



Male genome programming guided by histone acylations

Afsaneh Goudarzi

► To cite this version:

Afsaneh Goudarzi. Male genome programming guided by histone acylations. Biochemistry [q-bio.BM]. Université Grenoble Alpes, 2016. English. NNT : 2016GREAV059 . tel-01690716

HAL Id: tel-01690716
<https://theses.hal.science/tel-01690716>

Submitted on 23 Jan 2018

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

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

THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ
GRENOBLE ALPES**

Spécialité : **CHIMIE BIOLOGIE**

Arrêté ministériel : 25 mai 2016

Présentée par

Afsaneh GOUDARZI

Thèse dirigée par **Dr. Saadi KHOCHBIN** et
codirigée par **Dr. Sophie ROUSSEAUX**

préparée au sein du **CENTRE DE RECHERCHE UGA- INSERM U
1209- CNRS UMR 5309/ Institute for Advanced Bioscience
(IAB)**
dans l'**École Doctorale Chimie et Science du Vivant**

Male genome programming guided by histone acylations

Thèse soutenue publiquement le **22 November 2016**,
devant le jury composé de :

Pr Stefan NONCHEV

Université Grenoble Alpes, Grenoble

Président

Dr Matthieu GERARD

Directeur de recherche, Paris

Rapporteur

Dr Slimane AIT-SI-ALI

Directeur de recherche, Paris

Rapporteur

Dr Saadi KHOCHBIN

Directeur de recherche, Grenoble

Examineur

Table of contents

Abstract	4
Résumé.....	5
List of main abbreviations	6
Introduction.....	8
1- General aspects of chromatin	9
1-1- The history of chromatin discovery.....	9
1-2- Chromatin components.....	11
1-2-1- DNA.....	11
1-2-2- Nucleosome.....	11
1-3- Chromatin structure and function.....	14
1-3-1- Hierarchical structure of chromatin.....	14
1-4- Chromatin domains: Euchromatin and Heterochromatin.....	16
1-5- The dynamic nature of chromatin	17
1-5-1- Histone post-translational modifications (PTMs).....	17
1-5-2- Histone chaperones	17
1-5-3- ATP-dependent chromatin remodelers.....	20
1-5-4- Histone variants.....	20
1-5-4-1- H2A variants:.....	21
1-5-4-1-1- H2A.Z:	21
1-5-4-1-2- H2A.X:.....	22
1-5-4-1-3- Macro.H2A:.....	22
1-5-4-1-4- H2A.Bbd:	22
1-5-4-2- H2B variants:	23
1-5-4-3- H3 variants:	23
2- Overview of epigenetics.....	24
2-1- A brief history of evolving landscape of epigenetics	24
2-2- Epigenetics landscape	26
2-2-1- Chromatin modifications.....	26
2-2-1-1- Histone modifications	26
2-2-1-2- DNA modifications.....	28
2-2-2- Epigenetic regulators.....	29
2-2-2-1- Histone acetyltransferase (HAT).....	31
2-2-2-1-1- The GNAT family	31
2-2-2-1-2- The MYST family	31

2-2-2-1-3- The p300/CBP family	31
2-2-2-2- Histone deacetylase (HDAC)	33
3- Discovery of histone lysine acylation.....	34
3-1- Well-known histone modifications.....	35
3-1-1- Acetylation.....	35
3-1-2- Methylation.....	36
3-1-3- Phosphorylation	36
3-2- Recently identified histone modifications.....	37
3-2-1- Crotonylation	37
3-2-2- 2-hydroxyisobutyrylation	38
3-2-3- Butyrylation.....	40
3-2-4- β -hydroxybutyrylation	40
3-2-5- Succinylation, Malonylation, and Glutarylation	41
3-2-6- Propionylation	42
4- Mammalian spermatogenesis	44
4-1- Testis organization	44
4-2- Phases of spermatogenesis.....	45
4-2-1- Mitotic phase.....	46
4-2-2- Meiotic phase.....	46
4-2-3- Post-Meiotic phase.....	49
4-3- Stages of spermatogenesis.....	50
5- Chromatin structure and gene expression during spermatogenesis	52
5-1- Gene expression regulation.....	53
5-1-1- Meiotic sex chromosome inactivation.....	54
5-2- Chromatin remodeling and genome wide histone removal.....	56
5-2-1- Histone modifications	58
5-2-2- Brdt mediates histone removal.....	59
5-2-3- Histone variants.....	60
5-2-4- Transition proteins	61
5-2-5- Protamines.....	62
6- Fertilization and zygote formation.....	63
Thesis objectives.....	65
Results	67
1- Role of CBP and p300 in male gene expression program and histone hyperacetylation wave.	68
1-1- Expression of CBP/p300 during spermatogenesis	68

1-2- Role of the HATs CBP/p300 in the control of the histone hyperacetylation wave preceding histone removal	68
1-3- Role the the HATs CBP/p300 in controlling male germ cells gene expression programs	69
2- Investigation of butyrylation in parallel with acetylation on lysines 5 and 8 of histone H4 during spermatogenesis.	82
2-1- Genomic distribution of histone H4 K5K8 butyrylation and acetylation	82
2-2- Histone butyrylation activates transcription but inhibits the binding of Brdt	83
3- p300-catalyzed histone acylations	110
3-1- p300 discriminates between different acyl-CoAs	110
Discussion and Perspectives	153
- Genome-wide mapping of butyrylation and acetylation on H4K5 and H4K8 designates butyrylation as an active histone mark	154
- Regulation of gene expression and p300 activity: toward differential histone lysine acylations via metabolism	156
- Histone modifications alternation could impact gene targeting by chromatin binding factors	157
- CBP/p300 is involved in establishing the male epigenome and the corresponding gene expression programs	159
- The HAT p300 catalyzes differential histone acylation with the ability to discriminate between acyl-CoAs	160
General Conclusion	162
References	166
Annexes	177

Abstract

Male genome programming guided by histone acylations

The main focus of the investigations reported in this manuscript is the understanding of the regulatory events that are based on histone lysine modifications in post-meiotic male germ cells, where specific and chromosome-wide regulations of gene expression occur. In the first part of my work we designed a strategy to specifically investigate the role of the histone acetyl-transferases (HATs), Cbp and p300, in post-meiotic male germ cells. Accordingly, we generated double *Cbp* and *p300* conditional knock-out mice resulting in a partial depletion of *Cbp* and *p300* in post-meiotic cells. Although the mice were fertile and spermatogenesis seemed to take place normally, a transcriptomic analysis of early and late post-meiotic germ cells led to the identification of a specific subset of genes with an increased expression in late spermatogenic cells that is highly sensitive to the decreased amounts of Cbp and p300. In conclusion, these results have revealed an interesting new gene expression program specific to post-meiotic male germ cells that are specifically regulated by the considered HATs. Taking into account the occurrence of a variety of histone lysine acylations, we extended these investigations to a four-carbon histone lysine modification, butyrylation. Accordingly, we have undertaken a comprehensive comparative analysis of histone H4 acetylation and butyrylation on its K5 and K8 positions in differentiating male germ cells. Genome-wide mapping of H4K5ac, H4K5bu, H4K8ac and H4K8bu at two critical developmental stages, meiotic and post-meiotic haploid cells, shows an interchangeable use of acetylation and butyrylation in the Transcriptional Start Sites (TSSs) of the most highly expressed genes in both meiotic and haploid round spermatids. Interestingly, many of these promoters are also bound by the essential regulator of spermatogenic gene expression, the BET bromodomain-containing factor, Brdt. A detailed analysis of Brdt binding capacity of H4 tails bearing various combinations of K5 and K8 acetylation and butyrylation showed that H4K5 butyrylation severely interferes with Brdt-binding. Our results therefore indicate that not only Brdt is required for the activation of a meiotic and post-meiotic gene expression program, but also its turnover induced by H4K5 butyrylation is equally important. This work hence highlights how an interplay between two different acylations occurring on the same lysines can play an essential regulatory role by increasing the chromatin binding dynamics of a critical lysine acetyl-reader, Brdt. Finally, in a collaborative work with structural biologists we showed that while p300 is a robust acetylase, its activity gets weaker with increasing acyl chain length. These results suggest that *in vivo*, p300 would use a specific co-factor to ensure non-acetyl histone acylations.

Overall, these investigations shed an important light on how the male genome programming is guided by histone acylations and revealed for the first time a molecular network that regulates histone acylations and mediates its functional impact.

Chromatin, histone post-translational modification, Brdt, CBP, p300, gene expression, bio-informatics, proteomics, genomics.

Résumé

Les acylations des histones guident la programmation du génome mâle

Le principal intérêt de nos études présentées dans ce manuscrit, correspond à la compréhension des événements, reliés aux acylations d'histones au niveau des lysines (Ks) dans les cellules germinales post méiotiques, qui régulent spécifiquement l'expression des gènes à l'échelle du génome entier. Dans la première partie de mon travail, nous avons élaboré une stratégie afin d'analyser le rôle des « histones acetyl transférase » (HATs), Cbp et p300, dans les cellules germinales post méiotiques. Pour ce faire, nous avons généré une lignée de souris conditionnellement et partiellement invalidée pour les gènes *Cbp* et *p300* dans les cellules post méiotiques. Bien que les souris mâles sont fertiles et que la spermatogénèse semble se dérouler normalement, une analyse transcriptomique des cellules germinales haploïdes post méiotiques précoces et tardives nous a permis d'identifier une série de gènes dont l'expression est augmentée dans les cellules spermatogéniques tardives et qui sont sensibles à la diminution des niveaux de Cbp et p300. Ces résultats ont permis de révéler un programme spécifique d'expression de gènes dans les cellules germinales post méiotiques dépendant des HATs correspondantes. Prenant en compte qu'il existe une variété d'acylations des histones au niveau des lysines, nous avons étendu nos études à une modification à « quatre-carbone », la butyrylation. Nous avons alors initié une analyse comparative de l'acétylation et de la butyrylation de l'histone H4 en positions K5 et K8 dans les cellules germinales mâles en différenciation. Nous avons cartographié à l'échelle du génome les marques H4K5ac, H4K5bu, H4K8ac, et H4K8bu au niveau de deux étapes développementales critiques avec les cellules méiotiques et les cellules post méiotiques haploïdes. Cette cartographie montre que la majorité des gènes exprimés fortement, à la fois dans les cellules méiotiques et les spermatides précoces rondes haploïdes, qu'au niveau des sites d'initiation de la transcription (TSSs) l'acétylation et la butyrylation sont interchangeables. De façon intéressante, beaucoup de ces promoteurs correspondants sont aussi reconnus par un régulateur essentiel de l'expression des gènes lors de la spermatogénèse, le facteur à bromodomaine, Brdt. Une étude détaillée, des capacités de liaison du facteur Brdt sur les parties N-terminales de l'histone H4 portant des combinaisons variées d'acétylation et/ou de butyrylation en position K5 et K8, montre que la marque H4K5bu inhibe fortement la liaison du facteur Brdt. Nos résultats suggèrent qu'en addition à la fonction activatrice de Brdt vis-à-vis du programme d'expression de gènes méiotiques et post méiotiques, l'échange (« turnover ») induit par la butyrylation d'H4K5 est également important. Ce travail montre comment une interconnexion entre deux différentes acylations d'une même lysine peut jouer un rôle régulateur essentiel en augmentant la dynamique de liaison de la chromatine par un lecteur de lysine acétylé, Brdt. Enfin, au cours d'un travail collaboratif portant sur des approches structurales, nous avons montré, malgré le fait que p300 soit répertoriée comme une acétylase robuste, que son activité est réduite lorsque la longueur des chaînes acyl augmente. Ces résultats suggèrent qu'*in vivo*, p300 puisse utiliser un co-facteur spécifique pour assurer des acylations d'histones autre qu'une acétylation. Ces investigations mettent en lumière comment la programmation du génome mâle est guidée par diverses acylations d'histone et révèlent pour la première fois l'existence d'un réseau moléculaire qui régule ces acylations et transmet un impact fonctionnel.

