates the activity of DSIF

In human and yeast, DSIF is regulated by phosphorylation of the CTD of its Spt5 subunit. In yeast, this phosphorylation is mediated by the Bur1 cyclin-dependent kinase [129, 130]. We therefore tested whether phosphorylation of Spt5 by Bur1 affects Rpd3S occupancy *in vivo*. As expected, both the catalytic inactivation of Bur1 and the removal of its substrate (by deleting the CTD of Spt5) have a dramatic impact on Rpd3S occupancy. Both these mutants indeed show a decrease in Rpd3S occupancy along genes. These data suggest that phosphorylation of Spt5 by Bur1 negatively regulate the activity of DSIF on Rpd3S recruitment. In addition, the inactivation of Bur1 also causes Rpd3S to redistribute to promoter regions, a phenomenon that we also observed in *set2Δ* cells. Since Bur1 phosphorylates Rad6, which is also required for H3K36 methylation by Set2, it appears likely that this redistribution in Bur1 mutants is due to its effect on H3K36 methylation via Rad6. Why a lack in H3K36 methylation leads to an association of Rpd3S with promoters remains unknown, but is in agreement with the previous observation that histone acetylation decreases at promoters in these mutants.

Taken together, our in-depth analyses of Rpd3S genomic occupancy have revealed several key insights about its recruitment to genes *in vivo*. Our study highlights a complex network of protein-protein interactions mediated by phosphorylation of several substrates by at least three kinases. Interestingly, there is previous evidence in the literature linking these kinases to the function of DSIF. First, P-TEFb (Cdk9) and Bur1 can phosphorylate DSIF [129, 130, 331]. Second, Kin28, Ctk1 and Bur1 exhibit synthetic genetic interactions with Spt4 and Spt5 [332]. And third, the recruitment of Bur1 is stimulated by Kin28 [317, 333]. Finally, both DSIF and Kin28 have been shown to stimulate the recruitment of the Paf1 complex to the elongation complex [130, 334]. A lot more work will be required before we completely understand the interplay between these factors and Rpd3S.

3.1.6 Materials and Methods

3.1.6.1 Strains, Cell Growth and Crosslinking Conditions

All strains were grown in 50mL of YPD to an OD₆₀₀ of 0.6-0.8 before crosslinking, unless otherwise indicated. For ChIPchip, most strains were crosslinked with 1% formaldehyde for 30 min at room temperature on a wheel. The Rco1-9myc strains were crosslinked with 1% paraformaldehyde for 30 min at room temperature followed by 90 min at 4°C on a wheel. ATP analog-sensitive strains were treated with 6 μM of NAPP1 for 15 minutes prior to crosslinking.

3.1.6.2 Chromatin Immunoprecipitation and Antibodies

ChIP experiments were performed as per [285], with minor modifications. For myc-tag ChIP, we used 5μg of 9E11 antibody coupled to 2x10⁷ pan-mouse IgG DynaBeads (Invitrogen) per sample. For histone H4K5 acetylation ChIP, we used 4μL of a rabbit serum (Upstate 07-327) coupled to 2x10⁷ protein G DynaBeads per sample. Histone H3K36me3 was immunoprecipitated with 4μL of antibody (Abcam Ab9050) coupled to 2x10⁷ protein G DynaBeads per sample. Histone H4 levels were assayed with 2μL of an antibody raised against recombinant yeast histone H4 (a gift from Alain Verreault) coupled to 2x10⁷ protein G DynaBeads per sample. RNAPII ChIPs were done using 2μL of 8WG16 antibody coupled to 2x10⁷ pan-mouse IgG DynaBeads per sample. Note that in our ChIP-chip assays, 8WG16 generates profiles that are nearly identical as using a tagged RNAPII (data not shown). 8WG16 is therefore used here to measure total RNAPII levels on genes. RNAPII CTD serine 2 phosphorylation was assayed using 5μL of H5 antibody (Covance Research MMS-129R-200) coupled to 2x10⁷ protein G DynaBeads per sample.

The microarrays used for location analysis were purchased from Agilent Technologies (Palo Alto, California, United States) and contain a total of 44,290 Tm-

adjusted 60-mer probes covering the entire genome for an average density of one probe every 275 bp (± 100 bp) within the probed regions (catalog # G4486A and G4493A). Myc-tag ChIPs were hybridized against ChIPs from isogenic strains that did not contain the tag as controls. Acetylation, RNAPII and histone H4 ChIPs were hybridized against a sample derived from 400ng of input (non-immunoprecipitated) DNA. Acetylation levels were normalized to histone H4 levels by subtracting $\text{Log}_2(\text{Histone H4} / \text{input})$ from $\text{Log}_2(\text{H4K5 acetyl}/\text{input})$.

3.1.6.3 Data Analysis

The data were normalized and biological replicates were combined using a weighted average method as described previously [285]. The log_2 ratio of each spot of combined datasets was then converted to Z-score, similar to Hogan et al. [335], to circumvent the large differences in the immunoprecipitation efficiencies of the different factors. Visual inspection of the Z-scores was carried out on the UCSC Genome Browser (<http://genome.ucsc.edu/>). All data analyses described here were done using data from protein-coding genes longer than or equal to 500 bp. Median Z-score values for promoter and complete length of each annotation from SGD (version Feb. 02 2008) were calculated without interpolation and used in our clustering and Pearson correlation analyses. Promoters are defined as the shortest of either 250bp or half the intergenic region (half-IG) relative to the reference gene's 5' boundary. Selforganizing map (SOM) clustering was done with the Cluster software [336] and visualized with Java Treeview [337]. Only genes with no missing value were used for clustering.

Gene mapping was performed as in Rufiange et al. [60] on selected groups of genes described in the text. Briefly, data were mapped onto the 5' and 3' boundaries in 50 bp windows for each half-gene and adjacent half-IG regions. A sliding window of 300 bp was then applied to the Z-scores to smooth the curve.

3.1.7 Acknowledgments

We are grateful to all members of our laboratory for helpful discussion. We thank Marlene Oeffinger, Célia Jeronimo and Sara Hardy for their critical reading of the manuscript. The antibody against recombinant histone H4 was kindly provided by Alain Verreault. We also warmly thank Chhabi K. Govind and Alan G. Hinnebusch for sharing unpublished data.

SD did most of the experiments and analyses and wrote the paper with FR. LL generated some yeast strains and performed some ChIP experiments. PEJ performed some data analyses, helped with preparation of the figures and contributed to the writing of the paper. AF performed some ChIP experiments. FR wrote the paper with SD and the contribution of PEJ.

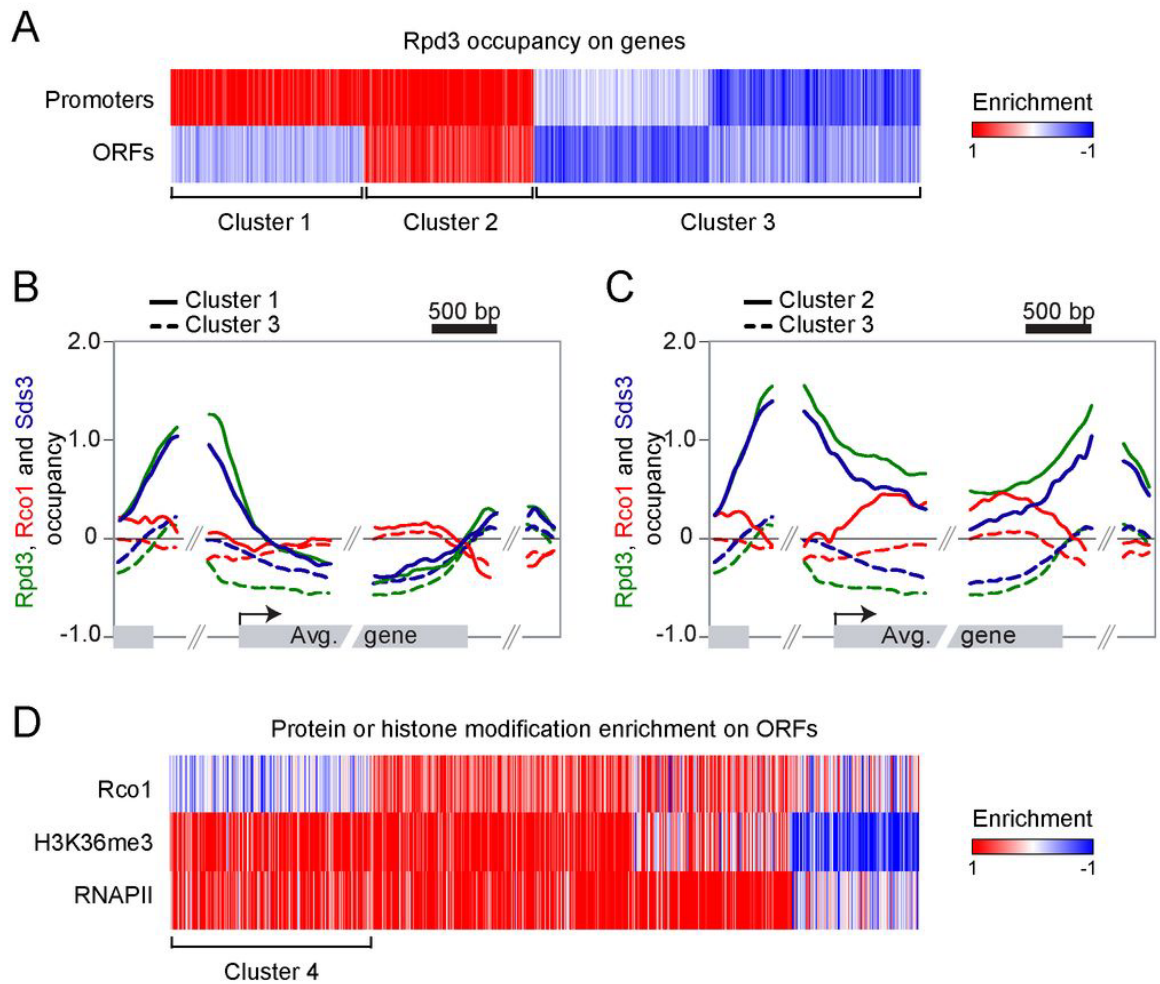


Figure 3-1 : Rpd3S and Rpd3L target overlapping sets of genes.

(A) Self-organizing map (SOM) clustering of Rpd3 binding on promoter and ORFs of the 3693 genes with no missing value. Red indicates enriched (bound) regions and blue represents depleted regions.

(B-C) Mapping of Rpd3 (green), Rco1 (red) and Sds3 (blue) on the 954 genes from cluster 1 (solid line, genes where Rpd3 is present both on ORFs and promoters) and on the 1906 genes from cluster 3 (dashed lines, genes where Rpd3 is not bound) (B) and on the 833 genes from cluster 2 (solid lines, genes where Rpd3 is present only on promoters) and cluster 3 (C).

(D) SOM clustering of Rco1, H3K36me3 and RNAPII on genes from the cluster 2 from panel “A”. Cluster 4 represents the 222 active genes with no Rco1 binding.

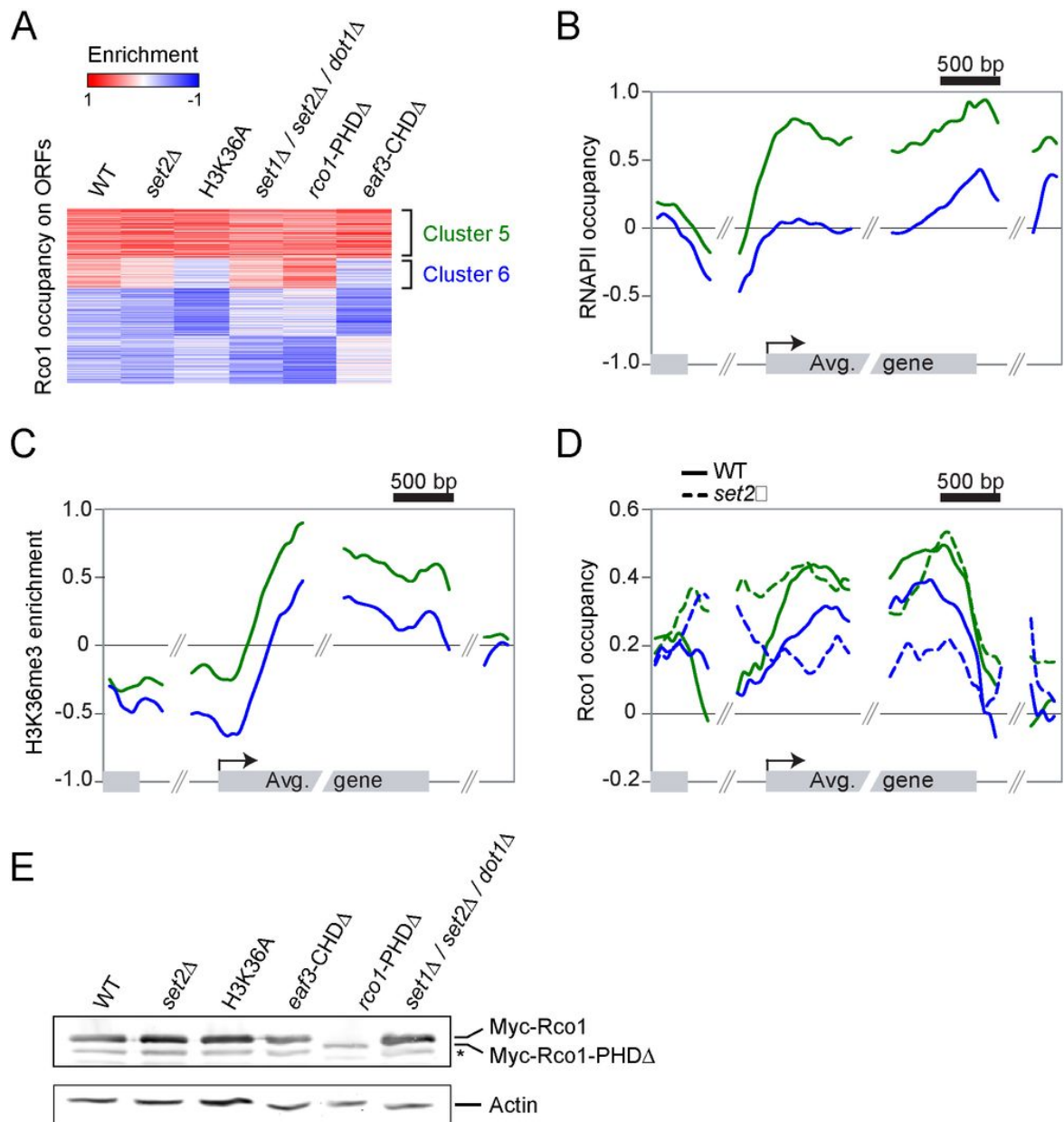


Figure 3-2 : The recruitment of Rpd3S to transcribed genes does not require Set2-dependent H3K36 methylation in vivo.

(A) SOM clustering of Rco1 enrichment on ORFs of all the 5007 genes with no missing value in WT and various mutants.

(B-C) Mapping of RNAPII occupancy (B) and H3K36me3 enrichment (C) on genes contained within clusters 5 (green, 1425 genes) and 6 (blue, 841 genes) from panel “A”.

(D) Mapping of Rco1 in WT (solid lines) and *set2* Δ (dashed lines) cells along genes contained within clusters 5 (green) and 6 (blue) from panel “A”.

(E) Western blot showing the Rco1 protein levels in the strains used in panels A-D. Note the decreased level of the PHD-truncated Rco1 protein. The star symbol (*) indicates a degradation product.

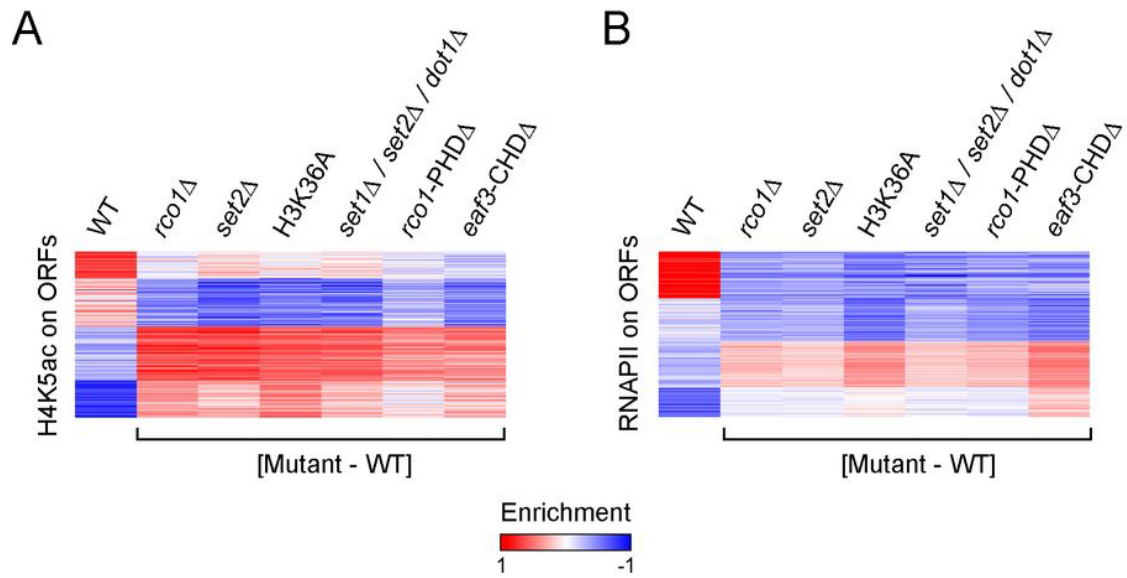


Figure 3-3 : Set2-dependent H3K36 methylation is required for the function of Rpd3S. (A-B) SOM clustering of H4K5ac enrichment (A) and RNAPII occupancy (B) on ORFs of all genes in WT cells, along with the difference calculated between the enrichment observed in various mutants and WT cells. The *rpd3*-H188A mutant is a catalytic-dead version of Rpd3.

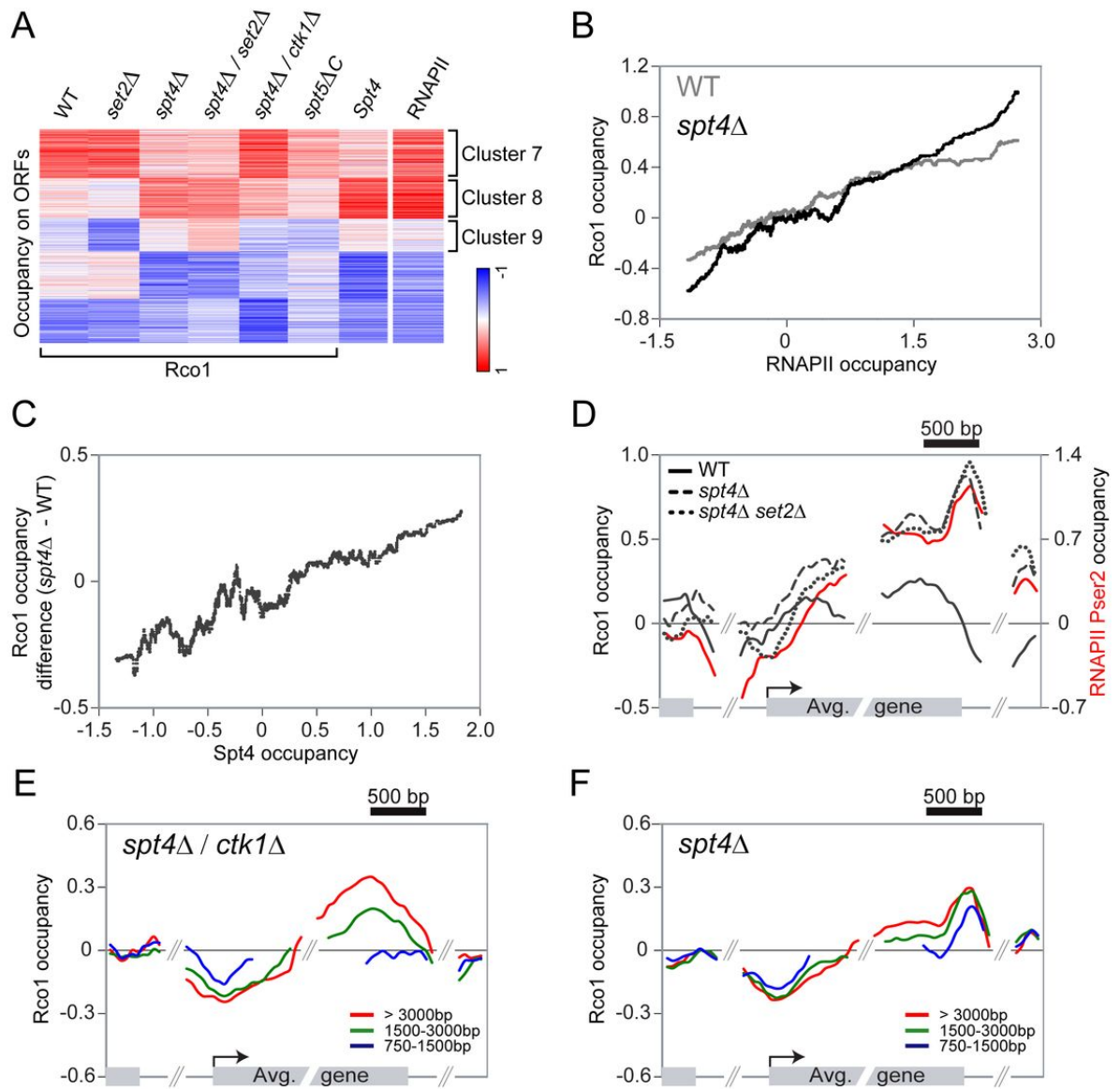


Figure 3-4 : Spt4 negatively regulates the recruitment of Rpd3S

(A) SOM clustering of Rco1 occupancy on ORFs of all genes in WT, and various mutants, together with Spt4 occupancy in WT on ORFs. The RNAPII ORF occupancy in WT was added after the clustering.

(B) Correlation between Rco1 occupancy and RNAPII occupancy in WT (grey) and *spt4Δ* cells (black). A sliding median window of 300 genes was applied to the data.

(C) Correlation between Spt4 occupancy and the difference of Rco1 occupancy in *spt4Δ* versus WT cells. A sliding median window of 300 genes was applied to the data.

(D) Mapping of Rco1 occupancy on the 953 genes contained within cluster 8 in WT (black solid line), *spt4Δ* (dashed line) and *spt4Δ / set2Δ* (dotted line) cells. The RNAPII phosphorylated on serine 2 enrichment is also shown (red solid line).

(E-F) Mapping of Rco1 occupancy on genes grouped according to their lengths in *spt4Δ / ctk1Δ* (E) and *spt4Δ / set2Δ* (F) cells (red line: the 465 genes >3000bp, green line: the 1703 genes between 1500-3000bp, blue line: the 2091 genes between 750-1500bp).

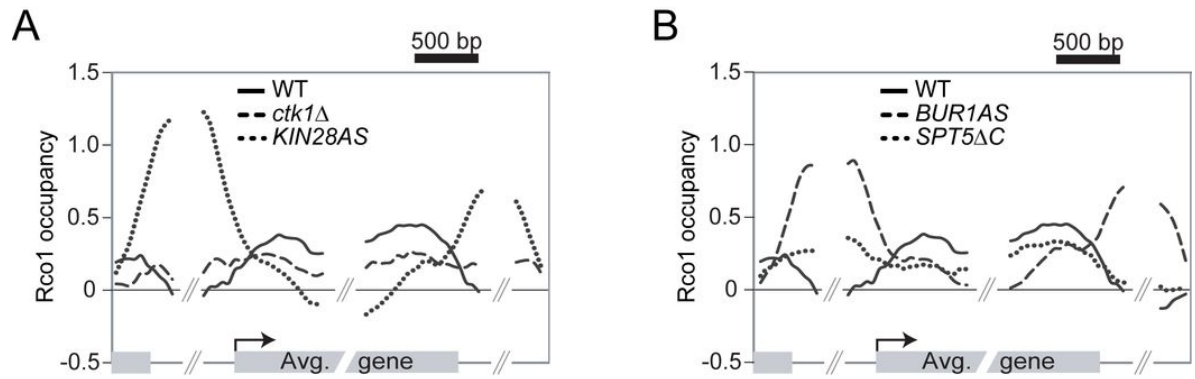


Figure 3-5 : Phosphorylation of the RNAPII and Spt5 CTD are important for the Rpd3S recruitment on genes.

(A) Mapping of Rco1 occupancy on the 2089 genes contained within clusters 7 and 8 in WT (solid line), *ctk1Δ* (dashed line), and *kin28AS* (dotted line) cells.

(B) As in 'A' using the *bur1AS* (dashed line) and *spt5ΔC* (dotted line) cells.

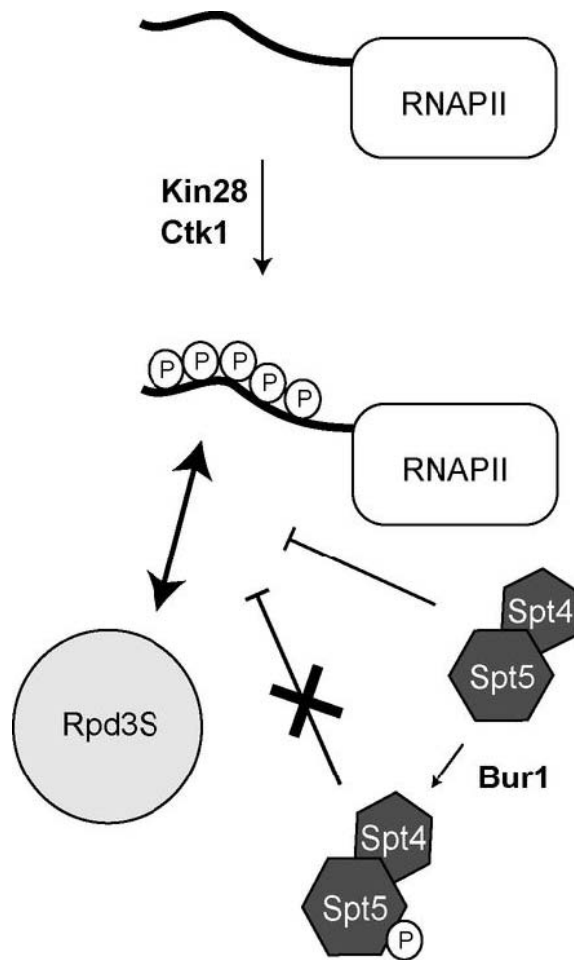


Figure 3-6 : A model summarizing how Rpd3S is recruited to active genes.

Phosphorylation of the RNAPII CTD by Kin28 and Ctk1 stimulates the recruitment of Rpd3S to the CTD. DSIF (Spt4/Spt5) counteracts the recruitment of Rpd3S but its phosphorylation by Bur1 alleviates its repressing activity.

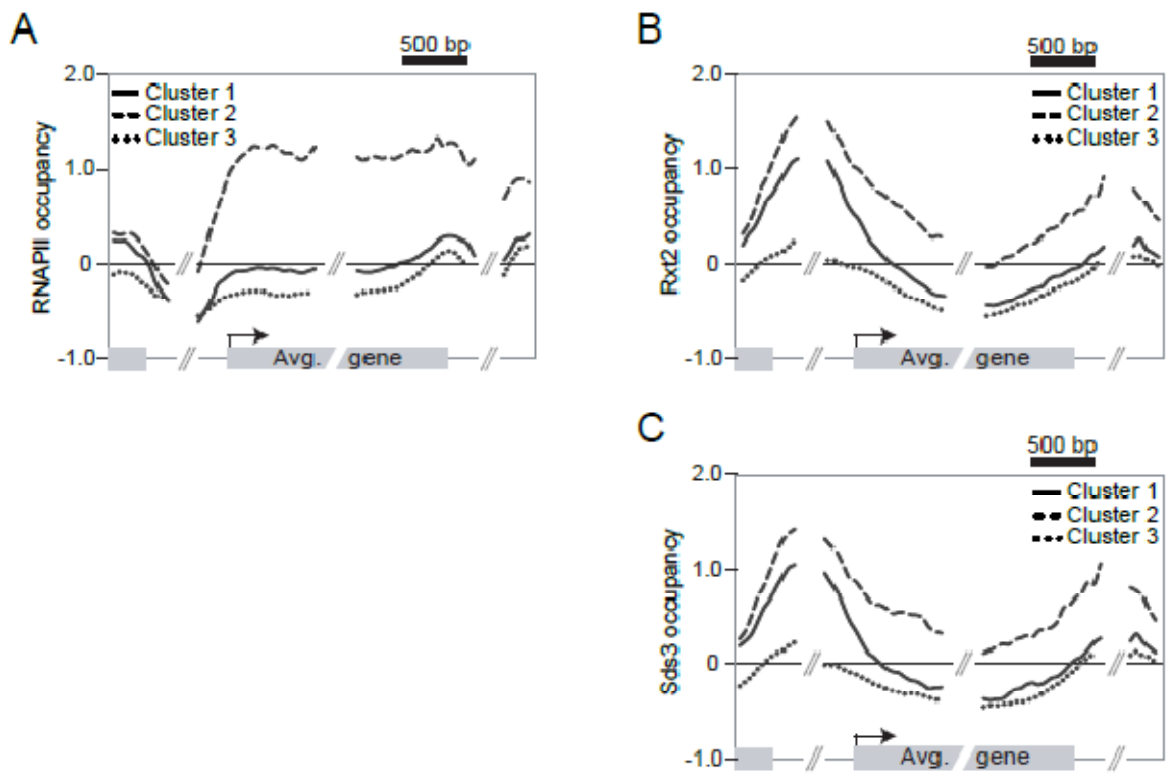


Figure 3-7 : A complement to Figure 3-1. Average signal of RNAPII, Rxt2 and Sds3 over the genes of the 3 clusters.

Mapping of RNAPII (A), Rxt2 (B) and Sds3 (C) occupancy on genes contained within clusters 1 (solid line, 954 genes), 2 (dashed line, 833 genes) and 3 (dotted line, 1906 genes).

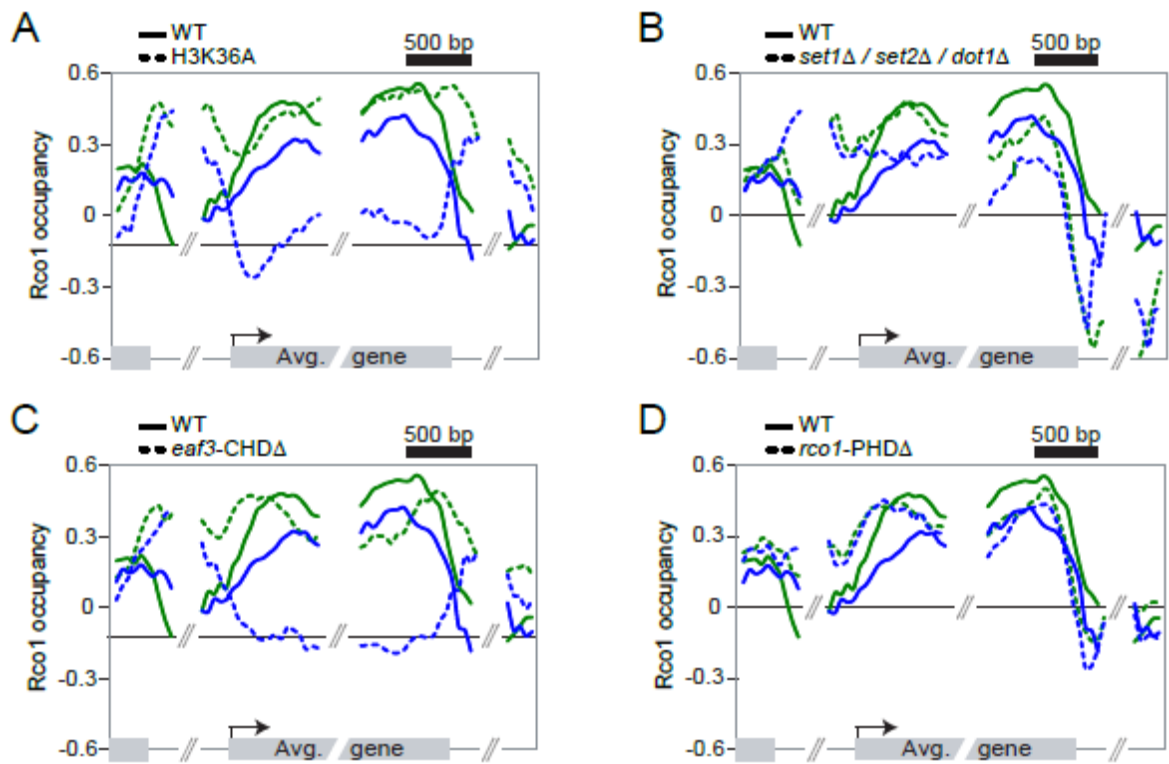


Figure 3-8 : A complement to Figure 3-2. The recruitment of Rpd3S to transcribed genes does not require Set2-dependent H3K36 methylation in vivo.

(A-D) Mapping of Rco1 occupancy on genes contained within clusters 5 (green, 1425 genes) and 6 (blue, 841 genes) from **Figure 3-2** in WT and various mutants.

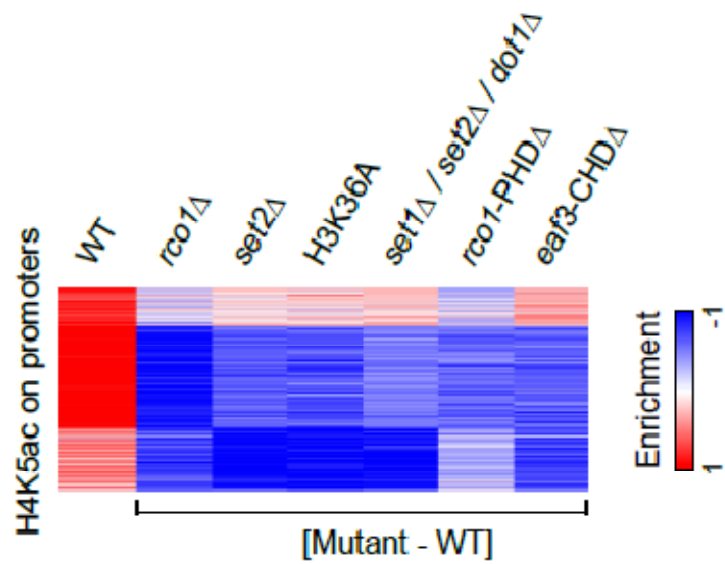


Figure 3-9 : A complement to Figure 3-3. Deletion of *SET2* or disruption of Rpd3S causes a general decrease in histone acetylation at promoters.

SOM clustering of the enrichment of H4K5ac on promoters of all genes in WT cells along with the difference calculated between the enrichment observed in WT and various mutants.

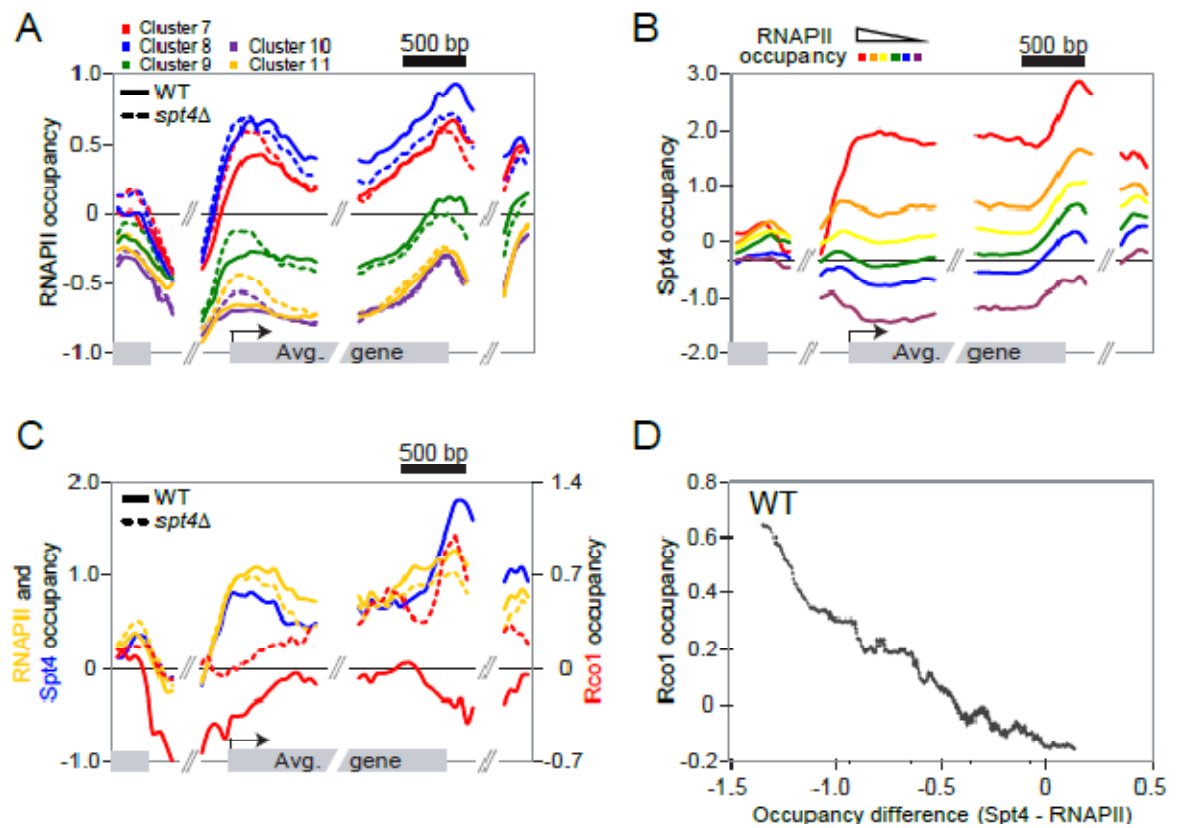


Figure 3-10 : A complement to Figure 3-4. Spt4 negatively regulates the recruitment of Rpd3S.

(A) Mapping of RNAPII occupancy in WT (solid lines) and *spt4Δ* (dashed lines) cells on genes contained within clusters 7 (red, 1136 genes), 8 (blue, 953 genes), 9 (green, 771 genes), and the non-identified following clusters 10 (purple, 1031 genes) and 11 (gold, 1109 genes) of the **Figure 3-4**.

(B) Mapping of Spt4 occupancy on genes binned by their RNAPII occupancy.

(C) Mapping of RNAPII (gold), Spt4 (blue) and Rco1 (red) occupancy in WT (solid line) and *spt4Δ* (dashed line) cells on the 222 genes contained within clusters 4 from **Figure 3-1D**.

(D) Anti-correlation between Rco1 occupancy and the difference between Spt4 and RNAPII occupancy in WT cells measured on the 2286 transcribed genes (RNAPII > 0). A sliding median window of 300 genes was applied to the data.

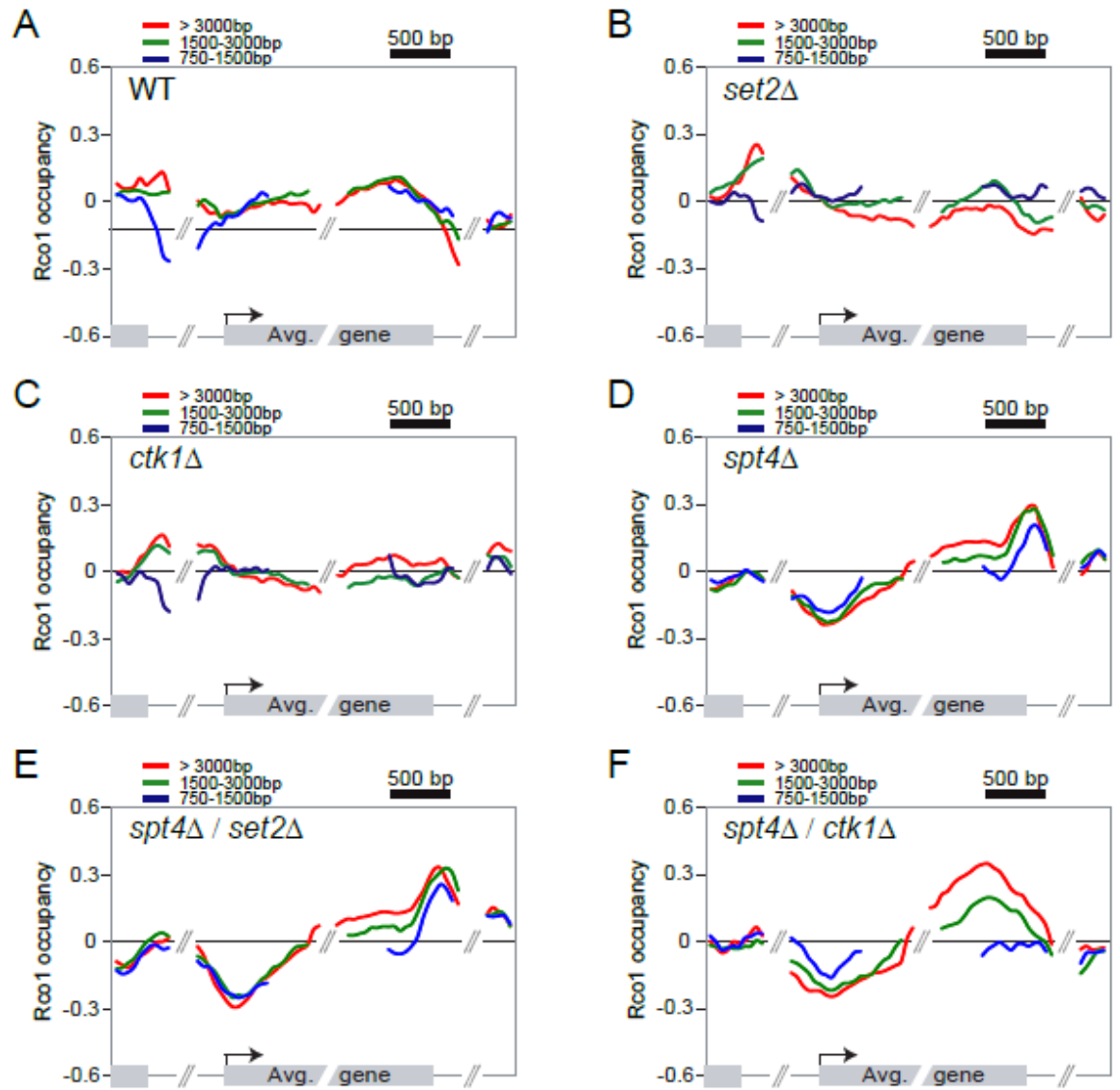


Figure 3-11 : Rpd3S ORF occupancy level is gene length-dependant in a *spt4Δ / ctk1Δ* mutant.