Chromatine, modification post traductionnel d'histone, Brdt, Cbp, p300, expression génique, bioinformatique, protéomique, génomique.

List of main abbreviations

ACL: ATP-citrate lyase
ACSS2: Acyl-CoA Synthetase
bp: Base pair
Brdt: Bromodomain testis-specific protein
CAF1: Chromatin assembly factor-1
CBP: CREB-binding protein
CENP-A: centromere protein-A
CHD: Chromodomain helicase DNA-binding
ChIP: Chromatin Immunoprecipitation
CREB: cAMP response element-binding protein
DBDs: DNA binding domains
DNA: Deoxyribonucleic acid
DSBs: DNA double-strand breaks
FACT: Facilitate Chromatin Transcription
Gcn5: General control non-derepressible 5
GNATs: Gcn5-related N-acetyltransferases
HATs: Histone acetyltransferases
HBO1: HAT bound to Orc1
HDACs: Histone deacetylases
HIRA: Histone regulator A
HMTs: Histone methyltransferases
H1s: H1 and its variants
HP1: Heterochromatin protein-1
ISWI: Imitation switch
INO80: Inositol requiring 80
Kbhb: Lysine β -hydroxybutyrylation
Kbu: Lysine butyrylation
Kcr: Lysine crotonylation
KDMs: lysine demethylases
Kglu: Lysine glutarylation
K_{hib}: Hydroxyisobutyrylation
Kma: Lysine malonylation
Kpr: Lysine propionylation
Ksu: Lysine succinylation
K_{2ohbu}: Lysine 2-hydroxybutyryl
K_{3ohbu}: Lysine 3-hydroxybutyryl
K_{3ohibu}: Lysine 3-hydroxyisobutyryl
K_{4ohbu}: Lysine 4-hydroxybutyryl
LSD1: Lysine-specific demethylase 1
MCD: Malonyl-CoA decarboxylase
Mecp2: Methyl-CpG binding protein 2

MOF: Males absent on the first
Morf: MOZ-related factor
MOZ: Monocytic leukemia zinc finger protein
MSCI: Meiotic sex chromosome inactivation
NAD: Nicotinamide-adenine-dinucleotide
Nap1: Nucleosome Assembly Protein-1
NASP: Nuclear Autoantigenic Sperm Protein
NRE: Negative regulatory element
PCC: Propionyl-CoA carboxylase
PRBP: Protamine 1 RNA binding protein
Prms: Protamines
PRMTs: Protein arginine methyltransferase
PTMs: Post-translational modifications
p300: E1A-associated protein of 300 kDa
SCAD: Short-chain acyl-CoA dehydrogenase
SCFAs: Short chain fatty acids
SWI/SNF: Switching defective/sucrose non fermenting
TBP: TATA-binding protein
Tet: Ten-eleven translocation
Tip60: Tat-interacting protein of 60 kDa
TPs: Transition proteins
TSA: Trichostatin A
TSSs: Transcriptional start site regions
UTR: Untranslated regions
YEATS: (Yaf9, ENL, AF9, Taf14, Sas5)
5mC: Methylation of the cytosine residues at their 5 position
5hmC: 5-Hydroxymethyl cytosine
5fC: 5-formyl cytosine
5caC: 5-carboxyl cytosine

Introduction

1- General aspects of chromatin

1-1- The history of chromatin discovery

Eukaryotic genome is composed of several chromosomes. In the cell nucleus, each chromosome contains a single molecule of deoxyribonucleic acid (DNA). DNA of eukaryotic cells is associated with several proteins, histones and non-histone proteins, that pack the DNA molecule into a structure known as chromatin. Chromatin not only condenses the DNA filament, but it also protects the genome and contributes to the regulation of gene expression as well as to many other processes depending on DNA. Chromatin, which is basically a complex formed of DNA, protein and RNA has a very dynamic structure, necessary to regulate the accessibility of DNA to ensure its various functions including transcription, recombination, DNA repair and replication (Venkatesh and Workman, 2015).

Our current conception of chromatin owes to many researchers whose works progressively resulted in the discovery of important features of chromatin during the last several decades (Figure 1). The very important discoveries of chromatin aspects will be very briefly highlighted and emphasized. In 1831 a Scottish botanist called Robert Brown used the term “Nucleus” to describe the opaque spot that he had observed in the epidermis of orchids under microscope. Later in 1879, the discovery of chromosome by Fleming enabled Wilson and Boveri (in 1902-1903) to argue that developmental processes are directed by the chromosomes (Felsenfeld, 2014). Following the discovery of chromosomes, the study from H.J Muller in 1930 on *Drosophila* reported that the difference between molted eyes in mutant flies compared to red eyes in wild type flies is due to a chromosome translocation. This observation led him to report a link between the activity of a gene and its location within the chromosome (Felsenfeld, 2014). Following the identification of proteins bound to DNA in the eukaryotic cell nucleus, named histones, in 1950, Stedman and Stedman considered that histones are capable of silencing gene expression and concluded that, depending on the different cell types, the presence of different kinds of histones give rise to different phenotypes. In that time, although it was believed that an active gene would be completely histone free, it was not clear how histones could be erased from transcriptionally active regions. Some years later, Allfrey and Mirsky’s speculated about histone acetylation and gene expression and argued that active chromatin regions could also be covered with histones but that these histones could be chemically modified. Since then, the study of the relationships between histone modifications and gene expression attracted a lot of attention.

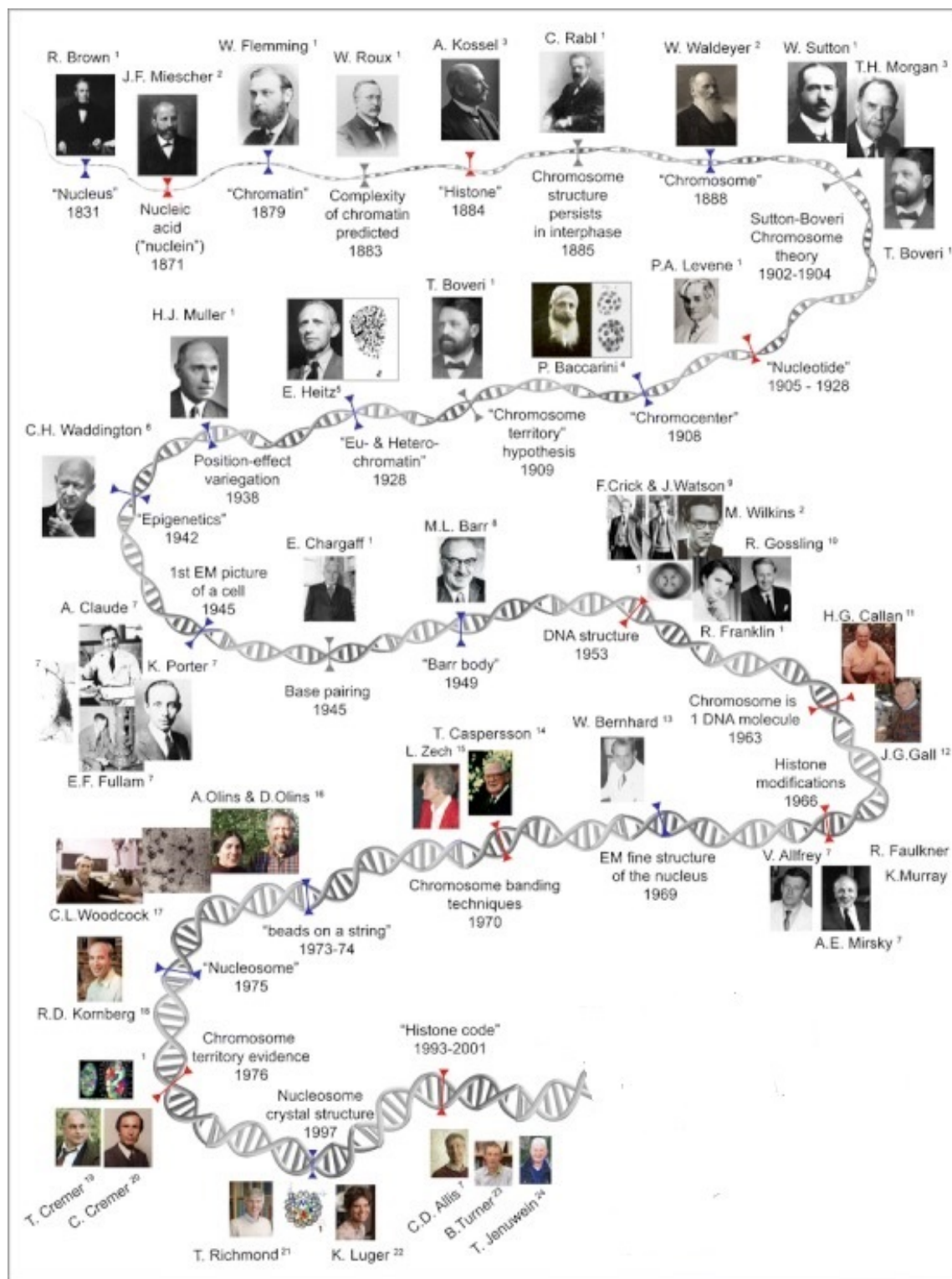


Figure 1: Timetable collection of landmark discoveries and concepts (gray) on molecular (red) and cellular (blue) aspects of chromatin. Adapted from (Jost et al., 2012).

1-2- Chromatin components

1-2-1- DNA

Deoxyribonucleic acid (DNA) was discovered by Friedrich Miescher in 1869 and its double helix structure was described by Watson and Crick in 1953. The basic building block of DNA is called nucleotide. The four different nucleotides present in the structure of DNA are called adenine (A), Thymine (T), Guanine (G) and Cytosine (C). Each nucleotide is composed of a nitrogenous base, a phosphate group and a sugar (deoxyribose). These bases can make hydrogen bonds from one strand of DNA to the other one, for instance, A and G respectively match with T and C. This base pairing forms the double helix of DNA which is coiled around a common axis (Figure 2A and (Alberts et al., 2002)).

1-2-2- Nucleosome

The nucleosome-based organization of the genome was first discovered by the observation of nuclease digested chromatin from rat liver cells. The resulting DNA fragmentation produces a 200 bp DNA fragment ladder, suggesting the possibility of DNA protection by a repetitive structure which we now know are formed of histones. The total length of the human genome is about 2 meters. This negatively charged DNA molecule repulses any interaction between DNA regions which inhibits its folding and hence could not reach the compact state required for fitting this long molecule inside the nucleus. This problem was solved when it was discovered that DNA is wrapped around a repetitive histone-based structure known as a nucleosome. There are four histones H2A, H2B, H3 and H4 which interact with each other to form an octameric complex where a tetramer of H3-H4 interacts with two dimers of H2A-H2B. Packaging of DNA into chromatin through nucleosome assembly decreases the 2 meters length of nude DNA to 0.09 mm in the chromatinized genome which can fit into a cell nucleus with a diameter of approximately 0.01 mm (Ooi and Henikoff, 2007). The structure of nucleosome, the building block of chromatin, is defined at the atomic resolution (1.9 Å) in which 147 base pairs (bp) of DNA are wrapped in 1.7 left-handed superhelical turns around an octamer of core histones (Figure 2B and (Davey et al., 2002)).

Histones have been discovered in 1889 by a German biochemist called Albrecht Kossel. Histones could be divided into two large categories: the core histones forming the

nucleosomes and the linker histones that cover DNA regions linking the nucleosomes together. The canonical core histones are small proteins (11-15kDa) that associate with DNA. They are highly basic and are conserved from yeast to human (Thatcher and Gorovsky, 1994). The four core histones including H2A, H2B, H3, and H4 share a similar structure in their globular domain which contains 3 α -helices connected by two loops known as “histone-fold” motif. It is believed that this region is involved in histone-histone and DNA-histone interactions. Within these α -helices the longest one (α 2) is located between two other shorter ones (α 1 and α 3) (Cutter and Hayes, 2015). In addition to the mentioned domain, the flexible and mobile N-terminal tails of the core histones that spread out of the nucleosome are accessible for chemical modifications and play important roles in chromatin functions (Figure 2C and (Gilbert et al., 2005). Moreover, these modifications can also occur in the central globular domain of the histones (Lawrence et al., 2015). Histone proteins interact with each other through their 3 α -helices to form H2A-H2B and H3-H4 heterodimers (Luger et al., 1997).

Linker histone family, also known as histone H1 family, is considered as the fifth histone. Linker histones (H1s: H1 and its variants) bind to DNA (close to entry/exit site of DNA) less strongly than core histones (Figure 3A), thus they can be dissociated from chromatin in the presence of (or by applying) low salt concentrations. H1s link adjacent nucleosomes to each other, stabilize wrapped DNA around core histones, induce higher-order chromatin structures and regulate gene expression (Cutter and Hayes, 2015).

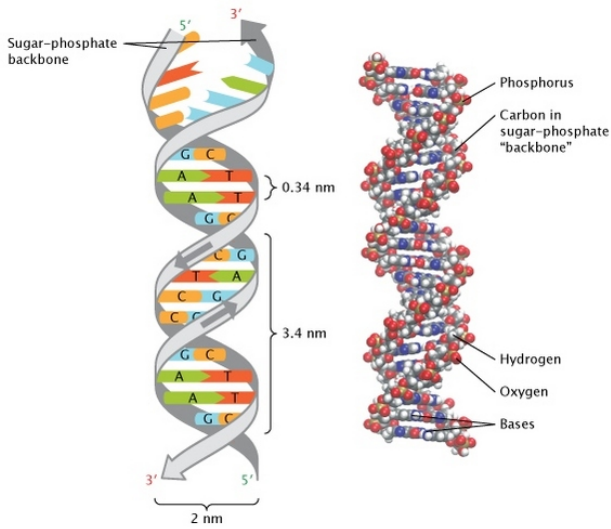
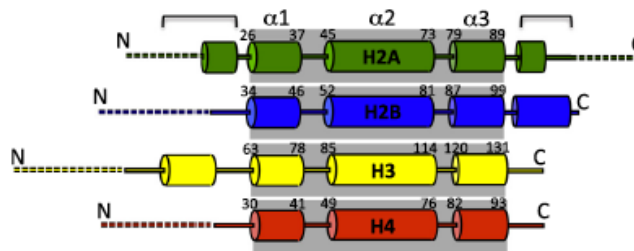
A**B****C**

Figure 2: Chromatin composition. **A)** DNA double helix structure, representing four nucleotides (A, T, C, G) and two anti-parallel strands of DNA which twist around each other. Taken from (© 2013 Nature Education). **B)** Crystal structure of nucleosome (on left, top view and on right, lateral view) containing four core histones in the center and 146 base pairs of DNA around them. Taken from (Luger et al., 1997). **C)** Scheme showing an alignment of the central histone fold domains of H2A, H2B, H3, H4. The grey box contains three α -helix of histone fold domains represented as cylinders and linked with lines. Tail domains and helices located outside the histone fold domain are indicated by dashed lines and brackets, respectively. Adapted from (Cutter and Hayes, 2015).

1-3- Chromatin structure and function

1-3-1- Hierarchical structure of chromatin

There are 400 negative charges in 200 bp of DNA wrapped around a histone octamer with $\cong 220$ residues positively charged, therefore the complete neutralization of all the DNA negative charges can not be achieved only through core histones and should involve other molecules such as linker histone H1, cations, etc. The first level of chromatin organization and DNA packaging is ensured by the nucleosomes forming a fiber of 11-nm diameter. This nucleosome fiber, which is also known as the 10-nm fiber, looks like “beads-on-a-string” under electron microscopy. In this structure, the length of Linker DNA present between the nucleosomes varies between 20 to 80 bp. Linker DNA has the propensity to associate with histone H1 (Figure 3A).

In 1976, Finch and Klug discovered the *in vitro* 30-nm fiber that was established by the folding of 10-nm fibers in a linker histone H1 or Mg^{+2} ion-dependent manner (Figure 3C). Therefore, the second level of DNA compaction and chromatin structure could be the 30-nm fiber which is considered to form a particular higher order structure following two models: “solenoids” and “zig-zag” according to the nucleosome arrangement in the fiber (Figure 3B and (Maeshima et al., 2014)). In the solenoids model, the linker DNA bends between adjacent nucleosomes that are positioned sequentially, around 6 to 8 nucleosomes per turn, to fold into “one-start helix”. In the zigzag model or “two-start helix”, nucleosomes are located in a zigzag manner and are connected by a straight linker DNA (Figure 3B and 3C and (Widom and Klug, 1985; Williams et al., 1986; Woodcock and Ghosh, 2010)).

Beyond the 30-nm fiber, the next levels of DNA compaction and higher-order structures of chromatin are found in interphase and metaphase chromosomes that can form structures of 700 nm and 1400 nm thickness, respectively (Figure 3C).

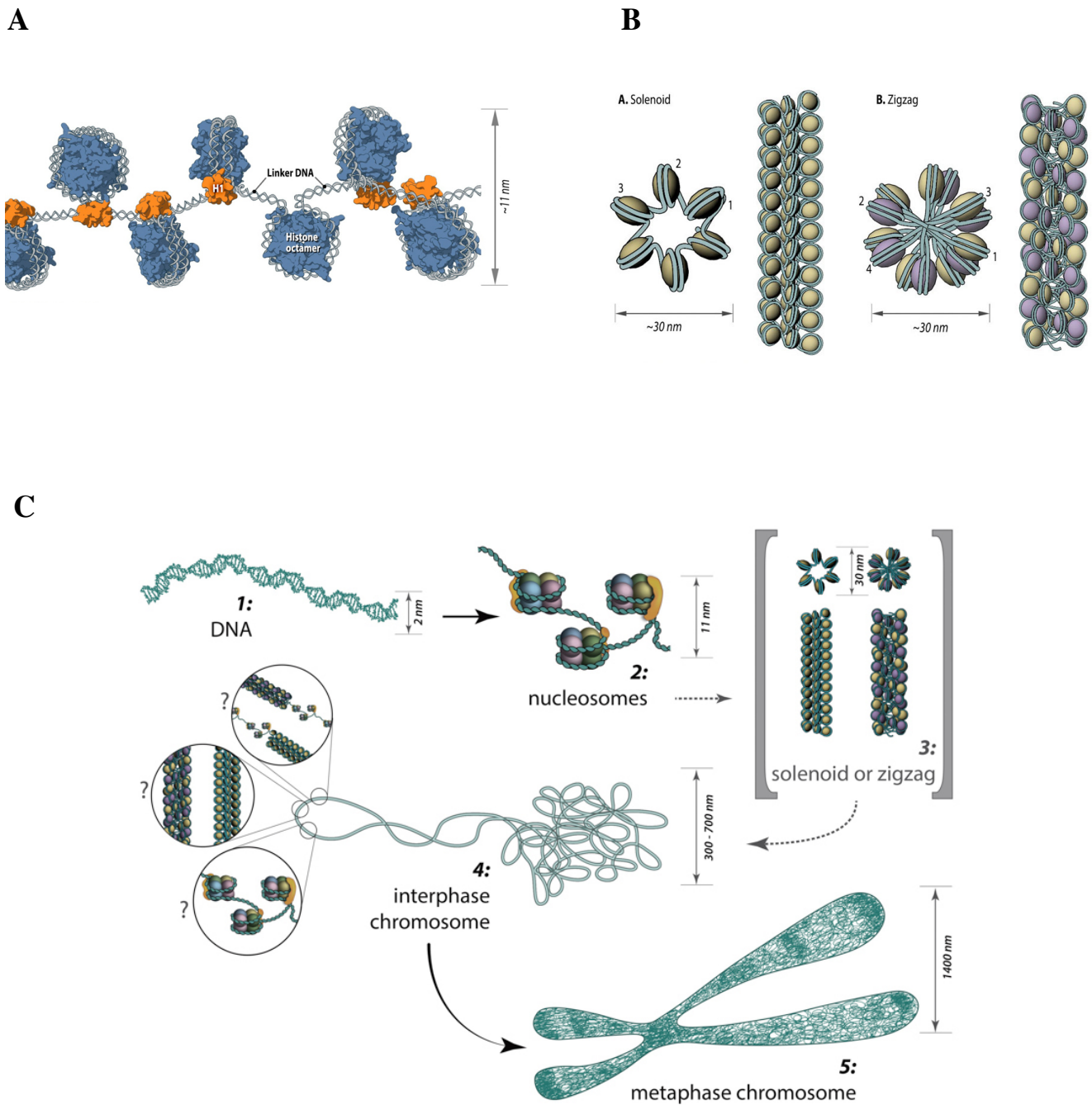


Figure 3: Chromatin organization levels and higher-order structure. A,B,C) Schematic representation of chromatin compaction from nucleosome to metaphasic chromosome. Taken from (MB:info site <https://www.mechanobio.info/topics/genome-regulation/dna-packaging/>).

1-4- Chromatin domains: Euchromatin and Heterochromatin

Two types of chromatin have been identified by staining DNA with a specific dye. The intensely stained DNA is called heterochromatin and, in contrast, the lighter stained portions are called euchromatin. Heterochromatin is a tightly condensed structure and was considered for long time as an inactive region of the genome. Euchromatin in contrast is less packed and is considered active since it contains dynamic regions and transcriptionally active genes (Figure 4 and (Lewin and Krebs, 2015)). Heterochromatin and euchromatin can now also be distinguished by histone modifications. Euchromatin regions are characterized by acetylation of histone H3 and H4, methylation of H3K4, whereas heterochromatin lacks active histone marks, such as acetylation, and contains Heterochromatin protein-1 (HP1) and methylation on H3K9. Repetitive DNA sequences such as pericentric and telomeric regions constitute portions of the genome that are covered by heterochromatin (Tamaru 2010). It is however important to note that heterochromatin itself was further divided into facultative and constitutive heterochromatin. The term of facultative heterochromatin refers to regions that are enriched in genes and could be activated upon specific signals and conditions. The best example of this category is the X-chromosome inactivation in female. In contrast, constitutive heterochromatin corresponds to gene-poor regions that are normally transcriptionally silent during the cell cycle (Saksouk et al., 2015).

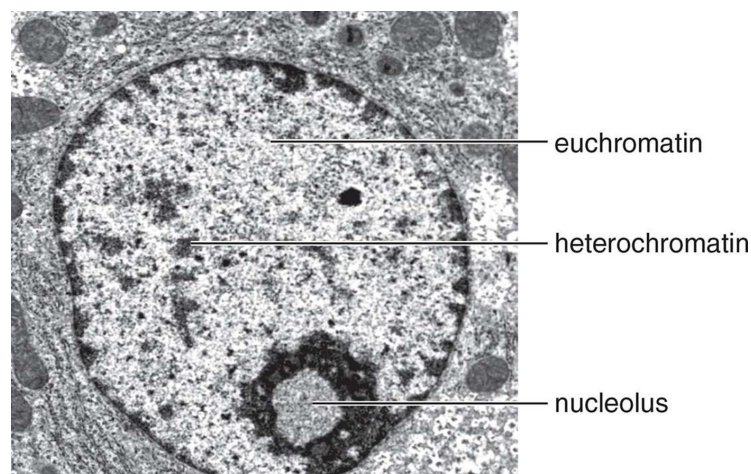


Figure 4: Structure of heterochromatin versus euchromatin. Dense and dark regions represent the heterochromatin, while the light and less compacted domains indicate euchromatin in the nucleus. Taken from (site http://schoolbag.info/chemistry/mcat_biochemistry/34.html).

1-5- The dynamic nature of chromatin

Although alterations of the genetic information occurs rarely in normal cells, the genome-bound histones and non-histone proteins and RNAs are subject to a wide variety of alterations including chemical modifications, reorganizations and active exchanges. These dynamic events are regulated and they control the function of the genome (Soshnev et al., 2016). Indeed the dynamic structure of chromatin regulates the accessibility of DNA to ensure transcription, recombination, DNA repair and replication (Venkatesh and Workman, 2015). Histone exchange is a mechanism that greatly contributes to chromatin fluidity. Histone replacement in a nucleosome could be total or partial. Several factors including histone post-translational modifications (PTMs), the action of energy-dependent chromatin remodelers and histone chaperones are involved in regulating chromatin dynamics (Venkatesh and Workman, 2015), which will be developed in more details below.