(A-D) Mapping of Rco1 occupancy on genes grouped according to their lengths in WT (A), *set2Δ* (B), *ctk1Δ* (C), *spt4Δ* (D), *spt4Δ / set2Δ* (E), and *spt4Δ / ctk1Δ* (F) cells (red line : the 465 genes >3000bp, green line: the 1703 genes between 1500-3000bp, blue line: the 2091 genes between 750-1500bp).

4 Détermination génomique des caractéristiques de liaison du complexe Set3C et de sa sous-unité Hst1

Dans la section précédente, j'ai étudié en profondeur les mécanismes régissant le recrutement et la fonction du complexe HDAC Rpd3S aux régions codantes transcriptionnellement actives. Rpd3S est important au rétablissement de la structure prétranscriptionnelle de la chromatine en déacétylant les nucléosomes présents sur ces régions. La perte de cette fonction mène à l'initiation opportuniste de la transcription au sein des régions codantes, ce qui génère des transcrits inutiles et potentiellement néfastes pour la cellule. J'ai démontré que bien que la fonction de Rpd3S sur l'ORF dépend de la méthylation du résidu H3K36 par l'histone méthyltransférase Set2, son recrutement, contrairement au modèle accepté dans la littérature, n'en dépend pas mais passe plutôt par une interaction avec le domaine C-terminal hyperphosphorylé de la sous-unité Rpb1 du complexe RNAPII.

Toutefois, Rpd3S n'est pas le seul complexe HDAC connu interagissant avec les régions codantes. Le groupe de Kim *et coll.* a démontré que le complexe Set3C pouvait lui aussi interagir avec des résidus méthyle-lysines, en l'occurrence H3K4 di- et triméthylé [140]. La perte de cette marque de méthylation ou du domaine PHD de la protéine Set3 donne une hyperacétylation aberrante des nucléosomes en 5' des régions codantes transcrites. La fonction de Set3C serait donc de déacétyler ces nucléosomes par le biais des deux HDACs qui en font partie : Hst1 et Hos2.

La HDAC Hos2 est nécessaire à l'activation des gènes *GAL* et *INO1* en situation de stress [139], et sa délétion cause la majeure partie de l'effet observé par Kim. Toutefois, bien que contribuant à cet effet, une délétion de Hst1 donne un phénotype différent des délétions des sous-unités de Set3C. Ce phénomène est probablement dû au rôle que joue Hst1 dans la répression de la transcription de plusieurs gènes. En effet, Hst1 interagit aussi avec les corépresseurs Sum1 et Rfm1 pour réprimer l'expression des gènes de synthèse du NAD⁺ [179] et de sporulation [178].

Hst1 interagit donc ponctuellement avec Set3C, mais a aussi une fonction complètement indépendante de ce complexe. Cette hypothèse, bien que supportée par des données génétiques, n'a jamais vraiment été confirmée. Robert et ses collègues [36] ont bien déterminé que Hst1 et Sum1 liaient les mêmes promoteurs, mais n'ont pas détecté Hst1 sur les régions codantes ou étudié Hos2.

L'étude présentée dans cette section se penche sur ces sujets. J'ai d'abord caractérisé à l'échelle génomique les sites de liaison de Hst1 et de Set3 et observé que leurs localisations sont généralement mutuellement exclusives : Hst1 lie surtout les promoteurs, probablement en complexe avec Sum1 ou Rfm1, tandis que Set3 lie seulement les régions codantes où l'on retrouve H3K4 di- ou triméthylé. Ces deux protéines sont toutefois présentes aux ORFs ayant une très forte activité transcriptionnelle. De plus, l'abolition de la méthylation de H3K4 n'affecte pas la liaison de Hst1 ou de Set3 à ces gènes, même si elle abolit la liaison de Set3 au reste de ses cibles.

La détermination des sites de liaison de Hos2 dans le génome a causé une grande surprise : cette HDAC, que l'on croyait jusqu'alors être une partenaire quasi indispensable de Set3, n'est présente qu'aux régions codantes les plus enrichies en Set3 et surtout où Hst1 est aussi présent sur ces régions. De plus, la localisation de Hos2 semble être restreinte aux gènes où l'on observe un niveau important de triméthylation en H3K4, mais pas où ce résidu est seulement diméthylé; l'importance biologique de cette observation reste toutefois à être déterminée.

Finalement, j'observe une redistribution de Set3 des régions codantes aux promoteurs dans les cellules déficientes en méthylation de H3K4, ce qui n'est pas du tout le cas dans les cellules de type sauvage. Cette redistribution corrèle très bien avec l'activité transcriptionnelle des gènes affectés. Ces données enrichissent le modèle proposé par Kim [140] sans le contredire de manière significative : la di- ou triméthylation de H3K4 est en effet essentielle à la liaison de Set3 à la plupart des régions codantes sauf celles transcriptionnellement très actives. Le mode de recrutement de Set3C sur ces régions reste

à caractériser. Une première piste prometteuse réside en un modèle d'interaction de Set3C au domaine C-terminal hyperphosphorylé de Rpb1, comme suggéré par Govind [316].

On sait depuis 2008 que plusieurs membres de Set3C interagissent avec le complexe Rpd3L [190]. Nos données, présentées au chapitre 3, indiquent que Rpd3L joue surtout un rôle corépressif aux promoteurs, mais aussi qu'il est présent sur les régions codantes transcriptionnellement très actives. On ne connaît pas la fonction de Rpd3L sur ces régions, mais ce sont les mêmes où l'on retrouve aussi Hst1, Hos2 et Set3. Cette colocalisation n'implique pas forcément une interaction ou une fonction commune, mais un modèle où ces deux complexes interagiraient est naturellement très attrayant et mérite une attention particulière.

4.1 Manuscrit :

**« Histone H3 Lysine 4 Methylation and RNAPII CTD
Phosphorylation Recruit Distinct Set3 Complexes to Different
Sets of Genes. »**

Simon Drouin and François Robert

Contribution des coauteurs :

S. Drouin: Design, exécution et analyse expérimentale. Rédaction conjointe avec F. Robert.

F. Robert: Rédaction conjointe avec S. Drouin.

4.1.1 Abstract

The Set3C complex contains two histone deacetylases (HDACs), Hst1 and Hos2, and has been characterized both as a repressor of meiosis and as being required for the activation of the GAL and INO1 genes. It has been previously shown that H3K4 dimethylation is sufficient to recruit Set3C to actively transcribed coding regions to enable Hst1- and Hos2-dependant nucleosomes deacetylation at specific genes. More recently, this model was challenged by work showing that RNAPII CTD phosphorylation triggers the recruitment of Set3C to active genes independently of histone H3 methylation. In addition to its function as part of Set3C, Hst1 is also known as a promoter-bound repressor of sporulation genes; the function of Hst1 in Set3C is therefore still under debate. Here, we show by genome-wide ChIP-chip that Hst1 and Set3 binding are mostly mutually exclusive except at the most transcribed genes. At these highly transcribed genes, Set3 and Hst1 occupy the coding region in a H3K4 methylation-independent manner. At most of its target genes, however, Set3 binds in the absence of Hst1 and its recruitment is strictly dependent on histone H3 di- or trimethylation. Surprisingly, we find that Hos2, thought to be an essential part of the Set3C complex, only binds the coding regions of the most transcribed genes where Hst1 and Set3 are present and not ORFs where Set3 binding requires H3K4 methylation. These results imply that two forms of the Set3 complex exist in the cell: a Set3-Hst1-Hos2 variant and a Set3-only variant. Finally, our data also show that these two complexes are recruited to chromatin by distinct mechanisms.

4.1.2 Introduction

Nucleosomes constitute the principal barrier to efficient transcription. Cells have evolved a quantity of means to circumvent this problem. One such solution involves the weakening of DNA-nucleosome electrostatic interactions through histone acetylation. However, hyperacetylated coding regions become too permissive and exhibit transcription

events from cryptic sites. We and others have shown that the Rpd3S complex is involved in restoring the neutral pre-transcription state of chromatin on the coding regions through its deacetylase activity [289]. However, Rpd3S is not the only HDAC to target actively transcribed coding regions. The Set3C complex has been shown to be recruited to these regions and affect acetylation levels, particularly in their 5' end.

The Set3C complex contains two histone deacetylases, Hst1 and Hos2 [136], and has been characterized both as a repressor of meiosis and as required for the activation of *GAL* genes and *INO1* [139]. The Set3C-dependant repression of meiosis genes is Hst1 independent. Furthermore, Hst1, along with its partner Sum1, represses middle-sporulation genes in a Set3C-independent manner.

Kim *et al.* have determined that Set3C-associated Hst1 and Hos2 play a role in deacetylating nucleosomes present in the 5' portion of ORFs, this activity being dependent on the binding of Set3 to di- or tri-methylated H3K4 [140]. However, it is unknown whether this mechanism is present at all genes or specific to a subset. In addition, a recent report from the Hinnebusch lab showed that the phosphorylation of the RNAPII CTD triggers the recruitment of Set3 to the *ARG1* gene [316]. Collectively, these data suggest that multiple forms of the Set3C may exist in the cell; fulfilling different functions, and being recruited by different mechanisms.

In order to solve the ambiguity about Set3- and Hst1-containing complexes, as well to better understand how they are recruited to genes, we have endeavored to characterize the binding pattern of Hst1, Set3 and Hos2.

We show that Hst1 and Set3 binding pattern are mostly mutually exclusive even though they have been shown to physically interact [136]. We also confirm that H3K4 di- or trimethylation is important for Set3C recruitment to most of its targets. However, it is dispensable for Set3 binding to the most actively transcribed regions. Curiously, we observe the presence of Hst1 along with Set3 and Hos2 on these highly transcribed regions, suggesting that the Hst1-Set3 interactions that have been described in the literature are

probably limited to these genes. Curiously, Hos2 is absent from coding regions where Set3 binds in a H3K4 methyl-dependant fashion, implying a HDAC-less Set3C variant. Finally we observe that Set3 is abundant on the promoter region of genes in a RNAPII abundance correlated fashion in cells lacking H3K4 methylation. Set3 is normally absent from promoters in wild type cells.

4.1.3 Results

4.1.3.1 Hst1, Set3 and Hos2 bind functionally distinct regions

We studied the genome-wide occupancy of myc-tagged Hst1, Set3 and Hos2 by ChIP-chip using a protocol adapted from Ren *et al.* [285]. As seen in **Figure 4-1A**, Hst1 is present at the promoter of 40% of all yeast genes, half of which also show significant Hst1 occupancy on the coding regions (**Figure 4-1A**, cluster 2 and **Figure 4-1B**). It is interesting to note that genes where only the promoter is bound by Hst1 are enriched in stress-response genes, including those previously found to be regulated by Hst1-Sum1 such as the NAD⁺ biosynthesis pathway genes [179] and meiosis [178]. Further confirming this hypothesis, we find that genes where Hst1 is bound at the promoter but not at the ORF and are also bound by Sum1 (**Figure 4-4A**).

We do not observe any binding of Set3 on the promoter regions (**Figure 4-1C**). Furthermore, Set3 ORF binding and Hst1 promoter association are mutually exclusive except at genes where Hst1 is also present on the coding region.

Surprisingly, we find that Hos2 does not bind all the genes that Set3 does: Hos2 binding is only observed on coding regions where Set3 abundance is highest and strongest where both Set3 and Hst1 are present (**Figure 4-1A**, **Figure 4-4D** and **E**). Furthermore, Hos2 occupancy is restricted to coding regions where H3K4 trimethylation is present

(**Figure 4-1A**; compare columns Hos2 (O) and H3K4Me3); whether this is significant has yet to be examined.

These results show that Hst1 is present in at least two forms in the cell: a Set3C-independent form present exclusively at gene promoters, most likely involved in transcription repression, and a Set3C-correlated form present on the same ORFs as Set3 and Hos2. Similarly, Set3C is also present in at least three forms: one that binds ORFs where Hst1 and Hos2 are both present, another that binds the same ORFs as Hos2 and a last that is Hst1- and Hos2-independent.

It has been previously shown that Set3C promotes transcription since mutants of this complex, including *set3Δ* and *hos2Δ*, display mycophenolic acid sensitivity, an elongation inhibitor [140]. Curiously, Set3 occupancy does not follow RNAPII abundance except for cluster 2, where Hst1 and Hos2 are also bound at the ORF (**Figure 4-1A**, RNAPII column). However that H3K4 di- and trimethylation marks, thought to recruit Set3C through the Set3 PHD domain [140], are present in significant levels at most Set3-bound genes (**Figure 4-1A**, H3K4Me2 and H3K4Me3 columns).

4.1.3.2 Set3 binding is severely affected by the loss of H3K4 methylation

Kim and colleagues have shown that H3K4 dimethylation by the Set1 histone methyltransferase is sufficient to recruit Set3C at coding regions [140]. More recently, this notion was challenged by the Hinnebusch group, who showed that Set3 can be recruited to genes via an interaction with the phosphorylated CTD of RNAPII. In both cases, however, the recruitment was only tested at a few genes. We were therefore curious to see the global effects of a *set1Δ* mutation, which abolishes H3K4 methylation, on Hst1, Set3 and Hos2 recruitment. We also tested a *spt4Δ* mutant, which still retains monomethyl H3K4 while losing di- and trimethylation, for Hst1 and Set3 binding. Spt4 is part of the yeast DSIF

complex, required for PAF complex recruitment and H2B monoubiquitination by Rad6. Rad6 is required for processing monomethylated H3K4 into its di- and trimethylated forms.

Hst1 promoter occupancy is not affected by a complete loss of H3K4 methylation (**Figure 4-2A**). A *spt4Δ* H3K4 monomethylation-only competent strain does not exhibit any binding pattern alterations either (**Figure 4-4B**). These results are not unexpected, although a bit surprising as the Hst1 ORF-bound fraction probably interacts with Set3C and thus should be affected by the loss of H3K4 methylation according to Kim et al.

Set3 binding, on the other hand, is severely affected by the loss of H3K4 methylation caused by the deletion of *set1Δ* (**Figure 4-2A**) and the presence of monomethyl H3K4 does not rescue this phenotype (**Figure 4-4C**), supporting the data from Kim *et al.* [140]. We observe a complete loss of Set3 binding in these mutants for most of the Set3 targets, indicating that H3K4 needs to be di- or trimethylated to recruit Set3 at these genes. However, our data also show that Set3 binds as well as or better than the wild-type on the ORF of a large fraction of genes in these mutants (**Figure 4-2A and 4-2C**). Interestingly, the genes where Set3 binding is unaffected in the *set1Δ* and *spt4Δ* mutants are the same where Hst1 is bound at the coding regions and also where RNAPII levels are highest (compare the blue line representing cluster 6 from **Figure 4-2A** in panels **B-E**). The binding pattern of Set3 (shown in **Figure 4-2C**) also changes quite drastically: Set3 levels in clusters 5 and 7 (red and green line, respectively) fall down to near-background levels, whereas cluster 6 (blue line) shows increased binding at the 5' end and, to our great surprise, enrichment at the promoter region (**Figure 4-2C**).

It will be very interesting to see if Hos2 occupancy is affected by the loss of H3K4 methylation as Set3, which would be the case if Set3 recruits Hos2 to H3K4 methylated genes through its PHD domain, or if it is unaffected by this loss like Hst1, implying that it is recruited to these ORFs with Set3 and Hst1 by an as yet uncharacterized mechanism.

These data clearly show that although H3K4 di- or trimethylation is required for Set3 binding at most coding regions, it is not required for Set3 and Hst1 recruitment at

genes with very high RNAPII occupancy. Note that RNAPII levels do not change significantly in the *spt4Δ* mutant when compared to the wild type (**Figure 4-2D**), indicating that the loss of H3K4 di- and trimethylation does not affect transcriptional activity as assayed by RNAPII abundance. The binding pattern changes observed for Set3 are therefore unrelated to changes in transcription activity.

4.1.3.3 Set3 binding mirrors RNAPII occupancy in *set1Δ* and *spt4Δ* mutants

The data shown above indicates that Set3, upon loss of H3K4 di- or trimethylation, redistributes to a subset of its target genes, namely those with both Hst1 ORF occupancy and the highest RNAPII levels. There also appears to be an enrichment of Set3 at these genes' promoters in these mutants whereas in the wild type strain Set3 does not bind promoters (**Figure 4-2C**).

We tested whether the distribution of Set3 is correlated to transcriptional activity as assayed by RNAPII abundance. Interestingly, Set3 and RNAPII enrichment on coding regions better correspond in the *set1Δ* and *spt4Δ* mutants than in the wild type (**Figure 4-3B**, compare blue line (WT) to red and green lines (*set1Δ* and *spt4Δ*, respectively)).

The behavior of Set3 on promoter regions is by far most surprising result, with Set3 accumulation on promoters in the *set1Δ* and *spt4Δ* mutants at levels closely mirroring RNAPII abundance (**Figure 4-3A**, red and green line, respectively).

These results indicate that H3K4 di- or trimethylation is important to recruit and sequester Set3 at genes in a transcription-level independent manner. In the absence of the methyl mark Set3 naturally associates with gene promoters and coding regions in a way closely related to their transcriptional level, as assayed by RNAPII abundance. The question remains as to what recruits Set3 to these genes in the absence of H3K4 methylation.

4.1.4 Discussion

The yeast HDAC Hst1 was best characterized in complex with the co-repressors Sum1 and Rfm1. These co-repressors recruit Hst1 to the promoter of specific genes to deacetylate nucleosomes and repress the transcription of said genes. This mechanism is required for the repression of NAD⁺ synthesis genes [338] and middle sporulation regulators [178, 179].

The function of Set3 is less well defined. The Set3C complex, which contains both Hst1 and Hos2 HDACs, is known to associate preferentially with highly transcribed genes. Set3C, along with Hos2, are required for the activation of the *GAL* genes and *INO1* in carbon-source restricted media [139]. Set3C is also involved in the Hst1-independent repression of early-middle sporulation genes [136]. The function of Hst1 in Set3C remains therefore unclear.

A recent study by Kim *et al.* [140] demonstrates that Set3 interacts with di- and trimethylated H3K4 to recruit the Set3C complex to the 5' regions of transcribed ORFs. This recruitment leads to the deacetylation of nucleosomes present on said ORFs by Hst1 and Hos2. We were curious to see if this hypothesis holds for all yeast genes, and also whether the repressive Hst1/Sum1/Rfm1 forms of Hst1 are involved in this mechanism.

We found that while Hst1 is present on the promoters of about 40% of all yeast genes and that Set3 exclusively binds coding regions. Curiously, Hst1 promoter binding and Set3 ORF binding are mostly mutually exclusive except at a subset of genes where Hst1 is present both at the promoter and the ORF and Set3 is present on the ORF. Surprisingly, we find that Hos2 and Set3 binding patterns only partially overlap: Hos2 binding is limited to coding regions with higher Set3 occupancy and strongest where Hst1 is also present on the ORF.

We see, similar to Kim *et al.*, that H3K4 di- or trimethylation is required for the binding of Set3 to most of its target regions. However, contrary to them, we find that about

one third of the Set3 targets do not require H3K4 methylation for proper recruitment. Finally, mutants lacking H3K4 methylation or possessing only H3K4 monomethylation capability show a transcription-rate dependant redistribution of Set3 to coding regions but also to promoters.

These observations raise many questions which we will endeavor to answer.

4.1.4.1 Hst1, Hos2 and Set3 interaction is limited to a small subset of genes

The HDAC Hst1 exists in at least two forms in the cell: as a promoter-bound repressor when complexed with Sum1 and/or Rfm1 [136, 179], and as a transcription-dependant deacetylase as part of Set3C [140].

We have shown that Hst1 binds about 40% of yeast gene promoters, amongst which we find previously characterized targets of the Hst-Sum1 complex such as middle sporulation regulators and NAD⁺ biosynthesis genes. Furthermore, we also find Hst1 bound to coding regions, but only for genes on whose promoters Hst1 is also bound; there are no genes bound exclusively at the ORF by Hst1. Finally, we found that all promoters where Hst1 binds exclusively on the ORF were also bound by Sum1, supporting the corepressive function of Hst1 (**Figure 4-4A**).

Set3 binding, on the other hand, is restricted to the coding region of 60% of yeast genes. These genes are usually H3K4 di- or trimethylated (**Figure 4-1A**). Curiously, Set3 is mostly absent at genes where Hst1 is found exclusively at promoters (**Figure 4-1A**, cluster 1) but highly enriched where Hst1 is at the ORF (**Figure 4-1A**, cluster 2). Interestingly, coding regions bound by Hst1 and Set3 are transcribed at a very high level, as assayed by RNAPII abundance (**Figure 4-1A** and **4-2F**) and show elevated H3K4 trimethyl levels (**Figure 4-1A**).

The Hos2 binding pattern we observe is very surprising, considering that it has been considered an essential part of the Set3C complex for years. We observe Hos2 enrichment

only at H3K4 trimethylated genes but not H3K4 dimethylated genes (**Figure 4-1A**). These genes are all very highly transcribed and mostly have Hst1 also present on the ORF. Note that Hos2 only binds genes where Set3 is abundant (**Figure 4-4**, panels **C-E**). These results clearly show that Hst1, Hos2 and Set3 binding profiles overlap only at a small subset of genes representing about 50% of the Hst1 targets, most of Hos2's or 33% of Set3's. This relatively limited overlap between Set3 and Hst1 fits the $\sim 1/4$ stoichiometry observed for Hst1 in the Set3C complex as opposed to 1/1 for the other members of this complex [136]. Hst1 therefore does not appear to be a full-time member of Set3C. This hypothesis is supported by studies where Set3C subunits were deleted: Hst1 mutants in these studies either have no effect or a phenotype different than for those of deletions for the other subunits of this complex [136, 140, 339, 340]. The partial overlap between Set3 and Hos2 is more puzzling, as many studies have observed Hos2 and Set3 genetically and physically interacting. However, this does not exclude that this apparently HDAC-less Set3 variant simply never has been observed and that all the interactions and phenotypes reported thus far are characteristic of the “complete” Set3C complex present on the most highly transcribed genes.

4.1.4.2 H3K4 di- or trimethylation is not required for Set3 binding at highly transcribed genes

Set3 contains a methyl-lysine PHD domain known to interact with methylated H3K4 *in vitro* [254]. Kim *et al.* later expanded on this by demonstrating on several genes that H3K4 dimethylation and the Set3 PHD domain are necessary for Set3 binding on coding regions and subsequent deacetylation by Hst1 and Hos2 [140].

We show that H3K4 di- or trimethylation is indeed required for the majority of Set3-ORF interactions (**Figure 4-2A** and **Figure 4-4C**). However, Set3 bound as well in these mutants than in the wild type over the subset of genes whose coding region was also

bound by Hst1 and Hos2. Furthermore, Hst1 binding is not affected by the loss of H3K4 methylation. While this is not surprising for the promoter-bound fraction, which represents the co-repressor function of Hst1, both Hst1 and Set3 binding on the coding region should be lost when H3K4 methylation is abolished according to Kim *et al.*'s model.

Although H3K4 methylation is not required for Set3 recruitment to highly transcribed genes, it is important to sequester Set3 to the coding region. Indeed, Set3 occupancy spreads to the promoter regions of these genes in H3K4 methylation-deficient strains whereas it is confined to the ORF in the wild type strain (**Figure 4-2C**). This redistribution is not due to changes in transcription levels nor has it any apparent effect on transcriptional activity (**Figure 4-2D**). We also observe increased Set3 binding at the 5' end of these coding regions in the *set1Δ* mutant relative to the wild type strain.

Curiously, Set3 binding on coding regions in *set1Δ* mutants mirror RNAPII distribution more closely than in the wild type (compare **Figure 4-2C** and **4-2D**). This is further confirmed by plotting Set3 abundance on the promoter and ORFs against RNAPII enrichment (**Figure 4-3A** and **4-3B**). There is a striking correlation between promoter- and ORF-bound Set3 in the *set1Δ* mutant which is absent or reduced, respectively, in the wild type.

These findings suggest that Set3 could interact directly with the transcriptional machinery, although whether this interaction has a positive or negative effect on transcription is not known at this point. However, other groups have described two opposite functions of Set3C on transcription. It was first discovered that Set3C represses early-middle sporulation genes' transcription independently of Hst1, and seeing as Set3C is only found at coding regions this effect is probably due to an elongation-related Set3C-mediated control of transcription. On the other hand, Set3C and Hos2 are required for the transcriptional activation of *GAL* and *INO1* genes. This function also has to be related to an elongation-related activity of Set3C, although opposite to that described above.

Furthermore, another argument supporting Set3-RNAPII interaction is the distribution of Set3 over the ORF in the wild type strain. If Set3 bound exclusively to di- and trimethylated H3K4 its binding pattern would overlap the distribution of these marks, as seen in **Figure 4-5A** and **4-5B**, meaning a large 5' peak and very little binding at the 3' end of the ORF. As we can see in **Figure 4-1C**, Set3 abundance is relatively constant over the whole coding region and does not relate to H3K4 di- and trimethyl patterns in the 3' end, especially for clusters 1 and 3 (red and green lines, respectively). H3K4 di- or trimethyl-dependant recruitment of Set3 is probably only a first step leading to the association of Set3 to the transcriptional machinery.

4.1.4.3 Further considerations

These conclusions could be improved in several ways. First, it would be interesting to see if we would lose the promoter-bound fraction of Hst1 when we mutate one or both of the Sum1 or Rfm1 co-repressors.

The Hos2 and Set3 binding phenotypes should also be further characterized. Mutants affecting different aspects of H3K4 methylation or Set3 methyl-lysine recognition will be used to this end. We will also use a *rad6Δ* mutant instead of the *spt4Δ* mutant we currently use to characterize Hos2 and Set3 behavior in a monomethyl-only H3K4 background. The *rad6Δ* mutant is less susceptible to indirect effects than the *spt4Δ* mutant. A *set1-RRMΔ* mutant will be used to assay Hos2 and Set3 binding in a strain incapable of trimethylating H3K4 but which retains H3K4 mono- and dimethylation capabilities. Note that the binding pattern of Hos2 in the wild type and H3K4 methylation defective strains also needs to be assessed to determine the integrity of the Set3C complex in these mutants.

Next, the HDACs Hst1 and Hos2 have both been described as being part of Set3C. Kim *et al.* have shown that acetylation levels are affected by the loss of H3K4 di- and trimethylation and also in a Set3 PHD domain mutation (W140A) which affects the methyl-

lysine binding pocket. We expect such an effect on genes where Set3 recruitment is abolished by a *set1Δ* mutation. However, the impact of this mutation on the acetylation of genes where Set3 and Hst1 both bind to the coding regions even in the absence of H3K4 methylation needs to be investigated.

We show that H3K4 methylation is not required for Set3 recruitment at a large subset of coding regions with high transcriptional activity. Related to this, H3K4 methylation is clearly only part of the mechanism tethering Set3 to the ORFs, as seen by the differences between H3K4 methylation and Set3 binding patterns in the wild type. The question remains as to what mechanism Set3 uses to tether itself to these regions.

Several mechanisms spring to mind. Set3 could interact with other methylated residues present over the whole coding region, such as H3K36 or H3K79. Deleting the histone methyltransferases responsible for these marks singly or in combination should give good insight into this hypothesis. Another alternative, supported by data from the Hinnebusch group [316], is that Set3C interacts with RNAPII directly, probably by recognizing patterns of C-terminal domain (CTD) phosphorylation. Mutations of the kinases responsible for Ser2, Ser5 and Ser7 CTD phosphorylation in wild type and *set1Δ* backgrounds should help determine if that is the case.

Finally, we observe many similarities between the Set3C complex and Rpd3L and Rpd3S such as the binding pattern over coding regions and the specificity for highly-transcribed genes. A study published in 2008 shows that the Rpd3L complex, also known to be found on coding regions of highly transcribed genes [289], interacts with at least three subunits of Set3C, namely Set3, Hos2 and Sif1 [190]. Preliminary analysis of Rpd3L subunit Sds3's binding profiles yields a profile surprisingly similar to that of Hst1, Hos2 and Set3 for cluster 8, where Hst1, Hos2, Set3 and Sds3 are found at the coding region (**Figure 4-6**). The implications of the possible interaction between two HDAC complexes with apparently opposing functions over genes where Set3 is recruited independently of H3K4 methylation imply novel very interesting mechanisms of transcription regulation.

4.1.5 Materials and Methods

4.1.5.1 Strains, Cell Growth and Crosslinking Conditions

All strains were grown in 50mL of YPD to an OD₆₀₀ of 0.6-0.8 before crosslinking, unless otherwise indicated. For ChIP-chip, strains were crosslinked with 1% formaldehyde for 30 min at room temperature on a wheel.

4.1.5.2 Chromatin Immunoprecipitation and Antibodies

ChIP experiments were performed as per [285], with minor modifications. For myc-tag ChIP, we used 5μg of 9E11 antibody coupled to 2×10^7 pan-mouse IgG DynaBeads (Invitrogen) per sample. RNAPII ChIPs were done using 2μL of 8WG16 antibody coupled to 2×10^7 pan-mouse IgG DynaBeads per sample. Note that in our ChIP-chip assays, 8WG16 generates profiles that are nearly identical as using a tagged RNAPII (data not shown). 8WG16 is therefore used here to measure total RNAPII levels on genes. The H3K4 methylation data were taken from Pokholok *et al.*, 2005 [341], and are normalized to H3 abundance.

The microarrays used for location analysis were purchased from Agilent Technologies (Palo Alto, California, United States) and contain a total of 44,290 Tm-adjusted 60-mer probes covering the entire genome for an average density of one probe every 275 bp (± 100 bp) within the probed regions (catalog # G4486A and G4493A). Myc-tag ChIPs were hybridized against ChIPs from isogenic strains that did not contain the tag as controls.

4.1.5.3 Data Analysis

The data were normalized and biological replicates were combined using a weighted average method as described previously [285]. The \log_2 ratio of each spot of combined datasets was then converted to Z-score, similar to Hogan et al. [335], to circumvent the large differences in the immunoprecipitation efficiencies of the different factors. Visual inspection of the Z-scores was carried out on the UCSC Genome Browser (<http://genome.ucsc.edu/>). All data analyses described here were done using data from protein-coding genes longer than or equal to 500 bp. Median Z-score values for promoter and complete length of each annotation from SGD (version Feb. 02 2008) were calculated without interpolation and used in our clustering and Pearson correlation analyses. Promoters are defined as the shortest of either 250bp or half the intergenic region (half-IG) relative to the reference gene's 5' boundary. Self-organizing map (SOM) clustering was done with the Cluster software [336] and visualized with Java Treeview [337]. Only genes with no missing value were used for clustering.

Gene mapping was performed as in Rufiange et al. [60] on selected groups of genes described in the text. Briefly, data were mapped onto the 5' and 3' boundaries in 50 bp windows for each half-gene and adjacent half-IG regions. A sliding window of 300 bp was then applied to the Z-scores to smooth the curve.

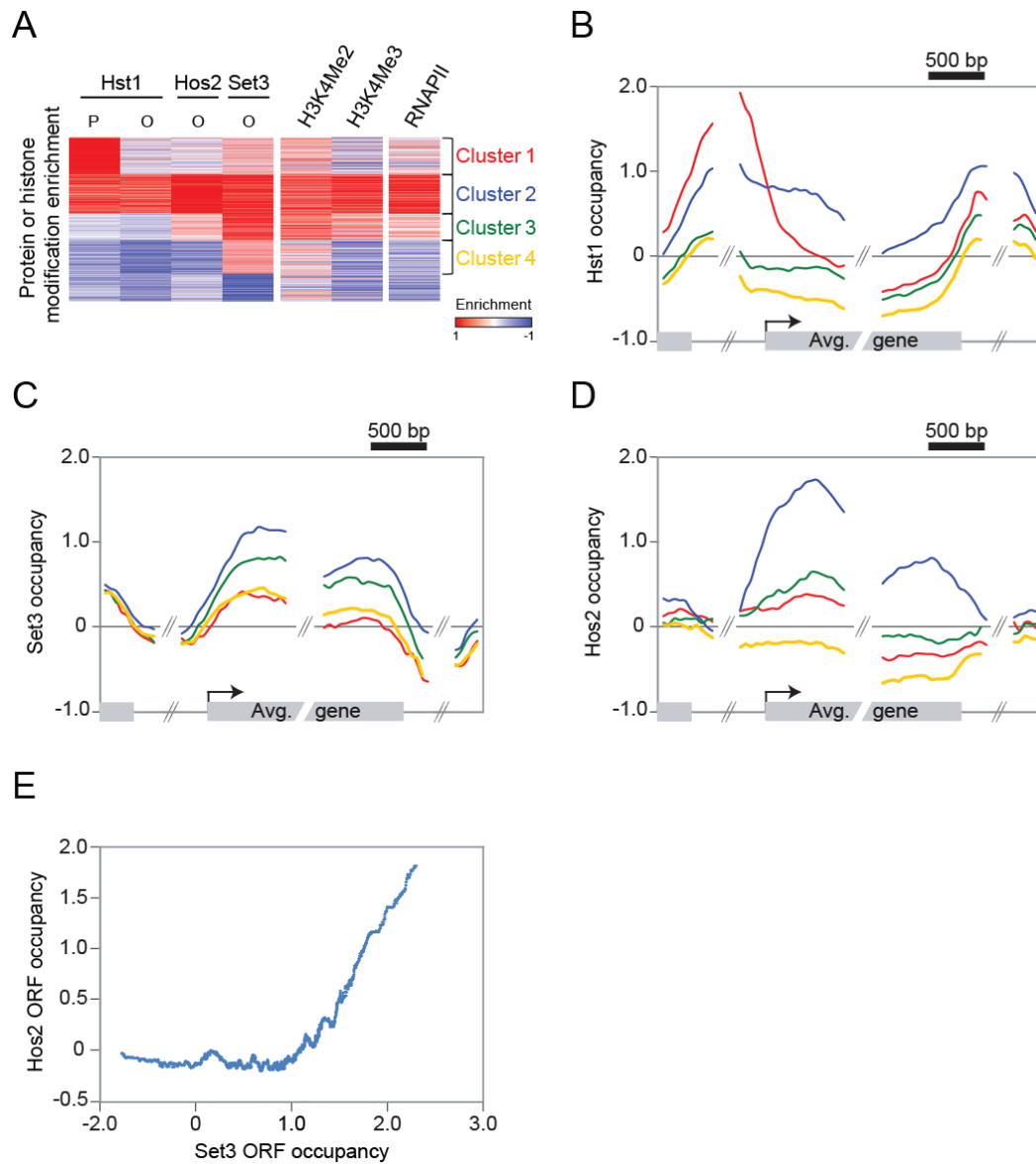


Figure 4-1 : Hst1, Hos2 and Set3 bind functionally distinct regions.

(A) Self-organizing map (SOM) clustering of Hst1 binding on promoter and ORFs, Hos2 binding on ORFs, and Set3 binding on ORFs. H3K4Me2, H3K4Me3 and RNAPII columns have not been clustered but are ordered in the same way. Red indicates enriched (bound) regions and blue represents depleted regions.

(B) Mapping of Hst1 occupancy on genes forming clusters 1 (red), 2 (blue), 3 (green) and 4 (gold) from panel “A”.

(C) Mapping of Set3 occupancy on genes from clusters 1 (red), 2 (blue), 3 (green) and 4 (gold) from panel “A”.

(D) Mapping of Hos2 occupancy on genes from clusters 1 (red), 2 (blue), 3 (green) and 4 (gold) from panel “A”.

(E) Hos2 ORF occupancy versus Set3 ORF occupancy. A sliding median window of 100 genes was applied to the data.

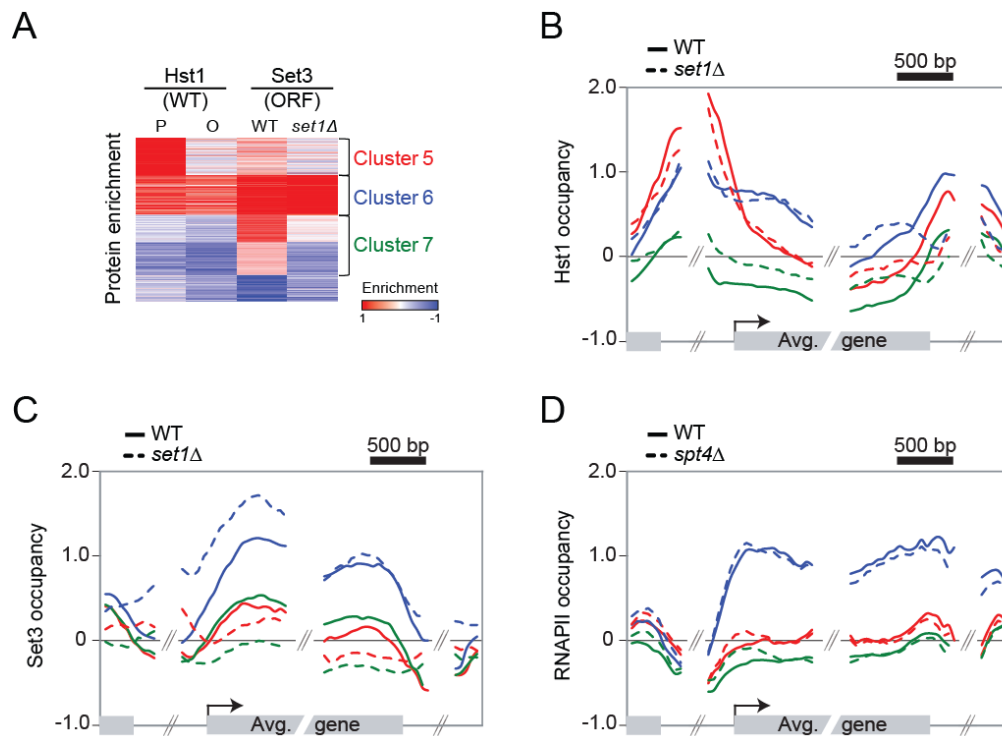


Figure 4-2 : Loss of H3K4 methylation affects Set3 but not Hst1 binding.

(A) Self-organizing map (SOM) clustering of WT Hst1 binding on promoter and ORFs and Set3 binding on ORFs in with WT and *set1Δ* strains. Red indicates enriched (bound) regions and blue represents depleted regions.

(B) Mapping of Hst1 occupancy in WT (solid lines) and *set1Δ* (dashed lines) cells on genes from clusters 5 (red), 6 (blue) and 7 (green) from panel “A”.

(C) Mapping of Set3 occupancy in WT (solid lines) and *set1Δ* (dashed lines) cells on genes from clusters 5 (red), 6 (blue) and 7 (green) from panel “A”.

(D) Mapping of RNAPII occupancy in WT (solid lines) and *spt4Δ* (dashed lines) cells on genes from clusters 5 (red), 6 (blue) and 7 (green) from panel “A”.

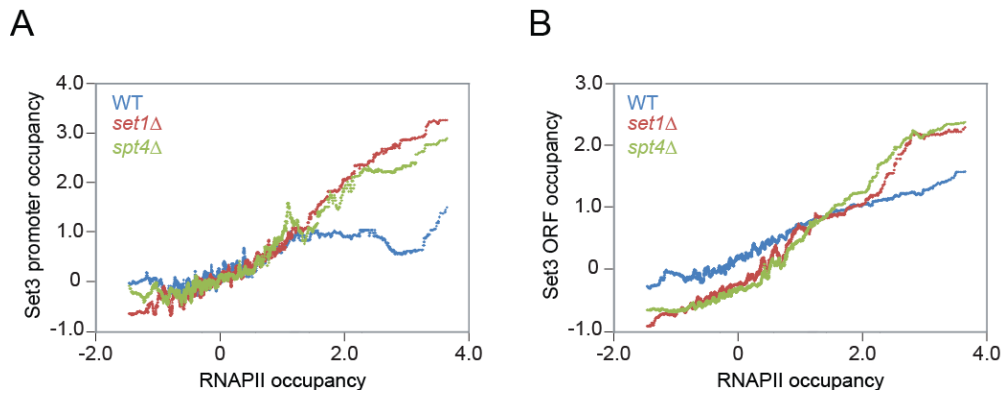


Figure 4-3 Set3 and RNAPII occupancies correlate better in H3K4 methylation-deficient backgrounds than in WT.

(A) Set3 promoter occupancy versus RNAPII ORF occupancy in WT (blue), *set1Δ* (red) and *spt4Δ* cells (green). A sliding median window of 100 genes was applied to the data.

(B) Set3 ORF occupancy versus RNAPII ORF occupancy in WT (blue), *set1Δ* (red) and *spt4Δ* cells (green). A sliding median window of 100 genes was applied to the data.

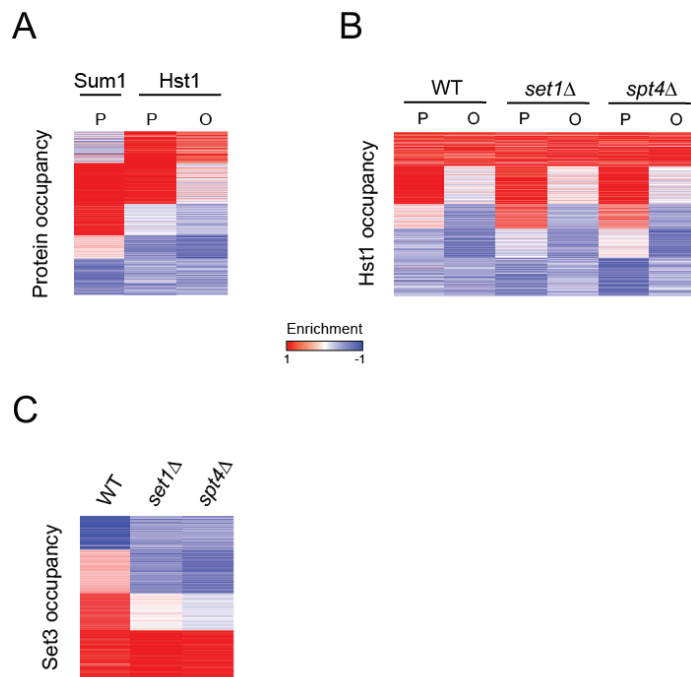


Figure 4-4 : A complement to Figure 4-1. Set3, but not Hst1, binding is affected by loss of H3K4 di- and trimethylation.