1-5-1- Histone post-translational modifications (PTMs)

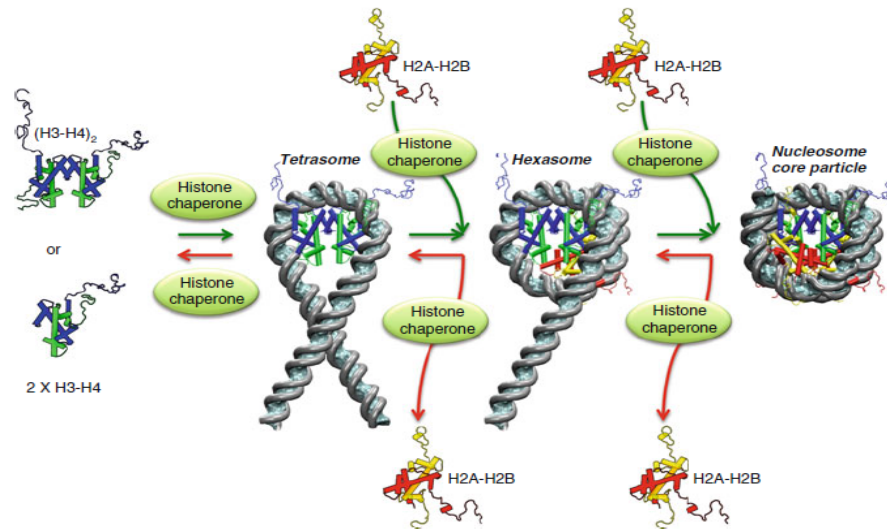
Depending on the amino-acids involved and their specific positions in histones, the occurrence of post-translational modifications on histones could weaken the histone-DNA interactions and alter the charge of histones with important regulatory and functional impacts (Venkatesh and Workman, 2015). Histone modifications will be further discussed in following sections.

1-5-2- Histone chaperones

The term 'histone chaperone' was first coined by R. Laskey to describe the role of a nuclear protein called nucleoplasmin in inhibiting the aggregation of histone-DNA during nucleosome assembly. Histone chaperones can function dependent and independent of DNA synthesis, and therefore are classified as replication-dependent and replication-independent. Additionally, histone chaperones can be also grouped into two classes, H3-H4 or H2A-H2B chaperones, based on the histone proteins that they bind to (Burgess and Zhang, 2013; Venkatesh and Workman, 2015). CAF1 (Chromatin assembly factor-1) and HIRA (histone regulator A) are two examples of histone chaperones which are respectively replication-

dependent and replication-independent (Quivy et al., 2001; Ray-Gallet et al., 2002). In general, in addition to their function in chromatin assembly (Figure 5A), these proteins participate in histone storage and transport. The essential contribution of histone chaperone proteins in the process of nucleosome assembly is during DNA replication. Chromatin assembly on the newly synthesized genome during the S phase occurs through several steps. During the process of DNA replication, in order to facilitate the function of the DNA replication machinery, nucleosomes should be disassembled. Once the newly synthesized DNA daughter strands appear, nucleosomes need to be reassembled into these nascent DNA strands. The histones required for the assembly of nucleosomes into replicating DNA are supplied either from newly-synthesized histones, for instance in somatic cells soluble histone proteins are synthesized during S phase, or from the old parental nucleosomes. The deposition of newly synthesized histones involves the function of histone chaperones. For instance, CAF-1 encompasses three subunits including p50, p60 and p160. CAF-1 is in charge of incorporating histones H3 and H4, while transient post-translational modification of H3 and H4 by cytosolic histone acetyltransferase occurs before their transportation into the nucleus. It is believed that CAF-1 is released from DNA after H3/H4 deposition and is not required for the H2A/H2B assembly (Figure 5B and (Krude, 1995)). NASP (Nuclear Autoantigenic Sperm Protein) appears to have a different role and controls the supply of histones and regulates their amounts. The histone chaperone Nap1 (Nucleosome Assembly Protein-1) is responsible for the shuttling of newly synthesized histones from the cytoplasm to the nucleus. Interestingly, the modification in the globular domain of histones that are not accessible to histone modifying enzymes is assisted by histone chaperones. H3K36 trimethylation depends on the function of both the histone methyltransferase Set2 and the histone chaperone Spt6 (Yoh et al., 2007; Youdell et al., 2008). FACT (Facilitate Chromatin Transcription) is one of the reported chaperones capable of binding to both H3-H4 and H2A-H2B (Belotserkovskaya et al., 2003).

A



B

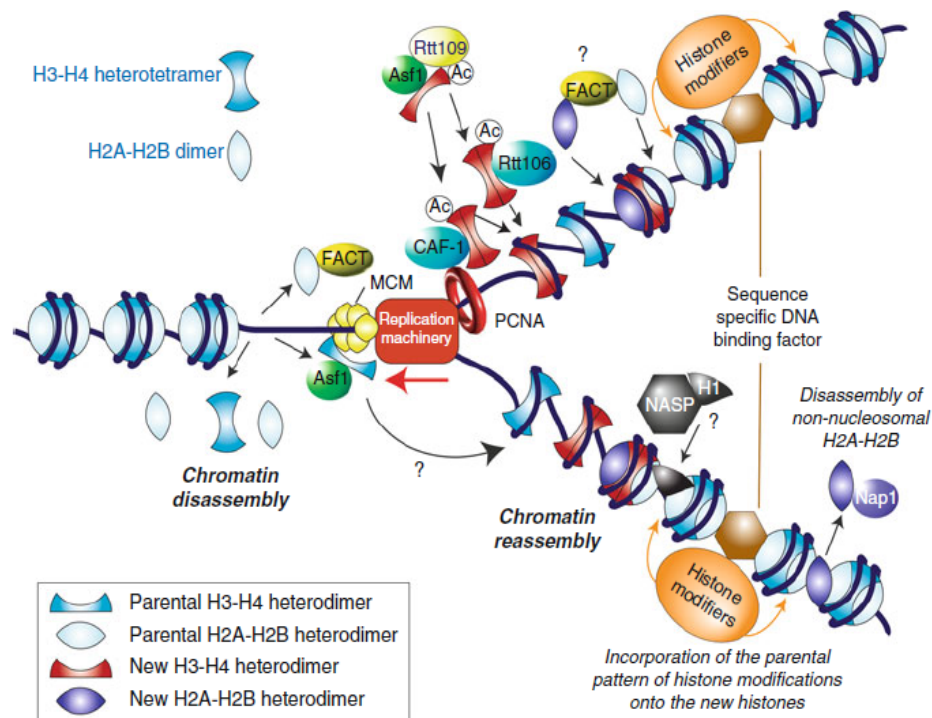


Figure 5: Histone chaperones in nucleosome assembly and disassembly. **A)** Schematic representation of the nucleosome core particle formation by the function of histone chaperones. Steps involved in chromatin assembly and disassembly are indicated by green and red arrows, respectively. Taken from (Workman and Abmayr, 2013). **B).** Model representing chromatin disassembly and assembly during DNA replication. Histones chaperones fulfill their task by removing histones in front of the replication fork and depositing histones behind it. Acetylation of H3K56 by the HAT activity of Rtt 109 facilitates interactions of histones with Rtt 106 and CAF-1 in yeast. Taken from (Workman and Abmayr, 2013).

1-5-3- ATP-dependent chromatin remodelers

All of the ATP-dependent chromatin remodelers contain an ATPase domain, therefore their function relies on the energy produced from ATP hydrolysis. Chromatin remodelers change the positions of nucleosomes along DNA by either sliding or evicting them. This action results in the generation of open DNA nucleosome free regions that can be targeted by the transcription machinery or by histone chaperones for nucleosome re-assembly. Briefly, these remodelers belong to four families: SWI/SNF (switching defective/sucrose non fermenting), CHD (chromodomain helicase DNA-binding), ISWI (Imitation switch), and INO80 (Inositol requiring 80).

This section will not be developed further here and the interested readers are invited to refer to the following recent reviews: ([Vignali et al., 2000](#); [Varga-Weisz, 2001](#); [Piatti et al., 2011](#); [Narlikar et al., 2013](#)).

1-5-4- Histone variants

One of the mechanisms that regulate DNA accessibility and result in chromatin remodeling is the substitution of core histones by histone variants in the nucleosome ([Bonisch and Hake, 2012](#)). In fact, histone variants incorporation within a nucleosome could affect many DNA-template dependent events such as transcription, recombination, repair, etc.

The sequences of histone variants are different from that of canonical histones. These differences vary from only a few amino acids to large sequence variations and additions of non-histone parts. Moreover, they also have distinct expression timing and deposition. Canonical and histone variants are expressed and incorporated in chromatin in a replication-dependent or replication-independent manner. Histone variants have been reported for all the core histone except histone H4. One of the important features of some of these variants is that they are located on specific genomic regions. Among the core histones, the majority of the identified histone variants belong to histone H2A, the main variants of this protein in the human are: H2A.Z, H2A.X, H2A.Bbd, and MacroH2A.

Additionally, most of the eukaryotes have H2A.Z and H2A.X as histone variants of H2A, whereas H2.Bbd and MacroH2A have been found only in vertebrates (Marino-Ramirez et al., 2005). The highest degree of variations between histone H2A variants usually originate from differences in their c-terminus (Ausio and Abbott, 2002).

1-5-4-1- H2A variants:

1-5-4-1-1- H2A.Z:

The addition of two amino acid residues in the C-terminal part of histone H2A.Z increases its acidic patch compared to canonical H2A. The acidic patch is thought to facilitate chromatin fiber folding and compaction (Suto et al., 2000). H2A.Z therefore is observed associated with heterochromatin protein HP1 α , that eventually leads to the establishment or maintenance of a specific conformation at constitutive heterochromatin domains. In addition, thanks to its extended acidic patch, H2A.Z also contributes to the formation of the higher-order chromatin fiber (Fan et al., 2002). Interestingly, despite sequence differences between H2A.Z and H2A, H2A.Z-containing nucleosomes show a globally similar structure compared to their canonical counterpart. However, several differences have been reported. For instance, the L1 loop which is an essential element for the interaction of the two H2A-H2B dimers within the nucleosome shows differences in the H2A.Z-containing nucleosomes versus H2A-containing nucleosomes, including a decreased stability of heterotypic nucleosome containing H2A.Z-H2B and H2A-H2B dimers due to a steric clash between their L1 loops at the opposite ends of the dyad axis (Suto et al., 2000; Bonisch and Hake, 2012). The surface of H2A.Z containing nucleosomes is also different from the canonical nucleosomes because of the altered amino acids in the docking domain, such as histidine, which has the ability to bind to metal ion. It also extends the acidic patch. As a consequence H2A.Z containing nucleosomes may interact with linker histones or other factors differently. H2A.Z functions in many biological processes: transcription regulation, DNA repair, heterochromatin formation, chromosome segregation and mitosis (Bonisch and Hake, 2012). Concerning transcription, H2A.Z presents a dual function in activating and repressing gene expression (Marques et al., 2010). Genome-wide mapping studies showed that H2A.Z is associated with gene promoters in *S. cerevisiae*, mammals and plants (Bonisch and Hake, 2012). Post-translational modifications (PTMs) such as acetylation, sumoylation and ubiquitination also occur on

H2A.Z, giving rise to different functions (Thambirajah et al., 2006). Ubiquitinated H2A.Z is localized on the inactive X chromosome (Xi) in mammals (Kalocsay et al., 2009), sumoylated H2A.Z associates with DNA repair in *S. cerevisiae* (Sarcinella et al., 2007) and the N-terminal acetylated form of H2A.Z destabilizes nucleosomes (Thambirajah et al., 2006). H2A.Z deposition into the chromatin is facilitated by the function of chromatin remodeling complexes, SWR-1 and SRCAP/TIP 60, respectively (Bonisch and Hake, 2012; Latrick et al., 2016).