(A) Self-organizing map (SOM) clustering of Sum1 promoter binding and Hst1 promoter and ORF binding in WT cells. Red indicates enriched (bound) regions and blue represents depleted regions.

(B) Self-organizing map (SOM) clustering of Hst1 promoter and ORF binding in WT, *set1Δ* and *spt4Δ* cells. Red indicates enriched (bound) regions and blue represents depleted regions.

(C) Self-organizing map (SOM) clustering of Set3 ORF binding in WT, *set1Δ* and *spt4Δ* cells. Red indicates enriched (bound) regions and blue represents depleted regions.

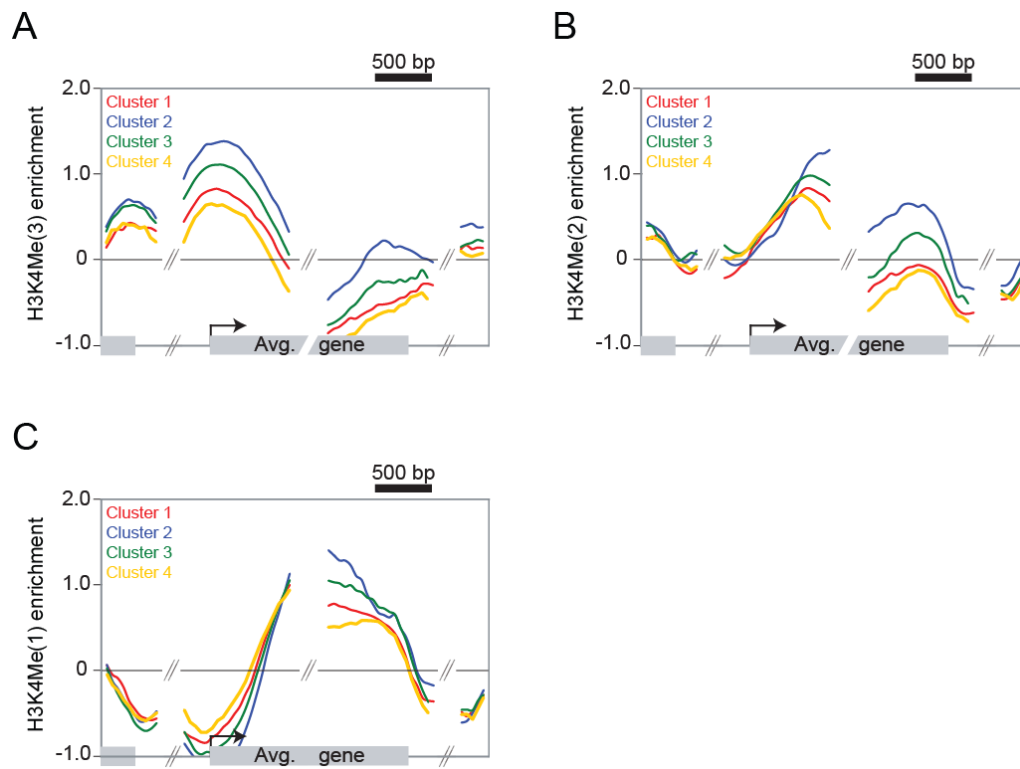


Figure 4-5 : H3K4 mono-, di- and trimethylation patterns in WT cells.

These data were taken from Pokholok *et al.*, 2005 [341], and are normalized to H3 abundance.

(A) Mapping of trimethyl H3K4 levels in WT cells on genes from clusters 1 (red), 2 (blue), 3 (green) and 4 (gold) from **Figure 4-1A**.

(B) Mapping of dimethyl H3K4 levels in WT cells on genes from clusters 1 (red), 2 (blue), 3 (green) and 4 (gold) from **Figure 4-1A**.

(C) Mapping of monomethyl H3K4 levels in WT cells on clusters 1 (red), 2 (blue), 3 (green) and 4 (gold) from **Figure 4-1A**.

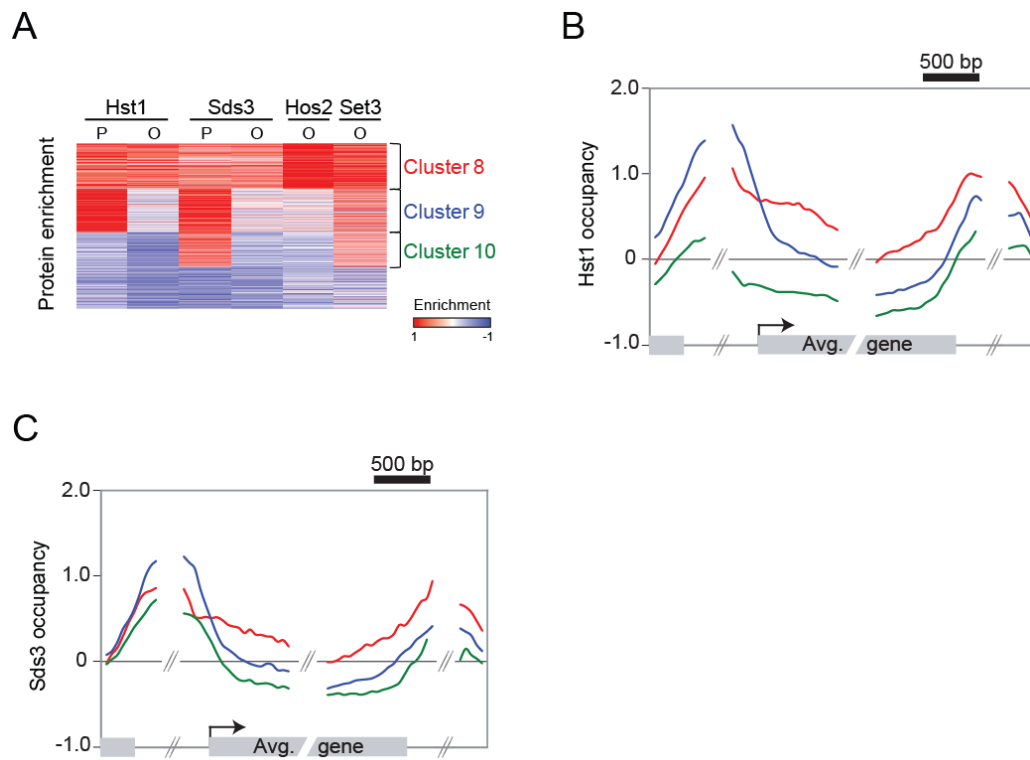


Figure 4-6 : Rpd3L and Set3C occupancies overlap at a set of ORFs.

(A) Self-organizing map (SOM) clustering of WT Hst1 and Rpd3L subunit Sds3 binding on promoter and ORF and Hos2 and Set3 occupancy on ORFs in WT strains. Red indicates enriched (bound) regions and blue represents depleted regions.

(B) Mapping of Hst1 occupancy on genes from clusters 8 (red), 9 (blue) and 10 (green) from panel “A”.

(C) Mapping of Sds3 occupancy on genes from clusters 8 (red), 9 (blue) and 10 (green) from panel “A”.

5 Conclusion

La chromatine constitue une stratégie évolutive incroyablement efficace pour stocker et protéger l'ADN. Toutefois, cette structure compacte représente aussi le principal obstacle aux processus de transcription et de réplication du bagage génétique. La régulation de la structure de la chromatine constitue par ce fait même la fonction la plus importante pour la survie de la cellule.

Les moyens mis en place pour moduler l'accès à l'ADN sont nombreux, complexes et strictement contrôlés. L'étude plus approfondie de ces mécanismes, débutée dans les années 1970, demeure jusqu'à ce jour très difficile de par leur nature vitale pour la survie de la cellule.

On peut diviser ces stratégies en quatre catégories générales, toutes interdépendantes à différents niveaux : le déplacement des nucléosomes catalysé par l'hydrolyse de l'ATP, le déplacement d'histones par des chaperonnes spécifiques, l'incorporation d'histones variants au sein du nucléosome et finalement les modifications covalentes post-traductionnelles imposées aux histones par une multitude d'enzymes. Mes travaux se sont concentrés sur ce dernier mécanisme chez la levure, spécifiquement sur l'impact fonctionnel de l'acétylation des résidus lysine, et plus particulièrement de leur déacétylation par les complexes HDAC Rpd3S et Set3C.

L'étude du rôle et de la régulation des HATs et HDACs responsable de la modulation de l'état d'acétylation des nucléosomes a connu une baisse d'intérêt considérable au début des années 2000. On avait alors démontré que tous les résidus lysine acétylables étaient équivalents, avec l'exception notable de H4K16. Le dogme à l'époque statuait que ces résidus agissaient de manière combinatoire pour déterminer le niveau d'activité transcriptionnelle des promoteurs : plus les nucléosomes situés au promoteur sont acétylés, plus l'activité transcriptionnelle est grande. Les complexes HDACs tels que Rpd3/Sin3 n'étaient alors considérés que comme de simples corépresseurs.

Ce n'est qu'avec la découverte d'un lien, maintenant très controversé, entre les mécanismes de vieillissement et l'activité HDAC de Sir2 que ce champ d'études a repris du

galon, même si le mode d'action attribué aux sirTuines était encore là limité à la corépression de gènes spécifiques. La caractérisation des différents rôles joués par les sirTuines chez la levure et particulièrement chez l'humain a monopolisé l'attention de plusieurs groupes et donné lieu à une véritable avalanche de publications débutant dans le milieu des années 2000 mais surtout à une nouvelle appréciation du rôle de l'acétylation dans le contrôle de l'environnement chromatinien et de la transcription.

La publication en 2005 d'études protéomiques à grande échelle chez la levure a mené à un questionnement concernant le rôle de Rpd3/Sin3 dans la régulation de la transcription. On a alors observé que Rpd3 se retrouve dans deux complexes structurellement et fonctionnellement indépendants : Rpd3L et Rpd3S. Le complexe Rpd3L serait responsable de la répression de la transcription des gènes déjà décrits comme ciblés par Rpd3/Sin3, tandis que le complexe Rpd3S jouerait un rôle aux régions codantes activement transcrites par l'ARN polymérase II.

La fonction de Rpd3S, inférée à partir de données obtenues lors d'expériences biochimiques *in vitro* en 2005, serait de déacétyler les nucléosomes après le passage de l'ARN polymérase II de manière à rétablir une structure chromatinienne plus compacte et réfractaire à la liaison de facteurs de transcriptions. On remarque en effet la présence d'événements de transcription cryptique et un niveau d'acétylation aberrant sur les régions codantes lorsque Rpd3S est muté. Nous n'avons toutefois que des observations indirectes pour supporter cette hypothèse; l'implication Rpd3S aux régions codantes activement transcrites n'avait jamais été observée jusque-là.

Mes travaux ont en grande partie confirmé le rôle répressif de Rpd3S aux ORFs transcrits. Toutefois, le mode de recrutement qui était proposé par le modèle courant, soit une interaction entre le chromodomaine d'Eaf3 et le résidu H3K36 méthylé par Set2, n'est pas valide *in vivo* : Rpd3S est plutôt associé au complexe transcriptionnel, particulièrement lorsque le domaine C-terminal de Rpb1 est hyperphosphorylé. De plus le complexe Spt4/Spt5, l'homologue chez la levure du facteur d'élongation humain DSIF, agit comme un antagoniste de cette interaction en limitant le chargement de Rpd3S sur le complexe d'élongation. Toutefois, le modèle n'avait pas tout faux, loin de là : des mutations de

Rpd3S, de Set2 ou du résidu H3K36 altèrent en effet les niveaux d'acétylation des nucléosomes présents aux promoteurs et aux régions codantes, démontrant que la fonction HDAC de Rpd3S dépend de la méthylation de H3K36. Ces nouvelles données corrigent et améliorent le modèle couramment accepté dans la littérature. Elles mettent aussi en évidence un nouveau niveau de complexité jusqu'alors méconnu dans la régulation de la transcription via la phosphorylation du domaine C-terminal de Rpb1 ainsi qu'une nouvelle fonction pour Spt4/Spt5, un complexe au rôle encore très peu caractérisé.

Le complexe Spt4/Spt5 n'agit pas de manière égale à tous les gènes transcrits : il agit principalement aux gènes où le niveau d'ARN polymérase II est le plus élevé et y réprime l'interaction entre Rpd3S et Rpb1. Ce groupe de gène représente approximativement la moitié de ceux où Rpd3S est présent aux régions codantes. Le reste des gènes où Rpd3S est présent démontre un niveau inférieur, mais quand même appréciable d'ARN polymérase II et l'interaction entre Rpd3S et l'ARN polymérase II à ces gènes est peu affectée par Spt4/Spt5. J'ai aussi découvert que la phosphorylation de la protéine Spt5 par la kinase Bur1/Bur2 régule l'activité de Spt4/Spt5. En effet, des mutants où Spt5 n'est pas phosphorylé ou phosphorylable démontrent une répression accrue de l'interaction entre Rpd3S et le domaine C-terminal de Rpb1, ce qui indique que la phosphorylation de Spt5 réprime l'inhibition du chargement de Rpd3S sur la machinerie transcriptionnelle par Spt4/Spt5.

La piste Rpd3S m'a aussi mené à une observation intrigante : le complexe Rpd3L, que l'on croyait uniquement présent aux promoteurs par des corépresseurs spécifiques, est aussi présent sur les mêmes régions codantes que Rpd3S. Cette observation remet sérieusement en question le rôle purement corépresseur de Rpd3L et soulève des points très intéressants relativement à sa fonction sur ces régions transcriptionnellement actives : est-elle redondante à celle de Rpd3S? Les complexes Rpd3S et Rpd3L sont-ils vraiment aussi distincts qu'on le croit? Une piste intéressante est apparue lorsque le groupe de Kim a publié en 2009 une étude démontrant que le complexe HDAC Set3C, à la fonction jusqu'alors mal définie, serait recruté aux régions codantes via l'interaction de sa sous-unité Set3 au résidu H3K4 di- ou triméthylé [140]. Une autre étude, publiée en 2008, suggère

l'existence d'un supercomplexe composé de Rpd3L ainsi que de plusieurs sous-unités de Set3C (« Rpd3L-Expanded ») [190]. De plus, Rpd3L ainsi que son homologue fonctionnel chez l'humain, mSin3, comportent des sous-unités pouvant interagir avec les résidus H3K4 méthylés, tout comme Set3, ce qui porte à penser que mSin3 pourrait combiner les fonctions de Rpd3L et Set3C. Ces observations m'ont évidemment mené à m'intéresser à Set3C pour déterminer si Rpd3L-Expanded serait le complexe que j'observe recruté aux régions codantes aussi occupées par Rpd3S.

Bien que la détermination de l'existence de Rpd3L-Expanded *in vivo* était mon but premier, mon attention a été détournée par l'étude des données intrigantes obtenues lorsque je voulais reproduire les résultats de Kim et de ses collègues : contrairement à ce qu'ils soutiennent, H3K4 di- ou triméthylé n'est pas essentiel au recrutement de Set3 à toutes les régions codantes activement transcrites. En effet, le recrutement de Set3 aux régions codantes où l'abondance d'ARN polymérase II est la plus élevée n'est pas affecté par la perte de cette marque de méthylation. On sait depuis des années que Set3C renferme deux HDACs, Hst1 et Hos2. La première possède une fonction corépressive aux promoteurs lorsqu'elle interagit avec Sum1 ou Rfm1, mais peut aussi être associée à Set3C, tandis que la seconde ne se retrouve que dans Set3C. Mes données indiquent que Hst1 est en effet majoritairement associée aux promoteurs indépendamment de Set3C, mais qu'une fraction est présente sur les mêmes régions codantes très transcrites que Set3 et que cette liaison s'effectue par mécanisme n'impliquant pas d'interaction avec le résidu H3K4 méthylé. Ces résultats sont intéressants, mais pas si surprenants quand on sait que Hst1 est considéré comme étant un membre moins stable de Set3C. Par contre, la présence de Hos2 presque exclusivement à ces mêmes régions constitue une surprise de taille, indiquant que la fraction de Set3 dont le recrutement dépend de la méthylation de H3K4 est soit en forme monomérique ou fait partie d'un complexe Set3C dépourvu de ses sous-unités HDACs. Il est d'autant plus intéressant de constater que Rpd3L est aussi présent sur ces régions, ce qui supporte l'hypothèse de l'existence du complexe Rpd3L-Expanded. Malgré ces avancées dans la caractérisation de Set3C, sa fonction et même sa composition exacte restent encore équivoques. On peut toutefois lui imputer un rôle semblable à celui de Rpd3S en se basant sur

deux observations : le chevauchement de la localisation de ces deux complexes aux régions codantes activement transcrites et leur activité catalytique identique. Toutefois, comme cette hypothèse n'est supportée que par des données circonstanciées, elle est évidemment à prendre avec un grain de sel.

La caractérisation des modes de recrutement et de la fonction de Rpd3S sont des morceaux importants dans le casse-tête de la régulation de la transcription par l'ARN polymérase II. Ces découvertes en camouflent cependant une autre d'importance capitale : nous avons finalement associé une fonction à deux des trois résidus lysine méthylés chez la levure, soit H3K4 et H3K36. On savait que la présence de cette modification était associée aux gènes activement transcrits et qu'elle ne servait que de mémoire transcriptionnelle étant donné sa longue demi-vie. L'association de Rpd3S au résidu H3K36 méthylé, bien que facultative au recrutement de ce complexe, est néanmoins essentielle à sa fonction HDAC et au rétablissement d'une structure chromatinienne répressive après le passage de l'ARN polymérase II. La méthylation de H3K4, quant à elle, est nécessaire au recrutement d'une forme de Set3C où les HDACs Hst1 et Hos2 sont absentes. Bien entendu, ces deux fonctions ne sont probablement que la pointe de l'iceberg et nous verrons sous peu l'attribution d'autres fonctions à ces résidus méthylés. D'ailleurs, la fonction exacte de H3K79 méthylé est encore inconnue, bien qu'on sache qu'elle est impliquée dans le maintien de l'intégrité génomique.

On croyait jusqu'à récemment que les marques de méthylation des résidus lysine ne pouvaient être éliminées que par le remplacement des histones modifiées, par exemple lors de la réplication de l'ADN. La découverte en 2004 de la première histone déméthylase chez l'humain, LSD1, a remis en question ce modèle et, ce faisant, secoué le monde de la chromatine. On sait maintenant que ces enzymes ont des homologues jusque chez la levure, où l'on en retrouve quatre : Jhd1, Jhd2, Gis1 et Rph1, ayant des affinités à H3K4 ou à H3K36 méthylé.

Bien que les cibles de ces protéines soient maintenant relativement bien caractérisées, on ne sait toujours rien de leur mode de recrutement, de leurs partenaires ou de leur fonction. Par contre, plusieurs pistes suggèrent des interactions fonctionnelles et

physiques entre les complexes HDACs et les histone déméthylases. Premièrement, chez l'humain LSD1 fait partie du complexe HDAC coREST, dont les sous-unités catalytiques, HDAC1 et HDAC2, sont des homologues de Rpd3. Le rôle de LSD1 dans ce complexe est de déméthyliser les résidus H3K4, ce qui active la fonction HDAC de coREST. Cette fonction est en contradiction apparente avec l'affinité de mSin3 pour le résidu H3K4 méthylé, mais il ne s'agit sans doute que d'une conséquence de la spécialisation de coREST pour la régulation de gènes dans le tissu nerveux. Deuxièmement, on retrouve l'histone déméthylase LID associée aux homologues de Rpd3S chez l'humain et la drosophile, soit, respectivement, SHMP et RLAF. Son rôle dans ces complexes serait, comme pour LSD1, de déméthyliser H3K4. Finalement, des interactions génétiques ont été répertoriées entre Jhd1 et le complexe PAF, un complexe recruté par DSIF et impliqué dans la méthylation de H3K4, ainsi qu'entre Jhd2 et Rco1, SAGA et NuA4 [62]. Toutes ces données indiquent que les histone déméthylases, tout comme les HDACs auxquelles elles sont associées, jouent des rôles potentiellement importants dans la régulation de la transcription et le maintien de l'identité des régions codantes. Il reste maintenant à savoir si les observations faites chez l'humain sont transposables à la levure.

La relâche de l'étude du rôle de l'acétylation dans le contrôle de la transcription au début des années 2000 aurait pu avoir des conséquences fâcheuses sur le statut de la levure comme organisme modèle pour l'étude de la chromatine. On présumait alors avoir résolu le problème de l'acétylation en l'associant uniquement à un contrôle plus ou moins direct de l'activité transcriptionnelle via la modulation de la compaction des nucléosomes situés aux promoteurs. Les dernières années nous ont démontré que cette vision des choses était très réductrice et que l'acétylation (ainsi que la méthylation) des nucléosomes jouait un rôle beaucoup plus fondamental dans le maintien de l'identité des régions codantes. Les duplications des familles de HDAC, leur spécialisation subséquente et leur grande redondance font de l'humain et des cellules mammifères en général des modèles peu commodes. La levure, de par sa relative simplicité et la grande homologie fonctionnelle retrouvée dans ces fonctions biologiques fondamentales, est un organisme idéal pour disséquer ces mécanismes pour ensuite les transposer et les confirmer chez l'humain.

Malgré ces avantages, plusieurs mécanismes essentiels sont curieusement sous-étudiés chez la levure par rapport à l'humain. La caractérisation de la régulation des différents complexes HDACs chez la levure fait partie de ce lot. Chez l'humain l'étude de ces mécanismes est relativement avancée, un état de choses dû en grande partie à la grande spécialisation des multiples HDACs ayant des rôles assez pointus. Cette spécialisation amène par la force des choses une régulation très précise que l'on doit comprendre en détail si l'étude des mécanismes en découlant nous intéresse. La perception de simplicité rattachée aux HDACs a causé un retard dans l'identification et la caractérisation de ces mécanismes. L'étude de ces méthodes de régulation chez la levure doit devenir prioritaire, car elle mènerait potentiellement à l'identification de nouvelles cibles thérapeutiques chez l'humain, ou du moins à une meilleure compréhension du mode d'action des agents inhibiteurs de HDACs présentement en développement et en essais cliniques.

6 Perspectives

La caractérisation de la fonction de Rpd3S, et en particulier du contrôle de son recrutement au domaine C-terminal hyperphosphorylé de Rpb1 par le complexe Spt4/Spt5, remet en question le dogme stipulant que les HDAC sont limitées à des rôles de corépresseurs de la transcription. La présence de multiples formes du complexe Set3C recrutées selon différents mécanismes aux régions codantes apporte un autre support à ce concept. L'étude de ces « nouvelles » fonctions permettra de comprendre les mécanismes encadrant l'activité de l'appareil transcriptionnel. Je suggérerai maintenant quelques pistes intéressantes découlant des travaux présentés dans cette thèse.

Bien que l'on connaisse beaucoup mieux le rôle de Rpd3S et les règles régissant son implication lors de la transcription, on ne sait que très peu de choses de son partenaire, Spt4/Spt5. Ce complexe, bien qu'essentiel, est encore relativement méconnu et gagnerait à être étudié plus en profondeur. La protéine Spt5 comporte un domaine C-terminal composé de séquences aminées répétées, ce qui n'est pas sans rappeler celui de Rpb1. Ce domaine, lorsque phosphorylé par Bur1/Bur2, est impliqué dans le relâchement de la répression du recrutement de Rpd3S. De plus, sa ressemblance au domaine C-terminal de Rpb1 en fait potentiellement une plate-forme d'interaction protéique intéressante. Il serait donc intéressant d'étudier plus à fond ce complexe à l'aide de techniques protéomiques, biochimiques et génomiques pour déterminer précisément son rôle dans la régulation des processus transcriptionnels.

Mes données démontrent que le complexe Rpd3L est surtout présent aux promoteurs de gènes, très probablement pour en réprimer l'expression. Ce résultat était attendu car il s'agit pour ce complexe d'un rôle bien établi et abondamment décrit dans la littérature. Ce qui est plus surprenant est sa présence aux régions codantes activement transcrites où Rpd3S est aussi localisé. La fonction de Rpd3L sur ces régions est inconnue, mais des données publiées par le groupe de Shevchenko suggèrent l'existence d'une interaction physique entre ce complexe et Set3C [190]. Mes observations ne permettent pas d'établir directement une colocalisation de ces deux complexes mais indiquent tout de

même qu'ils sont tous deux localisés sur les mêmes régions codantes (**Figure 4-6**). Encore là, des expériences biochimiques, génétiques et génomiques permettraient de mieux caractériser le rôle de Rpd3L sur ces régions et d'établir s'il y a bel et bien une interaction physique ou fonctionnelle entre ce complexe et Set3C.

Bien que le complexe Set3C ait été découvert au début des années 2000 [136], très peu de données sont disponibles à son sujet. Ce n'est que récemment que le groupe de Kim et ses collègues ont démontré que le résidu H3K4 di- ou triméthylé était important, voir essentiel pour le recrutement de ce complexe aux régions codantes. Mes données, présentées au chapitre 4, contredisent partiellement cette affirmation : la méthylation de H3K4 n'est essentielle au recrutement de Set3 que sur une partie de ses cibles, mais pas aux gènes où le taux d'occupation de l'ARN polymérase II est le plus élevé (**Figure 4-2**). Il est particulièrement intéressant de noter que les HDACs Hst1 et Hos2, membres du complexe Set3C, ne se retrouvent qu'aux régions codantes où l'on retrouve Set3 indépendamment de l'état de méthylation de H3K4. Toutes ces observations laissent présager l'existence d'une forme du complexe Set3C où Hst1 et Hos2 seraient absentes. Ce complexe « Set3-sans-HDAC » est recruté via l'interaction du domaine PHD de Set3 au résidu H3K4 di- ou triméthylé.

La caractérisation de Set3C ne fait que commencer : on ne sait toujours presque rien de son rôle et de son mode de recrutement aux régions codantes transcriptionnellement actives; il serait donc très important de trouver réponse à ces questions. Bien qu'on sache maintenant que le recrutement de Set3 à ces régions ne requiert pas la méthylation de H3K4, il faut tout de même vérifier si le recrutement des autres facteurs composant Set3C, particulièrement Hst1 et Hos2, est réellement indépendant de la méthylation de H3K4 ou si, comme pour Rpd3S et le résidu H3K36 méthylé, cette marque n'est importante que pour l'activité enzymatique de Set3C. Il faudra ensuite se pencher sur le mode de recrutement de Set3C aux gènes activement transcrits. Ces régions, où l'on retrouve à la fois Rpd3S, Rpd3L et Set3C, ont toutes en commun un niveau d'ARN polymérase II assez élevé par rapport au reste du génome. On peut donc imaginer que le recrutement de Set3C pourrait passer par un mécanisme similaire à celui de Rpd3S, via une interaction avec le domaine C-

terminal hyperphosphorylé de Rpb1. De plus, il n'est pas impossible que le complexe Spt4/Spt5 intervienne dans le recrutement de Set3C.

Ces idées et perspectives ne forment évidemment que la pointe de l'iceberg : la nouvelle vague d'études concernant l'implication des HDACs dans la régulation de la structure de la chromatine et des mécanismes associés n'en est qu'à ses débuts. Les raffinements qui seront apportés au modèle transcriptionnel chez la levure seront d'une importance capitale car ces mécanismes sont difficiles à étudier chez l'humain et jouent un rôle central dans plusieurs maladies complexes. J'espère que mes travaux ainsi que les suites en découlant apporteront quelques pièces à ce casse-tête colossal.

7 Bibliographie

1. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA*. Science, 1974. **184**(139): p. 868-71.
2. Klug, A., et al., *A low resolution structure for the histone core of the nucleosome*. Nature, 1980. **287**(5782): p. 509-16.
3. Richmond, T.J., et al., *Structure of the nucleosome core particle at 7 Å resolution*. Nature, 1984. **311**(5986): p. 532-7.
4. Arents, G., et al., *The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix*. Proc Natl Acad Sci U S A, 1991. **88**(22): p. 10148-52.
5. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 Å resolution*. Nature, 1997. **389**(6648): p. 251-60.
6. Kelly, T.J., et al., *Type B histone acetyltransferase Hat1p participates in telomeric silencing*. Mol Cell Biol, 2000. **20**(19): p. 7051-8.
7. Waterborg, J.H., *Dynamics of histone acetylation in Saccharomyces cerevisiae*. Biochemistry, 2001. **40**(8): p. 2599-605.
8. Katan-Khaykovich, Y. and K. Struhl, *Dynamics of global histone acetylation and deacetylation in vivo: rapid restoration of normal histone acetylation status upon removal of activators and repressors*. Genes Dev, 2002. **16**(6): p. 743-52.
9. Wittschieben, B.O., et al., *A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme*. Mol Cell, 1999. **4**(1): p. 123-8.
10. Winkler, G.S., et al., *Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels in vivo*. Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3517-22.
11. Winkler, G.S., et al., *RNA polymerase II elongator holoenzyme is composed of two discrete subcomplexes*. J Biol Chem, 2001. **276**(35): p. 32743-9.
12. Otero, G., et al., *Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation*. Mol Cell, 1999. **3**(1): p. 109-18.
13. Krogan, N.J. and J.F. Greenblatt, *Characterization of a six-subunit holo-elongator complex required for the regulated expression of a group of genes in Saccharomyces cerevisiae*. Mol Cell Biol, 2001. **21**(23): p. 8203-12.
14. Wittschieben, B.O., et al., *Overlapping roles for the histone acetyltransferase activities of SAGA and elongator in vivo*. EMBO J, 2000. **19**(12): p. 3060-8.
15. Kristjuhan, A., et al., *Spreading of Sir3 protein in cells with severe histone H3 hypoacetylation*. Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7551-6.
16. Tansey, W.P. and W. Herr, *TAFs: guilt by association?* Cell, 1997. **88**(6): p. 729-32.
17. Grant, P.A., et al., *A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation*. Cell, 1998. **94**(1): p. 45-53.
18. Lee, T.I., et al., *Redundant roles for the TFIID and SAGA complexes in global transcription*. Nature, 2000. **405**(6787): p. 701-4.

19. Basehoar, A.D., S.J. Zanton, and B.F. Pugh, *Identification and distinct regulation of yeast TATA box-containing genes*. Cell, 2004. **116**(5): p. 699-709.
20. Huisinga, K.L. and B.F. Pugh, *A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae*. Mol Cell, 2004. **13**(4): p. 573-85.
21. Georgakopoulos, T. and G. Thireos, *Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription*. EMBO J, 1992. **11**(11): p. 4145-52.
22. Marcus, G.A., et al., *Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors*. EMBO J, 1994. **13**(20): p. 4807-15.
23. Georgakopoulos, T., N. Gounalaki, and G. Thireos, *Genetic evidence for the interaction of the yeast transcriptional co-activator proteins GCN5 and ADA2*. Mol Gen Genet, 1995. **246**(6): p. 723-8.
24. Horiuchi, J., et al., *ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex*. Mol Cell Biol, 1995. **15**(3): p. 1203-9.
25. Brownell, J.E., et al., *Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation*. Cell, 1996. **84**(6): p. 843-51.
26. Kuo, M.H., et al., *Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines*. Nature, 1996. **383**(6597): p. 269-72.
27. Grant, P.A., et al., *Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex*. Genes Dev, 1997. **11**(13): p. 1640-50.
28. Ruiz-Garcia, A.B., et al., *Gcn5p is involved in the acetylation of histone H3 in nucleosomes*. FEBS Lett, 1997. **403**(2): p. 186-90.
29. Grant, P.A., et al., *Expanded lysine acetylation specificity of Gcn5 in native complexes*. J Biol Chem, 1999. **274**(9): p. 5895-900.
30. Gregory, P.D., et al., *Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the PHO5 promoter in yeast*. Mol Cell, 1998. **1**(4): p. 495-505.
31. Wang, L., L. Liu, and S.L. Berger, *Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo*. Genes Dev, 1998. **12**(5): p. 640-53.
32. Kuo, M.H., et al., *Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo*. Genes Dev, 1998. **12**(5): p. 627-39.
33. Kuo, M.H., et al., *Gcn4 activator targets Gcn5 histone acetyltransferase to specific promoters independently of transcription*. Mol Cell, 2000. **6**(6): p. 1309-20.
34. Pray-Grant, M.G., et al., *The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway*. Mol Cell Biol, 2002. **22**(24): p. 8774-86.
35. Sterner, D.E., R. Belotserkovskaya, and S.L. Berger, *SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription*. Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11622-7.

36. Robert, F., et al., *Global position and recruitment of HATs and HDACs in the yeast genome*. Mol Cell, 2004. **16**(2): p. 199-209.
37. Smith, E.R., et al., *ESAI is a histone acetyltransferase that is essential for growth in yeast*. Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3561-5.
38. Allard, S., et al., *NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tral1p*. EMBO J, 1999. **18**(18): p. 5108-19.
39. Clarke, A.S., et al., *Esa1p is an essential histone acetyltransferase required for cell cycle progression*. Mol Cell Biol, 1999. **19**(4): p. 2515-26.
40. Reid, J.L., et al., *Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase*. Mol Cell, 2000. **6**(6): p. 1297-307.
41. Bird, A.W., et al., *Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair*. Nature, 2002. **419**(6905): p. 411-5.
42. Boudreault, A.A., et al., *Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin*. Genes Dev, 2003. **17**(11): p. 1415-28.
43. Selleck, W., et al., *The Saccharomyces cerevisiae Piccolo NuA4 histone acetyltransferase complex requires the Enhancer of Polycomb A domain and chromodomain to acetylate nucleosomes*. Mol Cell Biol, 2005. **25**(13): p. 5535-42.
44. Chang, C.S. and L. Pillus, *Collaboration between the essential Esa1 acetyltransferase and the Rpd3 deacetylase is mediated by H4K12 histone acetylation in Saccharomyces cerevisiae*. Genetics, 2009. **183**(1): p. 149-60.
45. Lu, P.Y., N. Levesque, and M.S. Kobor, *NuA4 and SWRI-C: two chromatin-modifying complexes with overlapping functions and components*. Biochem Cell Biol, 2009. **87**(5): p. 799-815.
46. Guillemette, B. and L. Gaudreau, *Reuniting the contrasting functions of H2A.Z*. Biochem Cell Biol, 2006. **84**(4): p. 528-35.
47. Reifsnyder, C., et al., *Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases*. Nat Genet, 1996. **14**(1): p. 42-9.
48. Takechi, S. and T. Nakayama, *Sas3 is a histone acetyltransferase and requires a zinc finger motif*. Biochem Biophys Res Commun, 1999. **266**(2): p. 405-10.
49. John, S., et al., *The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex*. Genes Dev, 2000. **14**(10): p. 1196-208.
50. Howe, L., et al., *Histone H3 specific acetyltransferases are essential for cell cycle progression*. Genes Dev, 2001. **15**(23): p. 3144-54.
51. Howe, L., et al., *Yng1p modulates the activity of Sas3p as a component of the yeast NuA3 histone acetyltransferase complex*. Mol Cell Biol, 2002. **22**(14): p. 5047-53.
52. Taverna, S.D., et al., *Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs*. Mol Cell, 2006. **24**(5): p. 785-96.

53. Rosaleny, L.E., et al., *The Sas3p and Gcn5p histone acetyltransferases are recruited to similar genes*. *Genome Biol*, 2007. **8**(6): p. R119.
54. Kleff, S., et al., *Identification of a gene encoding a yeast histone H4 acetyltransferase*. *J Biol Chem*, 1995. **270**(42): p. 24674-7.
55. Parthun, M.R., J. Widom, and D.E. Gottschling, *The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism*. *Cell*, 1996. **87**(1): p. 85-94.
56. Ruiz-Garcia, A.B., et al., *HAT1 and HAT2 proteins are components of a yeast nuclear histone acetyltransferase enzyme specific for free histone H4*. *J Biol Chem*, 1998. **273**(20): p. 12599-605.
57. Ai, X. and M.R. Parthun, *The nuclear Hat1p/Hat2p complex: a molecular link between type B histone acetyltransferases and chromatin assembly*. *Mol Cell*, 2004. **14**(2): p. 195-205.
58. Poveda, A., et al., *Hif1 is a component of yeast histone acetyltransferase B, a complex mainly localized in the nucleus*. *J Biol Chem*, 2004. **279**(16): p. 16033-43.
59. Fillingham, J., et al., *Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109*. *Mol Cell Biol*, 2008. **28**(13): p. 4342-53.
60. Rufiange, A., et al., *Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1*. *Mol Cell*, 2007. **27**(3): p. 393-405.
61. Schneider, J., et al., *Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II*. *J Biol Chem*, 2006. **281**(49): p. 37270-4.
62. Collins, S.R., et al., *Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map*. *Nature*, 2007. **446**(7137): p. 806-10.
63. Driscoll, R., A. Hudson, and S.P. Jackson, *Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56*. *Science*, 2007. **315**(5812): p. 649-52.
64. Han, J., et al., *Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity*. *J Biol Chem*, 2007. **282**(39): p. 28587-96.
65. Han, J., et al., *The Rtt109-Vps75 histone acetyltransferase complex acetylates non-nucleosomal histone H3*. *J Biol Chem*, 2007. **282**(19): p. 14158-64.
66. Selth, L. and J.Q. Svejstrup, *Vps75, a new yeast member of the NAP histone chaperone family*. *J Biol Chem*, 2007. **282**(17): p. 12358-62.
67. Masumoto, H., et al., *A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response*. *Nature*, 2005. **436**(7048): p. 294-8.
68. Ozdemir, A., et al., *Characterization of lysine 56 of histone H3 as an acetylation site in *Saccharomyces cerevisiae**. *J Biol Chem*, 2005. **280**(28): p. 25949-52.
69. Maas, N.L., et al., *Cell cycle and checkpoint regulation of histone H3 K56 acetylation by Hst3 and Hst4*. *Mol Cell*, 2006. **23**(1): p. 109-19.
70. Celic, I., et al., *The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation*. *Curr Biol*, 2006. **16**(13): p. 1280-9.

71. Williams, S.K., D. Truong, and J.K. Tyler, *Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation*. Proc Natl Acad Sci U S A, 2008. **105**(26): p. 9000-5.
72. Varv, S., et al., *Acetylation of H3 K56 is required for RNA polymerase II transcript elongation through heterochromatin in yeast*. Mol Cell Biol, 2010. **30**(6): p. 1467-77.
73. Ehrenhofer-Murray, A.E., D.H. Rivier, and J. Rine, *The role of Sas2, an acetyltransferase homologue of Saccharomyces cerevisiae, in silencing and ORC function*. Genetics, 1997. **145**(4): p. 923-34.
74. Osada, S., et al., *The yeast SAS (something about silencing) protein complex contains a MYST-type putative acetyltransferase and functions with chromatin assembly factor ASF1*. Genes Dev, 2001. **15**(23): p. 3155-68.
75. Sutton, A., et al., *Sas4 and Sas5 are required for the histone acetyltransferase activity of Sas2 in the SAS complex*. J Biol Chem, 2003. **278**(19): p. 16887-92.
76. Shia, W.J., et al., *Characterization of the yeast trimeric-SAS acetyltransferase complex*. J Biol Chem, 2005. **280**(12): p. 11987-94.
77. Suka, N., K. Luo, and M. Grunstein, *Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin*. Nat Genet, 2002. **32**(3): p. 378-83.
78. Kimura, A., T. Umehara, and M. Horikoshi, *Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing*. Nat Genet, 2002. **32**(3): p. 370-7.
79. Shia, W.J., B. Li, and J.L. Workman, *SAS-mediated acetylation of histone H4 Lys 16 is required for H2A.Z incorporation at subtelomeric regions in Saccharomyces cerevisiae*. Genes Dev, 2006. **20**(18): p. 2507-12.
80. Vidal, M., et al., *Direct selection for mutants with increased K⁺ transport in Saccharomyces cerevisiae*. Genetics, 1990. **125**(2): p. 313-20.
81. Vidal, M. and R.F. Gaber, *RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in Saccharomyces cerevisiae*. Mol Cell Biol, 1991. **11**(12): p. 6317-27.
82. Vidal, M., et al., *RPD1 (SIN3/UME4) is required for maximal activation and repression of diverse yeast genes*. Mol Cell Biol, 1991. **11**(12): p. 6306-16.
83. Stillman, D.J., S. Dorland, and Y. Yu, *Epistasis analysis of suppressor mutations that allow HO expression in the absence of the yeast SW15 transcriptional activator*. Genetics, 1994. **136**(3): p. 781-8.
84. Taunton, J., C.A. Hassig, and S.L. Schreiber, *A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p*. Science, 1996. **272**(5260): p. 408-11.
85. Rundlett, S.E., et al., *HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription*. Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14503-8.
86. Kadosh, D. and K. Struhl, *Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters*. Cell, 1997. **89**(3): p. 365-71.

87. Kadosh, D. and K. Struhl, *Histone deacetylase activity of Rpd3 is important for transcriptional repression in vivo*. Genes Dev, 1998. **12**(6): p. 797-805.
88. Kasten, M.M., S. Dorland, and D.J. Stillman, *A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators*. Mol Cell Biol, 1997. **17**(8): p. 4852-8.
89. Kadosh, D. and K. Struhl, *Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo*. Mol Cell Biol, 1998. **18**(9): p. 5121-7.
90. Rundlett, S.E., et al., *Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3*. Nature, 1998. **392**(6678): p. 831-5.
91. Perez-Martin, J. and A.D. Johnson, *Mutations in chromatin components suppress a defect of Gcn5 protein in Saccharomyces cerevisiae*. Mol Cell Biol, 1998. **18**(2): p. 1049-54.
92. Burgess, S.M., M. Ajimura, and N. Kleckner, *GCN5-dependent histone H3 acetylation and RPD3-dependent histone H4 deacetylation have distinct, opposing effects on IME2 transcription, during meiosis and during vegetative growth, in budding yeast*. Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6835-40.
93. Deckert, J. and K. Struhl, *Targeted recruitment of Rpd3 histone deacetylase represses transcription by inhibiting recruitment of Swi/Snf, SAGA, and TATA binding protein*. Mol Cell Biol, 2002. **22**(18): p. 6458-70.
94. Dora, E.G., et al., *RPD3 (REC3) mutations affect mitotic recombination in Saccharomyces cerevisiae*. Curr Genet, 1999. **35**(2): p. 68-76.
95. Vogelauer, M., et al., *Histone acetylation regulates the time of replication origin firing*. Mol Cell, 2002. **10**(5): p. 1223-33.
96. Aparicio, J.G., et al., *The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in Saccharomyces cerevisiae*. Mol Cell Biol, 2004. **24**(11): p. 4769-80.
97. Huang, D., et al., *Dual regulation by pairs of cyclin-dependent protein kinases and histone deacetylases controls G1 transcription in budding yeast*. PLoS Biol, 2009. **7**(9): p. e1000188.
98. Takahata, S., Y. Yu, and D.J. Stillman, *The E2F functional analogue SBF recruits the Rpd3(L) HDAC, via Whi5 and Stb1, and the FACT chromatin reorganizer, to yeast G1 cyclin promoters*. EMBO J, 2009. **28**(21): p. 3378-89.
99. Ruiz-Roig, C., et al., *The Rpd3L HDAC complex is essential for the heat stress response in yeast*. Mol Microbiol, 2010.
100. De Nadal, E., et al., *The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes*. Nature, 2004. **427**(6972): p. 370-4.
101. Rohde, J.R. and M.E. Cardenas, *The tor pathway regulates gene expression by linking nutrient sensing to histone acetylation*. Mol Cell Biol, 2003. **23**(2): p. 629-35.
102. Humphrey, E.L., et al., *Rpd3p relocation mediates a transcriptional response to rapamycin in yeast*. Chem Biol, 2004. **11**(3): p. 295-9.
103. Tsang, C.K., et al., *Chromatin-mediated regulation of nucleolar structure and RNA Pol I localization by TOR*. EMBO J, 2003. **22**(22): p. 6045-56.