1-5-4-1-2- H2A.X:

H2A.X is characterized by the presence of SQ[E/D] Φ motif in its C-terminal region, where Φ indicates a hydrophobic amino acid. As a response to DNA double-strand breaks (DSB), Serine 139 becomes phosphorylated and this modified form of H2A.X is known as gamma H2A.X (γ -H2A.X). To date, there is no evidence confirming a direct chromatin destabilization by the phosphorylation of H2A.X, but it has been shown that γ -H2A.X helps recruiting DNA repair proteins (Downs et al., 2000; Celeste et al., 2003; Thiriet and Hayes, 2005; Fink et al., 2007).

1-5-4-1-3- Macro.H2A:

This is a unique histone variant which has a very extended C-terminal part containing non-histone sequences. The fascinating structure of this histone variant is that an H1-like linker region connects the histone domain of this protein located at N-terminal to the non-histone macro domain. This variant is globally involved in transcriptional repression (Angelov et al., 2003). Accordingly, it was found associated with the inactive X chromosome (Perche et al., 2000; Bonisch and Hake, 2012).

1-5-4-1-4- H2A.Bbd:

The newly identified variant of histone H2A is H2A.Bbd, showing a highly different structure compared to H2A. This protein is lacking a part of the docking domain which makes

it smaller compared to canonical H2A. H2A.Bbd is known as H2A.Lap1 in the mouse, because it lacks the acidic patch. H2A.Bbd's expression is almost restricted to testis and brain (Soboleva et al., 2011). Unlike canonical H2A, H2A.Bbd is excluded from the inactive X chromosome but enriched in actively transcribed regions (Chadwick and Willard, 2001).

1-5-4-2- H2B variants:

Most of H2B variants are preferentially or exclusively expressed in testis. Among these, TH2B and SubH2Bv are two examples. TH2B differs from H2B mainly in the N-terminal part (Govin et al., 2004). The TH2B functional importance has been recently shown to be involved in histone-to-protamine exchange occurring during late spermiogenesis (Montellier et al., 2013) as well as in facilitating genome reprogramming (Shinagawa et al., 2014; Yang et al., 2014). The functions of SubH2Bv and most of the other identified H2B variants are still obscure.

1-5-4-3- H3 variants:

In mammals, Histone H3 includes five variants. H3.1 and H3.2 that are considered canonical variants. The variant H3.3, the centromere-specific variant CenH3 (CENP-A in mammals) (Allshire and Karpen, 2008), and the testis-specific histone H3t are replacement histone variants (Witt et al., 1996). There is a very rich literature on the role of histone H3 variants mainly H3.3 and CENP-A. Here we will not discuss further these functional aspects but invite the readers to refer to the following references (Szenker et al., 2011; Filipescu et al., 2014; Jacob et al., 2014).

2- Overview of epigenetics

2-1- A brief history of evolving landscape of epigenetics

Several decades of research in molecular biology have shed light on the fact that cells' functions do not rely only on DNA sequence, and that other levels of information exist that are not directly encoded by DNA but still could be inherited/transmitted. There is evidence indicating that all the information is not contained in DNA. For instance, although most of our cells have the same DNA sequence, they express different genes and are able to transmit their specific functional state to their daughter cells. At the scale of an organism, we now know that, in the case of many identical twins, each individual could express different specific phenotypes. Additionally, phenotypes such as coat colors and body weight have been observed in mice having the same DNA sequence. In insects, ants and bees express very different phenotypes at the individual levels, for instance, between the queen and the workers and soldiers, while they all share the same genome. All these above-mentioned examples indicate the presence of mechanisms over the genome that are now designed under the term of “epigenetics”. Increasing knowledge and interest in epigenetics has revealed that gene functional states could be memorized and inherited, suggesting that this information, which is not directly encoded by DNA sequence, should rely on a code beyond the genetic code (Weinhold, 2006). Today the term epigenetics refers to heritable changes that regulate gene expression which are not directly dependent on DNA sequence (Trasler, 2009). The epigenetic information relies on chromatin structure and can switch gene expression on or off.

The history of epigenetics cannot be separated from the study of evolution and development. In the early twentieth century, developmental biology and genetics were considered as two separate domains. Two famous developmental biologists named Conrad H. Waddington and Ernst Hadorn, during mid-twentieth century, declared the commonality of developmental biology and genetics and proposed to combine them. This idea signed the birth of what is called today “epigenetics”. The word epigenetics that was coined by Waddington in 1942, came from a Greek word “epigenesis” which was used to describe the transmission of acquired characteristics in response to environmental changes across generations. Waddington's approach to demonstrate the inheritance of acquired characteristic was based on his knowledge of fruit fly embryo development. It is worth mentioning that Waddington speculated that the process of development could be a landscape in which cells can take

different pathways during their differentiation. This developmental landscape comprises a series of “valleys” and “forks”, each valley representing a development pathway, which would be separated from the others by hills and mountains. These valleys and hills represent “barriers” preventing committed cells from changing path. For instance, a ball at the top can follow different pathways and, depending on which valley it passes, it can differentiate into a specific cell or tissue (Figure 6A and 6B). He also proposed another scheme illustrating the idea that any change in environment could lead to a new characteristic that could be inherited and incorporated into the genome (Figure 6C and (Noble, 2015).

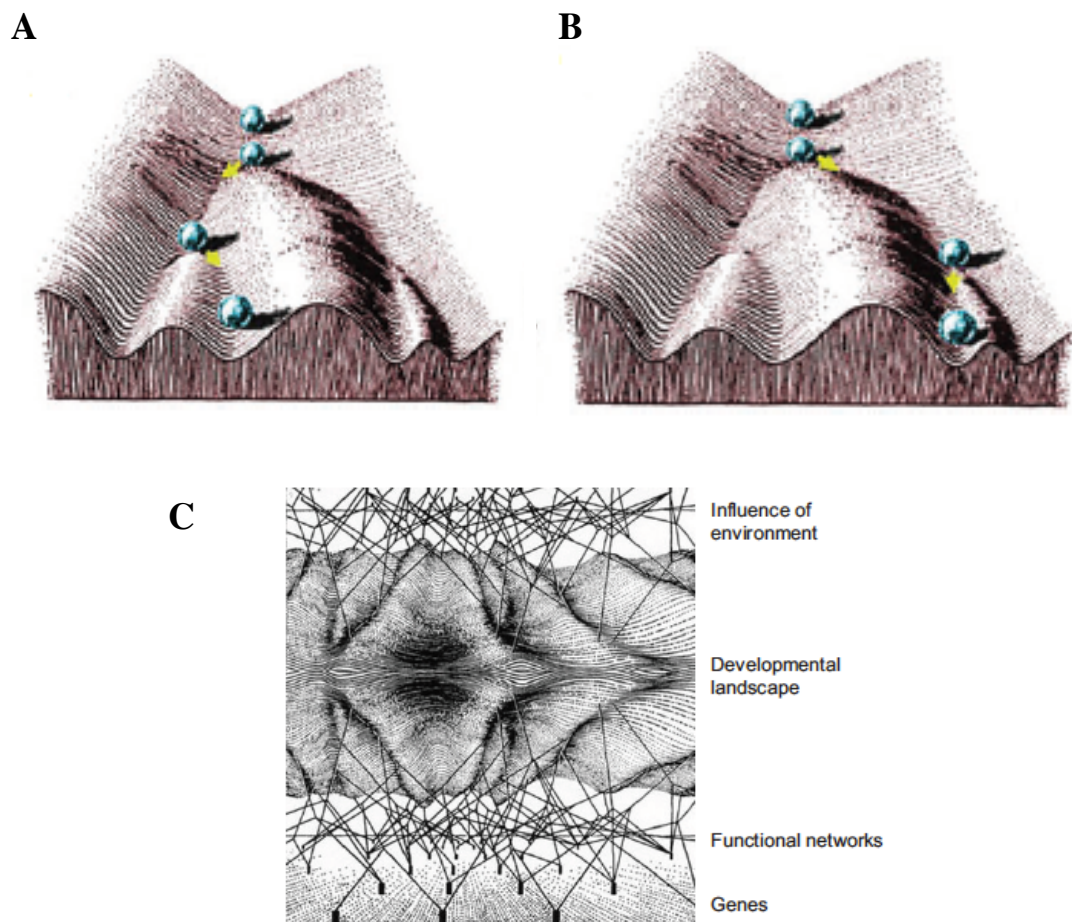


Figure 6: Epigenetic landscape. A,B) Schematic representation of a modified form of Waddington’s speculation on epigenetics. The ball at the top is the totipotent cell with the potential to differentiate to different cell types (balls at the bottom) depending on the developmental route (A or B) taken during the developmental process. Adapted from (Noble, 2015), modified diagram by K. Mitchell. **C)** Diagram illustrating how the environment could affect epigenetics and also the interactions underlying the epigenetic landscape. The slopes of valleys in the epigenetic landscape are controlled via genes (represented by pegs) and their interactions (represented by ropes), and any change to a gene could disrupt its interaction with other genes resulting in an overall effect on epigenetics. Taken from (Noble, 2015), modified from Waddington, 1957).

2-2- Epigenetics landscape

2-2-1- Chromatin modifications

The packing of DNA into chromatin prevents the binding of essential proteins and factors that are involved in driving different biological regulatory pathways, e.g., proteins such as topoisomerases and DNA polymerases during replication and transcription factors and RNA polymerases during transcription. Thus, DNA should be exposed and be accessible at anytime needed to promote binding by specific factors. DNA accessibility is facilitated by changes in the chromatin structure mediated by nucleosome disassembly factors, histone variants and/or histone post translational modifications (PTMs). Both DNA and histones are the substrates for enzymatically catalyzed chemical modifications, which are determinant elements in epigenetics and are sometimes called "epigenetic marks". Although DNA can be mainly methylated, histones are subject to numerous post-translational modifications ([Musselman et al., 2012](#)). Here below we briefly survey our knowledge of these so called epigenetic marks.

2-2-1-1- Histone modifications

Post-translational modifications of histones mostly concern residues that are located at the histone tails, since these tails extend out of the nucleosome and are more accessible to enzymes that deposit or erase these PTMs ([Figure 7A](#) and ([Musselman et al., 2012](#)). In addition to acetylation, a well-known histone modification, histones can undergo other modifications including methylation (me), phosphorylation (ph), ubiquitinylation (ub), sumoylation (su), ADP ribosylation, and deamination. In 2000, Brian D. Strahl & C. David Allis proposed the “histone code” hypothesis stating that the selective labeling of histone residues by post-translational modifications would decorate histone tails in specific patterns that could be read by “reader” proteins to regulate specific functions ([Figure 7C](#)). This hypothesis is in fact an extension of the “histone acetylation code”, hypothesis that was formulated earlier by Bryan Turner ([Turner, 2000](#)). It is of note that histone modifications are not restricted to histone tails, since they have also been found within globular domains with important functions ([Figure 7B](#) and ([Cosgrove et al., 2004](#); [Tropberger and Schneider, 2013](#)). In fact, histone modifications can crosstalk with one another. In other words, the ability of a residue to be modified by its modifying enzyme depends on the modification of another

residue. For instance, the phosphorylation of serine 10 on histone H3 stimulates the ability of Gcn5 to acetylate histone H3 at lysine 14 (H3K14) (Figure 7D and (Cheung et al., 2000; Lo et al., 2000)).

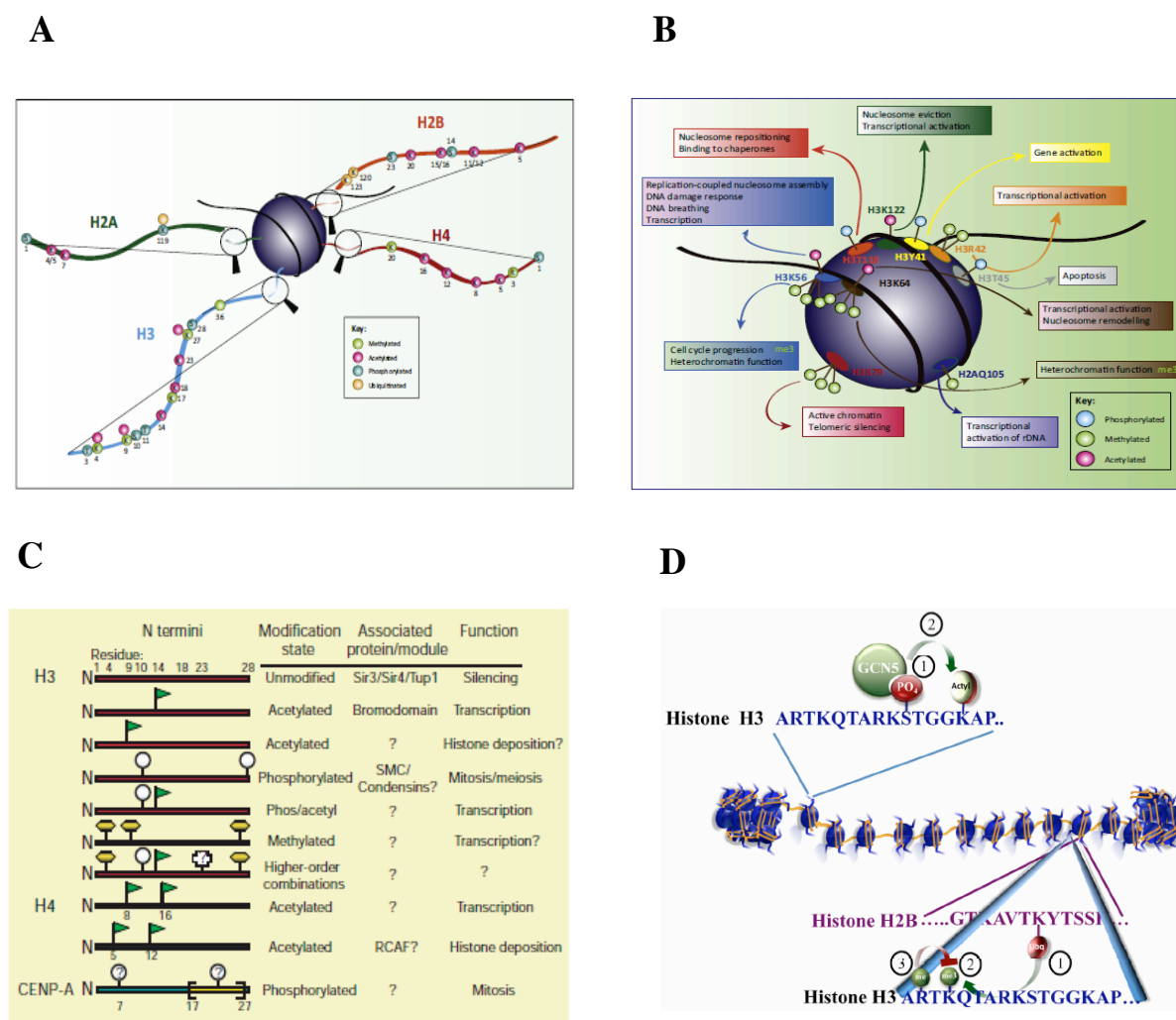


Figure 7: Histone post-translational modifications (PTMs). **A)** Histone tail modifications H2A, H2B, H3, and H4. Different modifications are shown with different colors. Taken from (Lawrence et al., 2016). **B)** The occurrence and function of PTMs within globular domains of histones are indicated. Taken from (Lawrence et al., 2016). **C)** The concept of ‘histone code’ is that modified histones can act individually or create a pattern and function in combinations, for instance H3 phosphorylation at serine 10 and 28 leading to help chromosome segregation and condensation during mitosis and meiosis. Taken from (Strahl and Allis, 2000). **D)** Schematic representation of histone cross-talk. Taken from (Lee et al., 2010).

2-2-1-2- DNA modifications

Chemical modifications of DNA, methylation of cytosine residues, is also essential for the development. Methylation of the cytosine residues at their 5 position (5mC) is an abundant DNA modification and 5mC is considered as the ‘fifth base’. DNA methylation is dynamic: it is enzymatically established and removed. The enzymes establishing DNA methylation are of two categories respectively responsible for either the maintenance of DNA methylation during DNA replication on the newly synthesized strand (DNMT1), or the addition of new methyl groups independently of DNA replication. The latter enzymes are known as *de novo* methyltransferases (DNMT3a and DNMTb).

Cytosine methylation mainly occurs at CpG dinucleotides in which cytosine is followed by guanine. There are also exceptions because non-CpG cytosines methylation has been reported to occur in neuron and embryonic stem cells (ESCs), which respectively represses transcriptional activity in mouse neurons via the recruitment of methyl-CpG binding protein 2 (Mecp2) and correlates with gene expression in ESCs (Guo et al., 2013).

In the genome there are regions that are highly enriched in CpG nucleotides. These are mostly in genes promoters and are known as ‘CpG islands’. Remarkably, these CpGs are generally protected against methylation and physiological or pathological methylation of these CpG usually suppresses gene expression. This phenomenon is important in the establishment of tissue-specific gene expression and silencing. It also plays an important role in cancer since aberrant methylation of CpG islands in regions upstream particular genes, especially tumor suppressor genes, can lead to their unprogrammed transcriptional silencing and contribute to malignant cell transformation (Feltus et al., 2006).

5mC oxidation through the activity of Ten-eleven translocation (Tet) proteins generates 5-Hydroxymethyl cytosine (5hmC) that can form 5-formyl cytosine (5fC) and 5-carboxyl cytosine (5caC) through further activity of Tet proteins (Figure 8A and (Breiling and Lyko, 2015)). The functions of 5mC and its derivatives are summarized in (Figure 8B). The abundance of 5hmC is tissue-dependent with the highest levels reported in the brain (Kriaucionis and Heintz, 2009). Methylation in repetitive DNA sequences prevents DNA transpositions and is associated with chromosome stability. Any increase and decrease in the

level of DNA methylation, hyper- and hypo- methylation, compared to normal tissue could cause aberrant gene expression and, as mentioned above, lead to malignancies.

DNA methylation and histone modifications can cross-talk and cooperate in the regulation of gene expression, even though they are regulated by different enzymes. For instance, H3K4me may block DNA methylation whereas H3K9me and H3K27me can promote it. For instance, the methyl group on lysine four of histone H3 inhibits the interaction of Dnmt3L which has a role in driving DNA methylation via the targeting non-methylated H3K4 (Ooi and Henikoff, 2007; Otani et al., 2009).

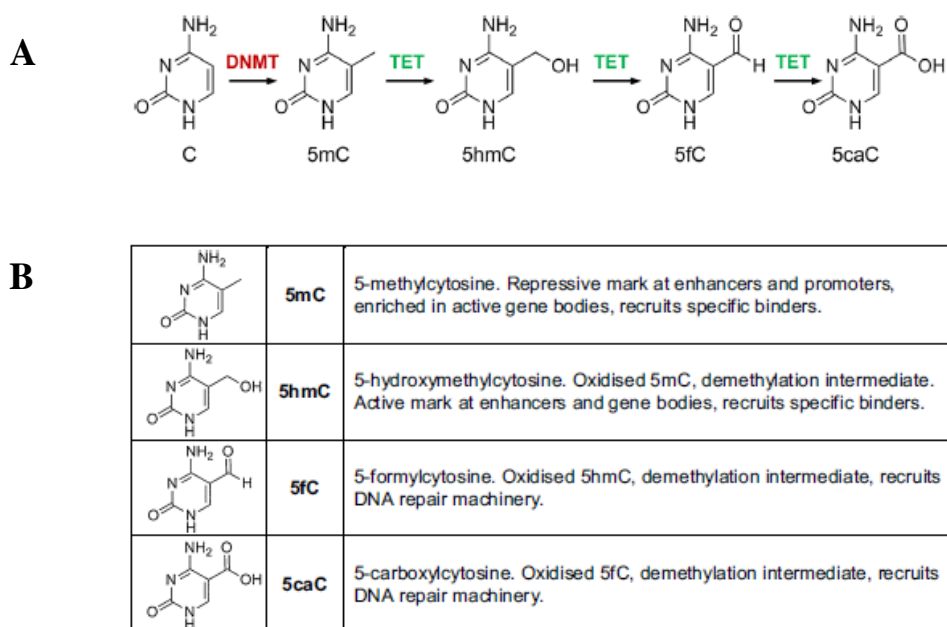


Figure 8: DNA covalent modifications. A) Cytosine methylation and its derivatives generated by the activity of TET proteins. 5mC : 5-methyl cytosine, 5hmC: 5-hydroxymethyl cytosine, 5fC: 5-formyl cytosine, 5caC: 5-carboxyl cytosine. Taken from (Breiling and Lyko, 2015). B) The functions of DNA modifications are listed. Taken from (Breiling and Lyko, 2015).

2-2-2- Epigenetic regulators

Epigenetic regulators are proteins involved in the establishment of a given epigenetic state or landscape and can be divided into three classes based on their functions:

A) Writers: enzymes that add modifications to histone residues or DNA such as histone acetyltransferases (HATs), histone methyltransferases (HMTs), protein arginine methyltransferase (PRMTs) and histone kinases.

B) Erasers: enzymes that catalyze removal of modifications such as histone deacetylases (HDACs), lysine demethylases (KDMs) and phosphatases.

C) Readers: specific protein modules that bind to epigenetically modified histones in a modification-dependent manner such as bromodomains, chromodomains, YEATs domains and PHD domains containing factors or proteins that can bind methylated DNA such as MeCP2 (Figure 9 and (Falkenberg and Johnstone, 2014; Andrews et al., 2016).

Histone PTMs affect chromatin structure and dynamics via two main mechanisms. They can either influence Histone-DNA and histone-histone contacts directly, or regulate the binding of PTM-specific binding proteins, that interact with modified histones through their specific domains such as bromo-, chromo-, and PHD domains.

This recognition of histone marks by readers, not only mediates transcription, but also is involved in DNA damage response, DNA replication and recombination, and chromatin remodeling. Indeed, PTM-binding protein can form a platform recruiting various regulatory proteins. Epigenetic “reader” misfunction leads to developmental abnormality and cancer (Bannister and Kouzarides, 2011; Tan et al., 2011; Musselman et al., 2012).

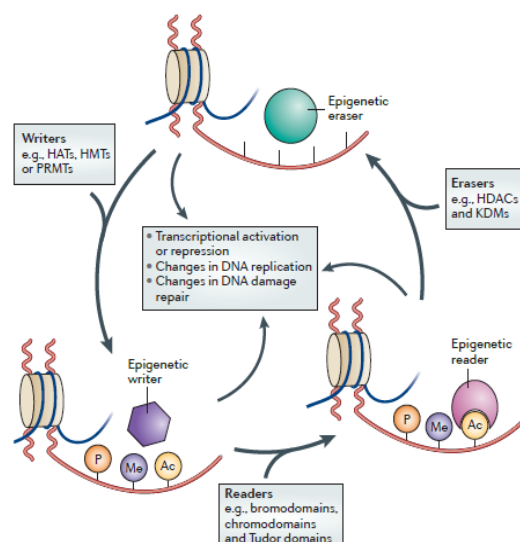


Figure 9: Writers, readers, and erasers of histone PTMs. Histone acetyltransferases (HATs), histone methyltransferase (HMTs), protein arginine methyltransferases (PRMs) and kinases add chemical groups to residues located in the tail of histones. Modified histones can be recognized by readers such as bromodomains, chromodomains, and Tudor domains. The erasers such as histone deacetylases (HDACs), lysine demethylases (KDMs), and phosphatases catalyze removal of added groups. Taken from (Falkenberg and Johnstone, 2014).