104. Vannier, D., D. Balderes, and D. Shore, *Evidence that the transcriptional regulators SIN3 and RPD3, and a novel gene (SDS3) with similar functions, are involved in transcriptional silencing in S. cerevisiae*. Genetics, 1996. **144**(4): p. 1343-53.
105. Smith, J.S., E. Caputo, and J.D. Boeke, *A genetic screen for ribosomal DNA silencing defects identifies multiple DNA replication and chromatin-modulating factors*. Mol Cell Biol, 1999. **19**(4): p. 3184-97.
106. Sun, Z.W. and M. Hampsey, *A general requirement for the Sin3-Rpd3 histone deacetylase complex in regulating silencing in Saccharomyces cerevisiae*. Genetics, 1999. **152**(3): p. 921-32.
107. Kozak, M.L., et al., *Inactivation of the Sas2 histone acetyltransferase delays senescence driven by telomere dysfunction*. EMBO J, 2010. **29**(1): p. 158-70.
108. Yang, B. and A.L. Kirchmaier, *Bypassing the catalytic activity of SIR2 for SIR protein spreading in Saccharomyces cerevisiae*. Mol Biol Cell, 2006. **17**(12): p. 5287-97.
109. Chou, C.C., Y.C. Li, and M.R. Gartenberg, *Bypassing Sir2 and O-acetyl-ADP-ribose in transcriptional silencing*. Mol Cell, 2008. **31**(5): p. 650-9.
110. Zhang, Y., et al., *SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex*. Mol Cell, 1998. **1**(7): p. 1021-31.
111. Lechner, T., et al., *Sds3 (suppressor of defective silencing 3) is an integral component of the yeast Sin3[middle dot]Rpd3 histone deacetylase complex and is required for histone deacetylase activity* 865. J Biol Chem, 2000. **275**(52): p. 40961-6.
112. Loewith, R., et al., *Pho23 is associated with the Rpd3 histone deacetylase and is required for its normal function in regulation of gene expression and silencing in Saccharomyces cerevisiae*. J Biol Chem, 2001. **276**(26): p. 24068-74.
113. Mallory, M.J. and R. Strich, *Ume1p represses meiotic gene transcription in Saccharomyces cerevisiae through interaction with the histone deacetylase Rpd3p*. J Biol Chem, 2003. **278**(45): p. 44727-34.
114. Puig, S., M. Lau, and D.J. Thiele, *Cti6 is an Rpd3-Sin3 histone deacetylase-associated protein required for growth under iron-limiting conditions in Saccharomyces cerevisiae*. J Biol Chem, 2004. **279**(29): p. 30298-306.
115. Colina, A.R. and D. Young, *Raf60, a novel component of the Rpd3 histone deacetylase complex required for Rpd3 activity in Saccharomyces cerevisiae*. J Biol Chem, 2005. **280**(52): p. 42552-6.
116. Carrozza, M.J., et al., *Stable incorporation of sequence specific repressors Ash1 and Ume6 into the Rpd3L complex*. Biochim Biophys Acta, 2005. **1731**(2): p. 77-87; discussion 75-6.
117. Carrozza, M.J., et al., *Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription*. Cell, 2005. **123**(4): p. 581-92.
118. Keogh, M.C., et al., *Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex*. Cell, 2005. **123**(4): p. 593-605.

119. Eisen, A., et al., *The yeast NuA4 and Drosophila MSL complexes contain homologous subunits important for transcription regulation*. J Biol Chem, 2001. **276**(5): p. 3484-91.
120. Brehm, A., et al., *The many colours of chromodomains*. Bioessays, 2004. **26**(2): p. 133-40.
121. Sims, R.J., 3rd, R. Belotserkovskaya, and D. Reinberg, *Elongation by RNA polymerase II: the short and long of it*. Genes Dev, 2004. **18**(20): p. 2437-68.
122. Joshi, A.A. and K. Struhl, *Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation*. Mol Cell, 2005. **20**(6): p. 971-8.
123. Li, B., et al., *Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin*. Science, 2007. **316**(5827): p. 1050-4.
124. Reid, J.L., Z. Moqtaderi, and K. Struhl, *Eaf3 regulates the global pattern of histone acetylation in Saccharomyces cerevisiae*. Mol Cell Biol, 2004. **24**(2): p. 757-64.
125. Mason, P.B. and K. Struhl, *The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo*. Mol Cell Biol, 2003. **23**(22): p. 8323-33.
126. Kaplan, C.D., L. Laprade, and F. Winston, *Transcription elongation factors repress transcription initiation from cryptic sites*. Science, 2003. **301**(5636): p. 1096-9.
127. Krogan, N.J., et al., *RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach*. Mol Cell Biol, 2002. **22**(20): p. 6979-92.
128. Lindstrom, D.L., et al., *Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins*. Mol Cell Biol, 2003. **23**(4): p. 1368-78.
129. Zhou, K., et al., *Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5*. Proc Natl Acad Sci U S A, 2009. **106**(17): p. 6956-61.
130. Liu, Y., et al., *Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex*. Mol Cell Biol, 2009. **29**(17): p. 4852-63.
131. Krogan, N.J., et al., *The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation*. Mol Cell, 2003. **11**(3): p. 721-9.
132. Krogan, N.J., et al., *Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II*. Mol Cell Biol, 2003. **23**(12): p. 4207-18.
133. Quan, T.K. and G.A. Hartzog, *Histone H3K4 and K36 methylation, Chd1 and Rpd3S oppose the functions of Saccharomyces cerevisiae Spt4-Spt5 in transcription*. Genetics, 2010. **184**(2): p. 321-34.
134. Watson, A.D., et al., *Ssn6-Tup1 interacts with class I histone deacetylases required for repression*. Genes Dev, 2000. **14**(21): p. 2737-44.
135. Davie, J.K., et al., *Tup1-Ssn6 interacts with multiple class I histone deacetylases in vivo*. J Biol Chem, 2003. **278**(50): p. 50158-62.

136. Pijnappel, W.W., et al., *The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program*. Genes Dev, 2001. **15**(22): p. 2991-3004.
137. Bilsland, E., M. Dahlen, and P. Sunnerhagen, *Genomic disruption of six budding yeast genes gives one drastic example of phenotype strain-dependence*. Yeast, 1998. **14**(7): p. 655-64.
138. Robyr, D., et al., *Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases*. Cell, 2002. **109**(4): p. 437-46.
139. Wang, A., S.K. Kurdistani, and M. Grunstein, *Requirement of Hos2 histone deacetylase for gene activity in yeast*. Science, 2002. **298**(5597): p. 1412-4.
140. Kim, T. and S. Buratowski, *Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions*. Cell, 2009. **137**(2): p. 259-72.
141. Carmen, A.A., S.E. Rundlett, and M. Grunstein, *HDA1 and HDA3 are components of a yeast histone deacetylase (HDA) complex*. J Biol Chem, 1996. **271**(26): p. 15837-44.
142. Wu, J., et al., *TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast*. Mol Cell, 2001. **7**(1): p. 117-26.
143. Wu, J., et al., *HDA2 and HDA3 are related proteins that interact with and are essential for the activity of the yeast histone deacetylase HDA1*. Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4391-6.
144. Green, S.R. and A.D. Johnson, *Promoter-dependent roles for the Srb10 cyclin-dependent kinase and the Hda1 deacetylase in Tup1-mediated repression in Saccharomyces cerevisiae*. Mol Biol Cell, 2004. **15**(9): p. 4191-202.
145. Carmen, A.A., et al., *Yeast HOS3 forms a novel trichostatin A-insensitive homodimer with intrinsic histone deacetylase activity*. Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12356-61.
146. Ahn, S.H., et al., *Histone H2B deacetylation at lysine 11 is required for yeast apoptosis induced by phosphorylation of H2B at serine 10*. Mol Cell, 2006. **24**(2): p. 211-20.
147. Landry, J., J.T. Slama, and R. Sternglanz, *Role of NAD(+) in the deacetylase activity of the SIR2-like proteins*. Biochem Biophys Res Commun, 2000. **278**(3): p. 685-90.
148. Shore, D., M. Squire, and K.A. Nasmyth, *Characterization of two genes required for the position-effect control of yeast mating-type genes*. EMBO J, 1984. **3**(12): p. 2817-23.
149. Rine, J. and I. Herskowitz, *Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae*. Genetics, 1987. **116**(1): p. 9-22.
150. Margolskee, J.P., *The sporulation capable (sca) mutation of Saccharomyces cerevisiae is an allele of the SIR2 gene*. Mol Gen Genet, 1988. **211**(3): p. 430-4.
151. Aparicio, O.M., B.L. Billington, and D.E. Gottschling, *Modifiers of position effect are shared between telomeric and silent mating-type loci in S. cerevisiae*. Cell, 1991. **66**(6): p. 1279-87.

152. Braunstein, M., et al., *Transcriptional silencing in yeast is associated with reduced nucleosome acetylation*. Genes Dev, 1993. **7**(4): p. 592-604.
153. Gottlieb, S. and R.E. Esposito, *A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA*. Cell, 1989. **56**(5): p. 771-6.
154. Shou, W., et al., *Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex*. Cell, 1999. **97**(2): p. 233-44.
155. Straight, A.F., et al., *Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity*. Cell, 1999. **97**(2): p. 245-56.
156. Moazed, D., et al., *Silent information regulator protein complexes in Saccharomyces cerevisiae: a SIR2/SIR4 complex and evidence for a regulatory domain in SIR4 that inhibits its interaction with SIR3*. Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2186-91.
157. Strahl-Bolsinger, S., et al., *SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast*. Genes Dev, 1997. **11**(1): p. 83-93.
158. Holmes, S.G., et al., *Hyperactivation of the silencing proteins, Sir2p and Sir3p, causes chromosome loss*. Genetics, 1997. **145**(3): p. 605-14.
159. Hoppe, G.J., et al., *Steps in assembly of silent chromatin in yeast: Sir3-independent binding of a Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation*. Mol Cell Biol, 2002. **22**(12): p. 4167-80.
160. Liou, G.G., et al., *Assembly of the SIR complex and its regulation by O-acetyl-ADP-ribose, a product of NAD-dependent histone deacetylation*. Cell, 2005. **121**(4): p. 515-27.
161. Kaeberlein, M., M. McVey, and L. Guarente, *The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms*. Genes Dev, 1999. **13**(19): p. 2570-80.
162. Kim, S., et al., *Modulation of life-span by histone deacetylase genes in Saccharomyces cerevisiae*. Mol Biol Cell, 1999. **10**(10): p. 3125-36.
163. Lin, S.J., P.A. Defossez, and L. Guarente, *Requirement of NAD and SIR2 for life-span extension by calorie restriction in Saccharomyces cerevisiae*. Science, 2000. **289**(5487): p. 2126-8.
164. Jiang, J.C., et al., *Distinct roles of processes modulated by histone deacetylases Rpd3p, Hda1p, and Sir2p in life extension by caloric restriction in yeast*. Exp Gerontol, 2002. **37**(8-9): p. 1023-30.
165. Lin, S.J., et al., *Calorie restriction extends Saccharomyces cerevisiae lifespan by increasing respiration*. Nature, 2002. **418**(6895): p. 344-8.
166. Anderson, R.M., et al., *Nicotinamide and PNCl govern lifespan extension by calorie restriction in Saccharomyces cerevisiae*. Nature, 2003. **423**(6936): p. 181-5.
167. Gallo, C.M., D.L. Smith, Jr., and J.S. Smith, *Nicotinamide clearance by Pnc1 directly regulates Sir2-mediated silencing and longevity*. Mol Cell Biol, 2004. **24**(3): p. 1301-12.
168. Lin, S.J., et al., *Calorie restriction extends yeast life span by lowering the level of NADH*. Genes Dev, 2004. **18**(1): p. 12-6.

169. Kaeberlein, M., et al., *Increased life span due to calorie restriction in respiratory-deficient yeast*. PLoS Genet, 2005. **1**(5): p. e69.
170. Tsuchiya, M., et al., *Sirtuin-independent effects of nicotinamide on lifespan extension from calorie restriction in yeast*. Aging Cell, 2006. **5**(6): p. 505-14.
171. Smith, D.L., Jr., et al., *Calorie restriction extends the chronological lifespan of *Saccharomyces cerevisiae* independently of the Sirtuins*. Aging Cell, 2007. **6**(5): p. 649-62.
172. Murakami, C.J., et al., *A method for high-throughput quantitative analysis of yeast chronological life span*. J Gerontol A Biol Sci Med Sci, 2008. **63**(2): p. 113-21.
173. Kaeberlein, M., et al., *Sir2-independent life span extension by calorie restriction in yeast*. PLoS Biol, 2004. **2**(9): p. E296.
174. Riesen, M. and A. Morgan, *Calorie restriction reduces rDNA recombination independently of rDNA silencing*. Aging Cell, 2009. **8**(6): p. 624-32.
175. Smith, D.L., Jr., et al., *Calorie restriction effects on silencing and recombination at the yeast rDNA*. Aging Cell, 2009. **8**(6): p. 633-42.
176. Brachmann, C.B., et al., *The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability*. Genes Dev, 1995. **9**(23): p. 2888-902.
177. Derbyshire, M.K., K.G. Weinstock, and J.N. Strathern, *HST1, a new member of the SIR2 family of genes*. Yeast, 1996. **12**(7): p. 631-40.
178. Xie, J., et al., *Sum1 and Hst1 repress middle sporulation-specific gene expression during mitosis in *Saccharomyces cerevisiae**. EMBO J, 1999. **18**(22): p. 6448-54.
179. Bedalov, A., et al., *NAD⁺-dependent deacetylase Hst1p controls biosynthesis and cellular NAD⁺ levels in *Saccharomyces cerevisiae**. Mol Cell Biol, 2003. **23**(19): p. 7044-54.
180. Vaquero, A., et al., *SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis*. Genes Dev, 2006. **20**(10): p. 1256-61.
181. Perrod, S., et al., *A cytosolic NAD-dependent deacetylase, Hst2p, can modulate nucleolar and telomeric silencing in yeast*. EMBO J, 2001. **20**(1-2): p. 197-209.
182. Durand-Dubief, M., et al., *Specific functions for the fission yeast Sirtuins Hst2 and Hst4 in gene regulation and retrotransposon silencing*. EMBO J, 2007. **26**(10): p. 2477-88.
183. Halme, A., et al., *Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast*. Cell, 2004. **116**(3): p. 405-15.
184. Wilson, J.M., et al., *Nuclear export modulates the cytoplasmic Sir2 homologue Hst2*. EMBO Rep, 2006. **7**(12): p. 1247-51.
185. Thaminy, S., et al., *Hst3 is regulated by Mec1-dependent proteolysis and controls the S phase checkpoint and sister chromatid cohesion by deacetylating histone H3 at lysine 56*. J Biol Chem, 2007. **282**(52): p. 37805-14.
186. Yang, B., A. Miller, and A.L. Kirchmaier, *HST3/HST4-dependent deacetylation of lysine 56 of histone H3 in silent chromatin*. Mol Biol Cell, 2008. **19**(11): p. 4993-5005.
187. Roper, S. and M. Esteller, *The role of histone deacetylases (HDACs) in human cancer*. Mol Oncol, 2007. **1**(1): p. 19-25.

188. de Ruijter, A.J., et al., *Histone deacetylases (HDACs): characterization of the classical HDAC family*. *Biochem J*, 2003. **370**(Pt 3): p. 737-49.
189. Yang, X.J. and E. Seto, *The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men*. *Nat Rev Mol Cell Biol*, 2008. **9**(3): p. 206-18.
190. Shevchenko, A., et al., *Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment*. *Genome Biol*, 2008. **9**(11): p. R167.
191. Cooper, K., *Rb, whi it's not just for metazoans anymore*. *Oncogene*, 2006. **25**(38): p. 5228-32.
192. Grandinetti, K.B. and G. David, *Sin3B: an essential regulator of chromatin modifications at E2F target promoters during cell cycle withdrawal*. *Cell Cycle*, 2008. **7**(11): p. 1550-4.
193. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1*. *Cell*, 2004. **119**(7): p. 941-53.
194. Lee, M.G., et al., *Functional interplay between histone demethylase and deacetylase enzymes*. *Mol Cell Biol*, 2006. **26**(17): p. 6395-402.
195. Hayakawa, T., et al., *RBP2 is an MRG15 complex component and down-regulates intragenic histone H3 lysine 4 methylation*. *Genes Cells*, 2007. **12**(6): p. 811-26.
196. Jelinic, P., J. Pellegrino, and G. David, *A novel mammalian complex containing Sin3B mitigates histone acetylation and RNAPII progression within transcribed loci*. *Mol Cell Biol*, 2010.
197. Lee, N., et al., *The H3K4 demethylase lid associates with and inhibits histone deacetylase Rpd3*. *Mol Cell Biol*, 2009. **29**(6): p. 1401-10.
198. Moshkin, Y.M., et al., *Histone chaperones ASF1 and NAPI1 differentially modulate removal of active histone marks by LID-RPD3 complexes during NOTCH silencing*. *Mol Cell*, 2009. **35**(6): p. 782-93.
199. Karagianni, P. and J. Wong, *HDAC3: taking the SMRT-N-CoRrect road to repression*. *Oncogene*, 2007. **26**(37): p. 5439-49.
200. Escaffit, F., et al., *Cleavage and cytoplasmic relocalization of histone deacetylase 3 are important for apoptosis progression*. *Mol Cell Biol*, 2007. **27**(2): p. 554-67.
201. Zhang, X., et al., *Histone deacetylase 3 (HDAC3) activity is regulated by interaction with protein serine/threonine phosphatase 4*. *Genes Dev*, 2005. **19**(7): p. 827-39.
202. Jeyakumar, M., et al., *Phosphorylation of thyroid hormone receptor-associated nuclear receptor corepressor holocomplex by the DNA-dependent protein kinase enhances its histone deacetylase activity*. *J Biol Chem*, 2007. **282**(13): p. 9312-22.
203. Longworth, M.S. and L.A. Laimins, *Histone deacetylase 3 localizes to the plasma membrane and is a substrate of Src*. *Oncogene*, 2006. **25**(32): p. 4495-500.
204. Oehme, I., et al., *Targeting of HDAC8 and investigational inhibitors in neuroblastoma*. *Expert Opin Investig Drugs*, 2009. **18**(11): p. 1605-17.
205. Martin, M., R. Kettmann, and F. Dequiedt, *Class IIa histone deacetylases: regulating the regulators*. *Oncogene*, 2007. **26**(37): p. 5450-67.