2-2-2-1- Histone acetyltransferase (HAT)

As presented above, HATs are responsible for transferring an acetyl group from acetyl-CoA to the ϵ -amino group of lysine side chains. HATs are very diverse and, depending on their catalytic domain, can be divided in three major groups: Gcn5 (general control non-derepressible 5)-related N-acetyltransferases (GNATs), E1A-associated protein of 300 kDa (p300)/CREB (cAMP-responsive element binding protein) binding protein (CBP), and MYST proteins. MYST HATs include Morf, Ybf2 (Sas3), Sas2 and Tip60 ([Sun et al., 2012](#)).

2-2-2-1-1- The GNAT family

GNATs family consists of Gcn5, PCAF, Elp3, Hat1, Hpa2 and Nut1, which are conserved from yeast to human. Human GCN5 and PCAF, two homologous proteins, through interacting to p300/CBP can control cell cycle and also function as transcriptional regulator ([Sun et al., 2012](#)).

2-2-2-1-2- The MYST family

MYST HATs include MOZ (monocytic leukemia zinc finger protein), Morf (MOZ-related factor), Tip60 (Tat-interacting protein of 60 kDa), HBO1 (HAT bound to Orc1), and MOF (males absent on the first) ([Timmermann et al., 2001](#)).

2-2-2-1-3- The p300/CBP family

These two enzymes are found in humans and higher eukaryotes. p300 with almost 300 kDa in size is also known as EP300 or KAT3B, related to its ability to bind to E1A which is an adenoviral oncogenic transcription factor. Some years after p300 discovery, in the search of a protein with the capacity of binding to CREB (cAMP response element-binding protein) the transcription factor, CBP was identified. p300 and CBP share similar sequence homology ([Figure 10 and \(Dancy and Cole, 2015\)](#)), and therefore are called as p300/CBP. p300/CBP are localized mainly in the nucleus. However, it has been shown that they can shuttle between the nucleus and the cytoplasm but the pathway has remained unknown. Histones and many other

proteins can be modified by p300/CBP. Indeed, p300/CBP regulate many cellular signaling pathways in different ways, not only due to their capacity to interact with different proteins through their various domains, but also because of their ability to modify themselves post-translationally. For instance, the subcellular localization of p300/CBP might be changed by sumoylation of three of its lysines, which are adjacent to the bromodomain and impose negative effect on its activity (Ryan et al., 2010). Autoacetylation of p300/CBP is another example, increasing p300/CBP acetyltransferase activity (Karukurichi et al., 2010).

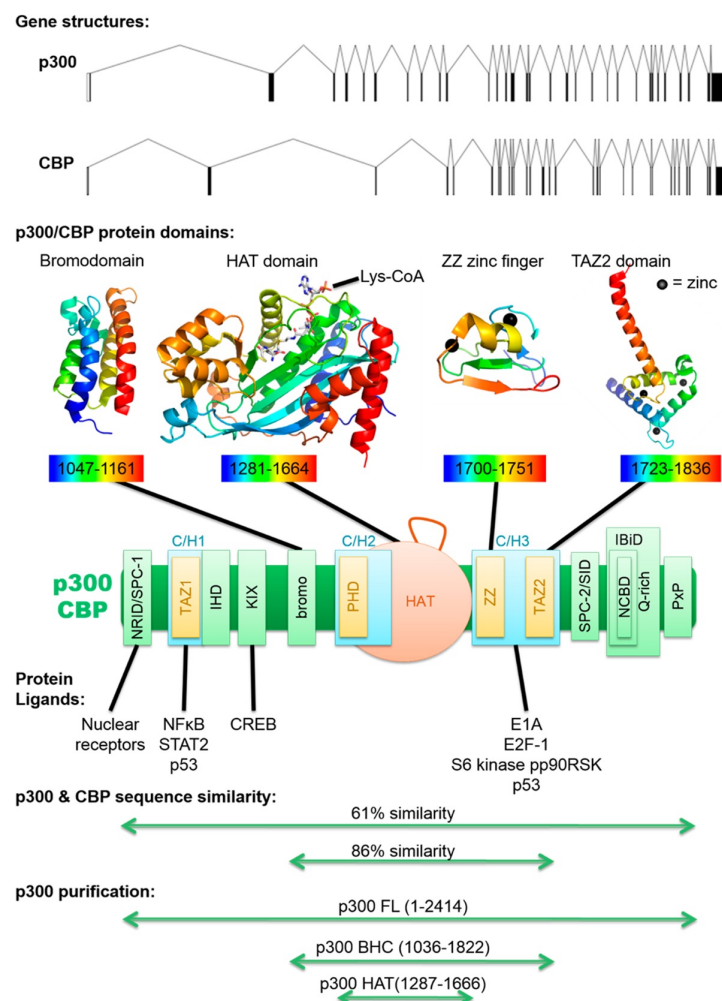


Figure 10: CBP/p300 gene and protein structures. Taken from (Dancy and Cole, 2015).

2-2-2-2- Histone deacetylase (HDAC)

The dynamic nature of histone acetylation depends on the opposite actions of HATs and HDACs, emphasizing the importance of these two enzymes in regulating mechanisms related to gene expression and many biological processes. In contrast to histone acetylation which opens chromatin structure and provides binding sites for chromatin binding proteins, HDACs activity represses gene expression.

HDAC proteins can be divided into two categories:

-“zinc-dependent” enzymes which are sensitive to trichostatin A (TSA), which inhibits their function.

-“nicotinamide-adenine-dinucleotide (NAD)–dependent” enzymes known as sirtuins which are resistant to TSA.

The Zinc-dependent group includes four families, whose structure, enzymatic functions, subcellular localization and gene expression patterns are different (Figure 11 and (Haberland et al., 2009)).

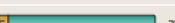
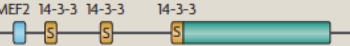
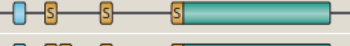
	Protein domains	Time of lethality	Phenotype	References
Class I	HDAC1  482	E10.5	Proliferation defects	12,53
	HDAC2  488	P1	Cardiac malformation	12,61
	HDAC3  428	E9.5	Gastrulation defects	69–71
	HDAC8  377	P1	Craniofacial defects	M.H. and E.O., unpublished observations
Class IIa	HDAC4  1,084	P7–P14	Chondrocyte differentiation defect in growth plate	33
	HDAC5  1,122	Viable	Exacerbated cardiac hypertrophy after stress	32
	HDAC7  912	E11	Endothelial dysfunction	34
	HDAC9  1,069	Viable	Exacerbated cardiac hypertrophy after stress	31
Class IIb	HDAC6  1,215	Viable	Increased tubulin acetylation	43
	HDAC10  669	ND	–	–
Class IV	HDAC11  347	ND	–	–

Figure 11: The histone deacetylase (HDAC) classification. HDAC conserved domain is shown in green rectangles, the blue square indicates myocyte enhancer factor 2(MEF)-binding sites. 14-3-3 is a phosphoserine binding protein. Taken from (Haberland et al., 2009).

3- Discovery of histone lysine acylation

New methodologies and high-sensitive mass spectrometric analyses of histones revealed that the number of histone modifications is much greater than previously anticipated. The explosion of newly discovered histone marks, including histone lysine crotonylation (Kcr), 2-hydroxyisobutyrylation (Khib), butyrylation (Kbu), malonylation (Kma), propionylation (Kpr), succinylation (Ksu), glutarylation (Kglu) and β -hydroxybutyrylation (Kbhb) (Chen et al., 2007; Tan et al., 2011; Xie et al., 2012; Dai et al., 2014; Tan et al., 2014; Xie et al., 2016) is now rising many unanticipated questions. Biologists should not only include these marks in their functional investigations but they also need to establish the functional interplay between these new PTMs and the already studied ones.

Some of the “hot” questions are as follows:

what is their functional significance? How do these modifications compete with abundant histone lysine acetylation? Which enzymes are involved in their deposition, removal? What are their readers? More importantly, what are their functional interplays with the other modifications? (Tan et al., 2011; Rousseaux and Khochbin, 2015).

As mentioned earlier, lysine acetylation changes lysine amino group charge from positive to neutral, increases the length of lysine side chain and its hydrophobicity. An acetylated lysine is capable of making hydrogen bond, electrostatic and van der Waals interactions. Newly identified short-chain lysine acylations are longer in their length compared to acetylation and can be divided into three groups with regard to differences in their chemical structure in comparison to lysine acetylation (Figure 12):

- 1) hydrophobic acyl groups: acylation of lysine residue with these acyl-groups (propionyl, crotonyl, butyryl) increases the hydrophobicity
- 2) Polar acyl groups: one of the features of these acyl groups is the presence of a hydroxyl group, which means that the addition of these acyl groups (β -hydroxybutyryl and 2-hydroxyisobutyryl) to a lysine residue gives the residue the potential to establish hydrogen bonds.
- 3) acidic acyl groups: like phosphorylation, these acylations (succinyl, malonyl, glutaryl) add negative charge to the modified lysine.

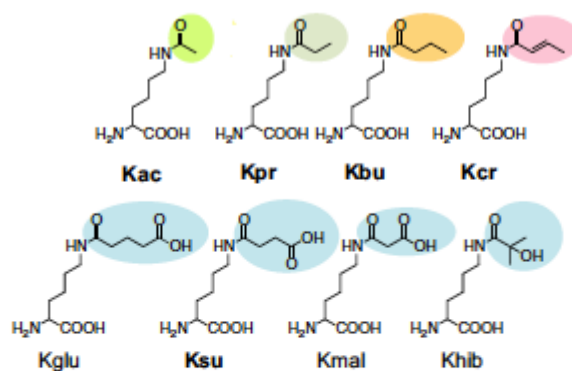


Figure 12: Chemical structure of lysine acylations. Adapted from (Flynn et al., 2015).

3-1- Well-known histone modifications

3-1-1- Acetylation

The first histone modification identified in the sixties is histone acetylation. This PTM is therefore one of the best-characterized in terms of mechanisms directing its establishment, removal and recognition (Shahbazian and Grunstein, 2007). Histone acetylation changes the charge of the modified histones by neutralizing the lysine's positive charge. When the modified lysine is present in specific positions in the histone fold domain, its acetylation could weaken its interactions with DNA or with the other histones. Signaling by histone acetylation is involved in a wide range of biological processes whose mis-regulation contributes to pathologies such as cancer, neurodegenerative disorders, etc.

Acetylation is a highly dynamic modification which is capable of directly changing the chromatin structure or of indirectly acting on chromatin by recruiting chromatin binding proteins.

3-1-2- Methylation

There are some differences between methylation and acetylation. Unlike acetylation, histone methylation does not change the charge of histones and the side chains of lysines and arginines are the main targets of this modification. Nevertheless, lysines can be modified in three manners, they can be mono-, di- or tri-methylated, but arginines can be mono- or di-methylated ([Bannister and Kouzarides, 2011](#)). Arginine methylation will not be discussed here. Most of the methylated lysine sites are reported at the N-terminal tails of histones, but H3K79me is an example of methylation in the globular domain of histone H3 which is catalyzed by the activity of a histone methyl transferase known as Dot1L ([Hublitz et al., 2009](#)).

Histone methylation can either increase or decrease gene expression depending on which residue accepts the methyl group. For instance, trimethylation of histone H3 at lysine 4 (H3K4me3), which is enriched in promoter and transcriptional start site regions (TSSs) enhances transcriptional activity but dimethylation and trimethylation of histone H3 at lysine 9, 27 (H3K9me2/3, H3K27me2/3) leads to silencing of gene expression ([Barski et al., 2007](#)) through the recruitment of HP1 (Heterochromatin Protein 1) and Polycomb repressive complex proteins. During many years it was believed that histone methylation was a stable mark until the discovery of first lysine demethylase known as lysine-specific demethylase 1 (LSD1) in 2004. Many other enzymes involved in Lysine methylation and demethylation have been discovered since then ([Nottke et al., 2009](#); [Bannister and Kouzarides, 2011](#)).