206. Lai, I.L., et al., *Histone deacetylase 10 relieves repression on the melanogenic program by maintaining the deacetylation status of repressors*. J Biol Chem, 2010. **285**(10): p. 7187-96.
207. Lee, J.H., et al., *Inhibition of histone deacetylase 10 induces thioredoxin-interacting protein and causes accumulation of reactive oxygen species in SNU-620 human gastric cancer cells*. Mol Cells, 2010. **30**(2): p. 107-12.
208. Valenzuela-Fernandez, A., et al., *HDAC6: a key regulator of cytoskeleton, cell migration and cell-cell interactions*. Trends Cell Biol, 2008. **18**(6): p. 291-7.
209. Afshar, G. and J.P. Murnane, *Characterization of a human gene with sequence homology to Saccharomyces cerevisiae SIR2*. Gene, 1999. **234**(1): p. 161-8.
210. Frye, R.A., *Characterization of five human cDNAs with homology to the yeast SIR2 gene: Sir2-like proteins (sirtuins) metabolize NAD and may have protein ADP-ribosyltransferase activity*. Biochem Biophys Res Commun, 1999. **260**(1): p. 273-9.
211. Jung-Hynes, B., R.J. Reiter, and N. Ahmad, *Sirtuins, melatonin and circadian rhythms: building a bridge between aging and cancer*. J Pineal Res, 2010. **48**(1): p. 9-19.
212. Bao, J. and M.N. Sack, *Protein deacetylation by sirtuins: delineating a post-translational regulatory program responsive to nutrient and redox stressors*. Cell Mol Life Sci, 2010. **67**(18): p. 3073-87.
213. Imai, S. and L. Guarente, *Ten years of NAD-dependent SIR2 family deacetylases: implications for metabolic diseases*. Trends Pharmacol Sci, 2010. **31**(5): p. 212-20.
214. Li, X., et al., *SIRT1 deacetylates and positively regulates the nuclear receptor LXR*. Mol Cell, 2007. **28**(1): p. 91-106.
215. Gerhart-Hines, Z., et al., *Metabolic control of muscle mitochondrial function and fatty acid oxidation through SIRT1/PGC-1alpha*. EMBO J, 2007. **26**(7): p. 1913-23.
216. Picard, F., et al., *Sirt1 promotes fat mobilization in white adipocytes by repressing PPAR-gamma*. Nature, 2004. **429**(6993): p. 771-6.
217. North, B.J. and E. Verdin, *Mitotic regulation of SIRT2 by cyclin-dependent kinase 1-dependent phosphorylation*. J Biol Chem, 2007. **282**(27): p. 19546-55.
218. Wang, F. and Q. Tong, *SIRT2 suppresses adipocyte differentiation by deacetylating FOXO1 and enhancing FOXO1's repressive interaction with PPARgamma*. Mol Biol Cell, 2009. **20**(3): p. 801-8.
219. Pandithage, R., et al., *The regulation of SIRT2 function by cyclin-dependent kinases affects cell motility*. J Cell Biol, 2008. **180**(5): p. 915-29.
220. Han, Y., et al., *Acetylation of Sirt2 by p300 attenuates its deacetylase activity*. Biochem Biophys Res Commun, 2008. **375**(4): p. 576-80.
221. Michishita, E., et al., *SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin*. Nature, 2008. **452**(7186): p. 492-6.
222. Yang, B., et al., *The sirtuin SIRT6 deacetylates H3 K56Ac in vivo to promote genomic stability*. Cell Cycle, 2009. **8**(16): p. 2662-3.
223. Mostoslavsky, R., et al., *Genomic instability and aging-like phenotype in the absence of mammalian SIRT6*. Cell, 2006. **124**(2): p. 315-29.
224. Grob, A., et al., *Involvement of SIRT7 in resumption of rDNA transcription at the exit from mitosis*. J Cell Sci, 2009. **122**(Pt 4): p. 489-98.

225. Vakhrusheva, O., et al., *Sirt7 increases stress resistance of cardiomyocytes and prevents apoptosis and inflammatory cardiomyopathy in mice*. *Circ Res*, 2008. **102**(6): p. 703-10.
226. Gregoret, I.V., Y.M. Lee, and H.V. Goodson, *Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis*. *J Mol Biol*, 2004. **338**(1): p. 17-31.
227. Villagra, A., et al., *The histone deacetylase HDAC11 regulates the expression of interleukin 10 and immune tolerance*. *Nat Immunol*, 2009. **10**(1): p. 92-100.
228. Liu, H., et al., *Histone deacetylase 11 regulates oligodendrocyte-specific gene expression and cell development in OL-1 oligodendroglia cells*. *Glia*, 2009. **57**(1): p. 1-12.
229. Jones, R.S. and W.M. Gelbart, *The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax*. *Mol Cell Biol*, 1993. **13**(10): p. 6357-66.
230. Nislow, C., E. Ray, and L. Pillus, *SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes*. *Mol Biol Cell*, 1997. **8**(12): p. 2421-36.
231. Porras-Yakushi, T.R., et al., *A novel SET domain methyltransferase modifies ribosomal protein Rpl23ab in yeast*. *J Biol Chem*, 2005. **280**(41): p. 34590-8.
232. Porras-Yakushi, T.R., J.P. Whitelegge, and S. Clarke, *A novel SET domain methyltransferase in yeast: Rkm2-dependent trimethylation of ribosomal protein L12ab at lysine 10*. *J Biol Chem*, 2006. **281**(47): p. 35835-45.
233. Webb, K.J., et al., *Identification of two SET domain proteins required for methylation of lysine residues in yeast ribosomal protein Rpl42ab*. *J Biol Chem*, 2008. **283**(51): p. 35561-8.
234. Miller, T., et al., *COMPASS: a complex of proteins associated with a trithorax-related SET domain protein*. *Proc Natl Acad Sci U S A*, 2001. **98**(23): p. 12902-7.
235. Roguev, A., et al., *The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4*. *EMBO J*, 2001. **20**(24): p. 7137-48.
236. Nagy, P.L., et al., *A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3*. *Proc Natl Acad Sci U S A*, 2002. **99**(1): p. 90-4.
237. Briggs, S.D., et al., *Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae*. *Genes Dev*, 2001. **15**(24): p. 3286-95.
238. Krogan, N.J., et al., *COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression*. *J Biol Chem*, 2002. **277**(13): p. 10753-5.
239. Fingerman, I.M., et al., *Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in Saccharomyces cerevisiae*. *J Biol Chem*, 2005. **280**(31): p. 28761-5.
240. Bernstein, B.E., et al., *Methylation of histone H3 Lys 4 in coding regions of active genes*. *Proc Natl Acad Sci U S A*, 2002. **99**(13): p. 8695-700.

241. Ng, H.H., et al., *Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity.* Mol Cell, 2003. **11**(3): p. 709-19.
242. Wood, A., et al., *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, 2003. **278**(37): p. 34739-42.
243. Ng, H.H., S. Dole, and K. Struhl, *The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B.* J Biol Chem, 2003. **278**(36): p. 33625-8.
244. Nakanishi, S., et al., *Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1.* J Cell Biol, 2009. **186**(3): p. 371-7.
245. Sun, Z.W. and C.D. Allis, *Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast.* Nature, 2002. **418**(6893): p. 104-8.
246. Dover, J., et al., *Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6.* J Biol Chem, 2002. **277**(32): p. 28368-71.
247. Dehe, P.M., et al., *Histone H3 lysine 4 mono-methylation does not require ubiquitination of histone H2B.* J Mol Biol, 2005. **353**(3): p. 477-84.
248. Xiao, T., et al., *The RNA polymerase II kinase Ctk1 regulates positioning of a 5' histone methylation boundary along genes.* Mol Cell Biol, 2007. **27**(2): p. 721-31.
249. Fischle, W., et al., *Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains.* Genes Dev, 2003. **17**(15): p. 1870-81.
250. Min, J., Y. Zhang, and R.M. Xu, *Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27.* Genes Dev, 2003. **17**(15): p. 1823-8.
251. Shi, X., et al., *ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression.* Nature, 2006. **442**(7098): p. 96-9.
252. Wysocka, J., et al., *A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling.* Nature, 2006. **442**(7098): p. 86-90.
253. Pray-Grant, M.G., et al., *Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation.* Nature, 2005. **433**(7024): p. 434-8.
254. Shi, X., et al., *Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36.* J Biol Chem, 2007. **282**(4): p. 2450-5.
255. Muramoto, T., et al., *Methylation of H3K4 Is required for inheritance of active transcriptional states.* Curr Biol, 2010. **20**(5): p. 397-406.
256. Strahl, B.D., et al., *Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression.* Mol Cell Biol, 2002. **22**(5): p. 1298-306.
257. Li, B., et al., *The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II.* J Biol Chem, 2003. **278**(11): p. 8897-903.
258. Schaft, D., et al., *The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation.* Nucleic Acids Res, 2003. **31**(10): p. 2475-82.

259. Xiao, T., et al., *Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast*. *Genes Dev*, 2003. **17**(5): p. 654-63.
260. Nelson, C.J., H. Santos-Rosa, and T. Kouzarides, *Proline isomerization of histone H3 regulates lysine methylation and gene expression*. *Cell*, 2006. **126**(5): p. 905-16.
261. Youdell, M.L., et al., *Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36*. *Mol Cell Biol*, 2008. **28**(16): p. 4915-26.
262. Lin, L.J., et al., *Asf1 can promote trimethylation of H3 K36 by Set2*. *Mol Cell Biol*, 2010. **30**(5): p. 1116-29.
263. Singer, M.S., et al., *Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae*. *Genetics*, 1998. **150**(2): p. 613-32.
264. San-Segundo, P.A. and G.S. Roeder, *Role for the silencing protein Dot1 in meiotic checkpoint control*. *Mol Biol Cell*, 2000. **11**(10): p. 3601-15.
265. Feng, Q., et al., *Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain*. *Curr Biol*, 2002. **12**(12): p. 1052-8.
266. Lacoste, N., et al., *Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase*. *J Biol Chem*, 2002. **277**(34): p. 30421-4.
267. Ng, H.H., et al., *Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association*. *Genes Dev*, 2002. **16**(12): p. 1518-27.
268. Park, J.H., et al., *A core nucleosome surface crucial for transcriptional silencing*. *Nat Genet*, 2002. **32**(2): p. 273-9.
269. van Leeuwen, F., P.R. Gafken, and D.E. Gottschling, *Dot1p modulates silencing in yeast by methylation of the nucleosome core*. *Cell*, 2002. **109**(6): p. 745-56.
270. Ng, H.H., et al., *Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: a potential mechanism for position-effect variegation*. *Proc Natl Acad Sci U S A*, 2003. **100**(4): p. 1820-5.
271. Verzijlbergen, K.F., et al., *Multiple histone modifications in euchromatin promote heterochromatin formation by redundant mechanisms in Saccharomyces cerevisiae*. *BMC Mol Biol*, 2009. **10**: p. 76.
272. Game, J.C., M.S. Williamson, and C. Baccari, *X-ray survival characteristics and genetic analysis for nine Saccharomyces deletion mutants that show altered radiation sensitivity*. *Genetics*, 2005. **169**(1): p. 51-63.
273. Wysocki, R., et al., *Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkpoint functions of Rad9*. *Mol Cell Biol*, 2005. **25**(19): p. 8430-43.
274. Toh, G.W., et al., *Histone H2A phosphorylation and H3 methylation are required for a novel Rad9 DSB repair function following checkpoint activation*. *DNA Repair (Amst)*, 2006. **5**(6): p. 693-703.
275. Bannister, A.J., R. Schneider, and T. Kouzarides, *Histone methylation: dynamic or static?* *Cell*, 2002. **109**(7): p. 801-6.
276. Morillon, A., et al., *Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription*. *Mol Cell*, 2005. **18**(6): p. 723-34.
277. Tsukada, Y., et al., *Histone demethylation by a family of JmjC domain-containing proteins*. *Nature*, 2006. **439**(7078): p. 811-6.

278. Fang, J., et al., *The Saccharomyces cerevisiae histone demethylase Jhd1 fine-tunes the distribution of H3K36me2*. Mol Cell Biol, 2007. **27**(13): p. 5055-65.
279. Klose, R.J., et al., *Demethylation of histone H3K36 and H3K9 by Rph1: a vestige of an H3K9 methylation system in Saccharomyces cerevisiae?* Mol Cell Biol, 2007. **27**(11): p. 3951-61.
280. Liang, G., et al., *Yeast Jhd2p is a histone H3 Lys4 trimethyl demethylase*. Nat Struct Mol Biol, 2007. **14**(3): p. 243-5.
281. Tu, S., et al., *Identification of histone demethylases in Saccharomyces cerevisiae*. J Biol Chem, 2007. **282**(19): p. 14262-71.
282. Mersman, D.P., et al., *Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression*. Genes Dev, 2009. **23**(8): p. 951-62.
283. Radman-Livaja, M., et al., *Replication and active demethylation represent partially overlapping mechanisms for erasure of H3K4me3 in budding yeast*. PLoS Genet, 2010. **6**(2): p. e1000837.
284. Kuo, M.H. and C.D. Allis, *In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment*. Methods, 1999. **19**(3): p. 425-33.
285. Ren, B., et al., *Genome-wide location and function of DNA binding proteins*. Science, 2000. **290**(5500): p. 2306-9.
286. Iyer, V.R., et al., *Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF*. Nature, 2001. **409**(6819): p. 533-8.
287. Hegde, P., et al., *Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray*. Cancer Res, 2001. **61**(21): p. 7792-7.
288. Lee, T.I., et al., *Transcriptional regulatory networks in Saccharomyces cerevisiae*. Science, 2002. **298**(5594): p. 799-804.
289. Drouin, S., et al., *DSIF and RNA Polymerase II CTD Phosphorylation Coordinate the Recruitment of Rpd3S to Actively Transcribed Genes*. PLoS Genet, 2010. **6**(10): p. e1001173.
290. Shahbazian, M.D. and M. Grunstein, *Functions of site-specific histone acetylation and deacetylation*. Annu.Rev.Biochem., 2007. **76**:75-100.: p. 75-100.
291. Yang, X.J. and E. Seto, *The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men*. Nat.Rev.Mol.Cell Biol., 2008. **9**(3): p. 206-218.
292. Wang, H., et al., *Recruitment of Cln3 cyclin to promoters controls cell cycle entry via histone deacetylase and other targets*. PLoS.Biol., 2009. **7**(9): p. e1000189.
293. Kremer, S.B. and D.S. Gross, *SAGA and Rpd3 chromatin modification complexes dynamically regulate heat shock gene structure and expression*. J Biol Chem, 2009. **284**(47): p. 32914-31.
294. Veis, J., et al., *Activation of the G2/M-specific gene CLB2 requires multiple cell cycle signals*. Mol Cell Biol, 2007. **27**(23): p. 8364-73.
295. Sharma, V.M., et al., *Histone deacetylases RPD3 and HOS2 regulate the transcriptional activation of DNA damage-inducible genes*. Mol Cell Biol, 2007. **27**(8): p. 3199-210.

296. Inai, T., M. Yukawa, and E. Tsuchiya, *Interplay between chromatin and trans-acting factors on the IME2 promoter upon induction of the gene at the onset of meiosis*. Mol Cell Biol, 2007. **27**(4): p. 1254-63.
297. Washburn, B.K. and R.E. Esposito, *Identification of the Sin3-binding site in Ume6 defines a two-step process for conversion of Ume6 from a transcriptional repressor to an activator in yeast*. Mol Cell Biol, 2001. **21**(6): p. 2057-69.
298. Kurdistani, S.K., et al., *Genome-wide binding map of the histone deacetylase Rpd3 in yeast*. Nat Genet, 2002. **31**(3): p. 248-54.
299. Fazio, T.G., et al., *Widespread collaboration of Isw2 and Sin3-Rpd3 chromatin remodeling complexes in transcriptional repression*. Mol Cell Biol, 2001. **21**(19): p. 6450-60.
300. Zhou, J., et al., *Histone deacetylase Rpd3 antagonizes Sir2-dependent silent chromatin propagation*. Nucleic Acids Res, 2009. **37**(11): p. 3699-713.
301. Sandmeier, J.J., et al., *RPD3 is required for the inactivation of yeast ribosomal DNA genes in stationary phase*. EMBO J, 2002. **21**(18): p. 4959-68.
302. Oakes, M.L., et al., *Role of histone deacetylase Rpd3 in regulating rRNA gene transcription and nucleolar structure in yeast*. Mol Cell Biol, 2006. **26**(10): p. 3889-901.
303. Knott, S.R., et al., *Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in Saccharomyces cerevisiae*. Genes Dev, 2009. **23**(9): p. 1077-90.
304. Lamb, T.M. and A.P. Mitchell, *Coupling of Saccharomyces cerevisiae early meiotic gene expression to DNA replication depends upon RPD3 and SIN3*. Genetics, 2001. **157**(2): p. 545-56.
305. Merker, J.D., et al., *The histone methylase Set2p and the histone deacetylase Rpd3p repress meiotic recombination at the HIS4 meiotic recombination hotspot in Saccharomyces cerevisiae*. DNA Repair (Amst), 2008. **7**(8): p. 1298-308.
306. Lickwar, C.R., et al., *The Set2/Rpd3S pathway suppresses cryptic transcription without regard to gene length or transcription frequency*. PLoS One, 2009. **4**(3): p. e4886.
307. Li, B., et al., *Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription*. Genes Dev, 2007. **21**(11): p. 1422-30.
308. Lee, J.S. and A. Shilatifard, *A site to remember: H3K36 methylation a mark for histone deacetylation*. Mutat Res, 2007. **618**(1-2): p. 130-4.
309. Biswas, D., S. Takahata, and D.J. Stillman, *Different genetic functions for the Rpd3(L) and Rpd3(S) complexes suggest competition between NuA4 and Rpd3(S)*. Mol Cell Biol, 2008. **28**(14): p. 4445-58.
310. Wada, T., et al., *DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs*. Genes Dev, 1998. **12**(3): p. 343-56.
311. Price, D.H., *Poised polymerases: on your mark...get set...go!* Mol.Cell., 2008. **30**(1): p. 7-10.

312. Saunders, A., L.J. Core, and J.T. Lis, *Breaking barriers to transcription elongation*. *Nat.Rev.Mol.Cell Biol.*, 2006. **7**(8): p. 557-567.
313. Cui, Y. and C.L. Denis, *In vivo evidence that defects in the transcriptional elongation factors RPB2, TFIIIS, and SPT5 enhance upstream poly(A) site utilization*. *Mol Cell Biol*, 2003. **23**(21): p. 7887-901.
314. Bucheli, M.E. and S. Buratowski, *Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II*. *EMBO J*, 2005. **24**(12): p. 2150-60.
315. Kaplan, C.D., M.J. Holland, and F. Winston, *Interaction between transcription elongation factors and mRNA 3'-end formation at the *Saccharomyces cerevisiae* GAL10-GAL7 locus*. *J Biol Chem*, 2005. **280**(2): p. 913-22.
316. Govind, C.K., et al., *Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes*. *Mol Cell*, 2010. **39**(2): p. 234-46.
317. Qiu, H., C. Hu, and A.G. Hinnebusch, *Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters*. *Mol.Cell.*, 2009. **33**(6): p. 752-762.
318. Chu, Y., et al., *The BUR1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by SET2*. *Mol Cell Biol*, 2006. **26**(8): p. 3029-38.
319. Peterlin, B.M. and D.H. Price, *Controlling the elongation phase of transcription with P-TEFb*. *Mol.Cell*, 2006. **23**(3): p. 297-305.
320. Wood, A., et al., *The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS*. *Mol.Cell.*, 2005. **20**(4): p. 589-599.
321. Keogh, M.C., V. Podolny, and S. Buratowski, *Bur1 kinase is required for efficient transcription elongation by RNA polymerase II*. *Mol Cell Biol*, 2003. **23**(19): p. 7005-18.
322. Qiu, H., C. Hu, and A.G. Hinnebusch, *Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters*. *Mol Cell*, 2009. **33**(6): p. 752-62.
323. Ng, H.H., et al., *Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79931*. *J Biol Chem*, 2002. **277**(38): p. 34655-7.
324. Berretta, J. and A. Morillon, *Pervasive transcription constitutes a new level of eukaryotic genome regulation*. *EMBO Rep.*, 2009. **10**(9): p. 973-982.
325. Gavin, A.C., et al., *Proteome survey reveals modularity of the yeast cell machinery*. *Nature*, 2006. **440**(7084): p. 631-636.
326. Keogh, M.C., et al., *The *Saccharomyces cerevisiae* histone H2A variant Htz1 is acetylated by NuA4*. *Genes Dev*, 2006. **20**(6): p. 660-5.
327. Aygun, O., J. Svejstrup, and Y. Liu, *A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin*. *Proc.Natl.Acad.Sci.U.S.A.*, 2008. **105**(25): p. 8580-8584.
328. Lambert, J.P., et al., *A novel proteomics approach for the discovery of chromatin-associated protein networks*. *Mol.Cell Proteomics.*, 2009. **8**(4): p. 870-882.

329. Lavoie, S.B., et al., *The peptidyl-prolyl isomerase Pin1 interacts with hSpt5 phosphorylated by Cdk9*. J.Mol.Biol., 2001. **312**(4): p. 675-685.
330. Arevalo-Rodriguez, M., et al., *Cyclophilin A and Ess1 interact with and regulate silencing by the Sin3-Rpd3 histone deacetylase*. EMBO J, 2000. **19**(14): p. 3739-49.
331. Wada, T., et al., *Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro*. EMBO J., 1998. **17**(24): p. 7395-7403.
332. Lindstrom, D.L. and G.A. Hartzog, *Genetic interactions of Spt4-Spt5 and TFIIIS with the RNA polymerase II CTD and CTD modifying enzymes in Saccharomyces cerevisiae*. Genetics, 2001. **159**(2): p. 487-97.
333. Viladevall, L., et al., *TFIIH and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast*. Mol.Cell., 2009. **33**(6): p. 738-751.
334. Qiu, H., et al., *The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II*. Mol Cell Biol, 2006. **26**(8): p. 3135-3148.
335. Hogan, G.J., C.K. Lee, and J.D. Lieb, *Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters*. PLoS.Genet., 2006. **2**(9): p. e158.
336. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns* 1113. Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14863-8.
337. Saldanha, A.J., *Java Treeview--extensible visualization of microarray data*. Bioinformatics., 2004. **20**(17): p. 3246-3248.
338. Hirao, M., et al., *Identification of selective inhibitors of NAD⁺-dependent deacetylases using phenotypic screens in yeast*. J Biol Chem, 2003. **278**(52): p. 52773-82.
339. Cohen, T.J., et al., *Hos2p/Set3p deacetylase complex signals secretory stress through the Mpk1p cell integrity pathway*. Eukaryot Cell, 2008. **7**(7): p. 1191-9.
340. Mou, Z., A.E. Kenny, and M.J. Curcio, *Hos2 and Set3 promote integration of Ty1 retrotransposons at tRNA genes in Saccharomyces cerevisiae*. Genetics, 2006. **172**(4): p. 2157-67.
341. Pokholok, D.K., et al., *Genome-wide map of nucleosome acetylation and methylation in yeast*. Cell, 2005. **122**(4): p. 517-27.