3-1-3- Phosphorylation

Like acetylation, histone phosphorylation alters the electrostatic properties of the histones. The hydroxyl group of some amino-acids such as serine, threonine and tyrosine are the main targets of kinases that add a phosphate group. This modification can be erased by the function of phosphatase enzymes. This PTM is involved in several biological processes. The best examples are the phosphorylation of H2A.X (Y-H2A.X) in DNA repair, and that of H3S10 during mitosis and transcriptional regulation ([Musselman et al., 2012](#)).

3-2- Recently identified histone modifications

3-2-1- Crotonylation

Crotonylation, a recently discovered histone PTM, is an evolutionarily conserved mark. It shows marked differences compared to acetylation. From a structural point of view, the crotonyl group is different from the acetyl group, because it has a flattened C-C π -bond, while the methyl group of an acetyl group is tetrahedral and rotatable (Figure 13A). This specific structural aspect allows to anticipate that Kcr could be regulated with different enzymes compared to Kac. The identification of Kcr sites in core histones (Figure 13B), showed that some residues can be modified with either ac group or cr group. This mark has been identified in five tested eukaryotic species which include yeast *S. cerevisiae*, *C. elegans*, *Drosophila* S2 cells, mouse embryonic fibroblast (MEF) cells, as well as human HeLa cells. It seems that HDACs 1,2,3 and 6 do not have considerable decrotonylase activities. However, Bao et al showed that three members of the sirtuins family (Sirt1, Sirt2, Sirt3) can catalyze removal of the crotonyl mark *in vitro* and that, among them, only Sirt3 showed a significant *in vivo* decrotonylase activity. Sirt1-Sirt3 target crotonylation on all core histones but the robust decrease of H3K4cr and H3K27cr revealed that these sites could be the preferred ones (Bao et al., 2014).

The genomic distribution of histone Kcr is very similar between male germ cells and human somatic cells. Histones labeled with crotonylation are enriched in the TSS of active genes. Interestingly, X-linked genes that escape repression also show an increase in histone crotonylation compared to their repressed state, although this aspect will be discussed and developed further in the discussion section (Tan et al., 2011). A recent detailed study of crotonylation revealed that histone crotonylation directly enhances transcription to a greater extent compared to acetylation. The concentration of cellular crotonyl-CoA which is the donor for crotonylation, is regulated through metabolism and is a key determinant in controlling the histone crotonylation level. Additionally, p300 mediates histone crotonylation and its activity shifts toward acetylation or crotonylation depending on the respective availability of the corresponding acyl-CoAs. For instance any decrease in acetyl-CoA levels enhances p300 crotonyl-transferase activity (Sabari et al., 2015). Although the crotonyl group may not be recognized by bromodomains (except BRD9 which binds weakly), another domain known as YEATS (Yaf9, ENL, AF9, Taf14, Sas5) domain can strongly bind to this

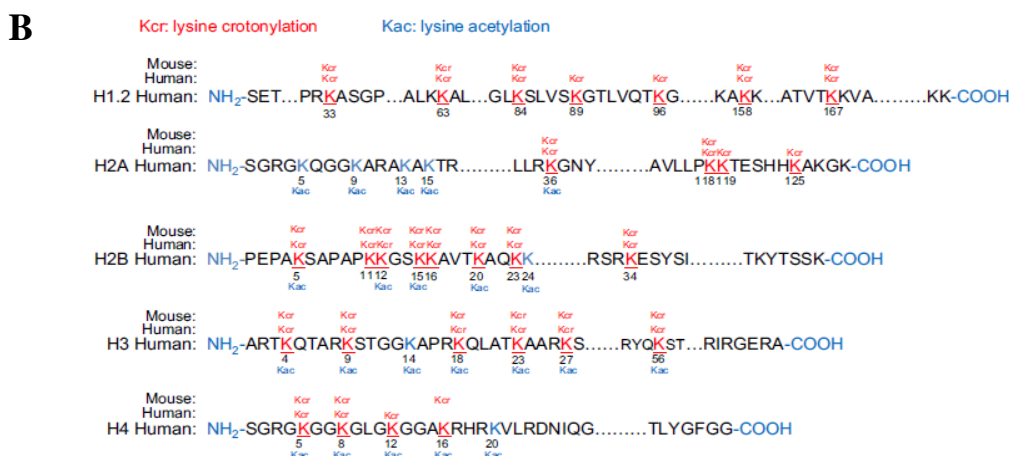


Figure 13: Short-chain lysine acylations. A) Comparison between crotonyl- and acetyl- groups. B) Lysine acetylation (Kac) and lysine crotonylation (Kcr) sites identified in human and mouse histones. Taken from (Tan et al., 2011).

3-2-2- 2-hydroxyisobutyrylation

The recent mass spectrometry study from Dai et al, on histones from mouse germ cells followed by HPLC analysis showed that the observed mass shift of +86.0354 at a histone lysine residue corresponds to $C_4H_7O_2$. Several proposed structures for this modification were 2-hydroxyisobutyrylation (K_{hib}), 2-hydroxybutyryl (K_{2ohbu}), 3-hydroxybutyryl (K_{3ohbu}), 3-hydroxyisobutyryl (K_{3ohibu}) and 4-hydroxybutyryl (K_{4ohbu}) groups (Figure 14). Further investigations finally demonstrated that this modification is Khib.

Khib has several unique features compared to Kac and Kcr. 63 lysine sites were identified in human and mouse histones with lysine 2-hydroxyisobutyrylation, 27 of these lysine sites were labeled with only Khib but not with acetylation or crotonylation.

Khib has a different structure and is of larger size than Kac, Kcr or Kme and it is not restricted to N-terminal tail domains, since it is also detected in other regions of the core histones. For instance, the stretch of residues between Lys 34 and Lys 120 in the globular domain of histone H2B provides 8 lysine sites that accept this mark, which are placed at the surface of the nucleosome. Histone Khib may thus have a particular role in regulating the inter- or intra- nucleosome interactions. Khib is conserved among the four tested eukaryotic cells, including Hela cells, Mouse embryonic fibroblast cells, *Drosophila* S2 cells and *Saccharomyces cerevisiae* cells with a stoichiometry of 5.33% and 7.97% for H4K91 and H1.2K62, respectively, which correspond to a higher stoichiometry than acetylation at various sites, for example compared to the 0.1% stoichiometry reported for H3K56ac.

It has been shown that genes associated with this new histone mark have high expression levels. The genomic location and the function of Khib in the context of male germ cell differentiation will be discussed in details in the next chapter (Dai et al., 2014).

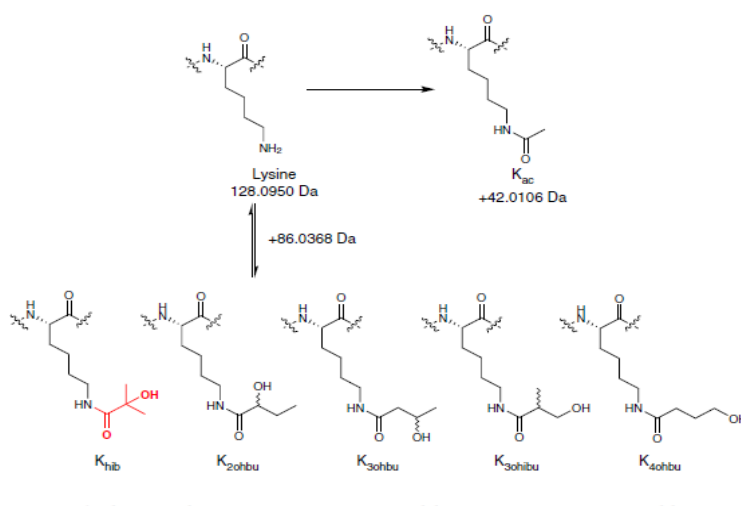


Figure 14: Several possible structures for the mass shift of +86.0368 Da. Taken from (Dai et al., 2014).

3-2-3- Butyrylation

Another novel histone modification is butyrylation, which is well-studied in the present work. The structural differences between acetyl lysine and butyryl lysine are expected to result in distinct functions of butyrylation compared to acetylation. To answer the important question concerning which enzymes can catalyze histone butyrylation, an *in vitro* study from Yingming Zhao's group in 2007, indicated that, among several tested acetyltransferases including CBP, p300, MOF and Tip 60, only CBP and p300 were capable of transferring [¹⁴C] butyryl-CoA to core histones. Furthermore, the same approach showed the occurrence of butyrylation by CBP and p300 on non-histone proteins like p53. In search for other non-histone proteins with butyrylation marks, they took advantage of the fact that the two acetyltransferases CBP and p300 acetylate themselves, and demonstrated that autobutyrylation of these two proteins expanded the list of butyrylated non-histone proteins (Chen et al., 2007).

A screen of human bromodomains containing proteins revealed that BRD9, CECR2 and the second bromodomains of TAF9 can bind to a butyryl mark with relatively good affinity. BRD4 have also shown a weak interaction with butyrylated peptides (Vollmuth and Geyer, 2010; Flynn et al., 2015).

Any misregulation in metabolic pathways and improper function of related enzymes could affect the nature of protein acylation. For instance, short-chain acyl-CoA dehydrogenase (SCAD) deficient mice as well as SCAD deficient fibroblasts show an increased level of protein lysine butyrylation. SCAD is an enzyme involved in fatty acid catabolism and its deficiency leads to accumulation of butyryl-CoA (Pougovkina et al., 2014). Butyrylation, since it is the focus of the current work will be discussed in details in the discussion section.

3-2-4- β -hydroxybutyrylation

β -hydroxybutyrylation is another type of newly discovered histone modification. β -hydroxybutyrate is a ketone body with remarkably elevated levels during starvation. β -hydroxybutyrate not only serves as an energy source, but also has been successfully used in the treatment of several neurological diseases such as epilepsy, Parkinson, and Alzheimer

(McNally and Hartman, 2012), yet the mechanisms of its action need to be clarified. A recent study from Zhao's group reported the existence of histones modified by β -hydroxybutyrate. Among the identified lysine sites are H4K8, H4K12, H3K4, H3K9, and H3K56 in mouse and human cells. A ChIP-seq analysis in liver cells indicated that Kbhb is associated with gene promoters and that H3K9bhb in particular is highly distributed at the TSS of active genes and present more or less the same abundance compared to H3K9ac and H3K4me3 on gene promoters.

Interestingly, a high level of β -hydroxybutyrate in mouse-starved during 3 days led to the enhancement of histone lysine β -hydroxybutyrylation (Kbhb), suggesting that cellular concentration of β -hydroxybutyrate can regulate the extent of this modification on histone proteins. Moreover, in the liver of fasting mice, Kbhb enhancement on H3K9bhb correlated to high levels of gene expression. Additionally, H3K9bhb-activated genes with no significant changes in H3K9ac levels showed different functions compared to genes induced by H3K9ac, suggesting that H3K9bhb could possess unique functions compared to acetylation (Xie et al., 2016).

3-2-5- Succinylation, Malonylation, and Glutarylation

Succinylation, Malonylation and Glutarylation are acidic acyl modifications that share some common features with acetylation in terms of structure but also possess several differences. In agreement with the presence of a carboxyl group in the structure of Succinyl-CoA, Malonyl-CoA, and Glutaryl-CoA (Figure 12), the addition of these modifications on an ϵ -amine of lysine residue alters the charge of the lysine from +1 to -1 under physiologic PH (Zhang et al., 2011), while acetylation neutralizes a lysine's charge from +1 to 0.

Lysine modified by these acyl groups loses its interaction with any molecules with a negative charged moiety in their structure such as nucleotides (phosphate group adds negative charge to nucleotide), proteins etc.

Succinyl-CoA, Malonyl-CoA, and Glutaryl-CoA derived from metabolic pathways are considered to be the donors of the corresponding acyl groups to lysines. SIRT5 from the sirtuins family was shown to regulate the removal of succinyl, malonyl, glutaryl groups. The

observed enhancement of lysine succinylation in SIRT5-deficient mouse embryonic fibroblasts and in SIRT5 KO mice strongly supported the previously proposed desuccinylase activity of SIRT5 (Park et al., 2013). It is however worthy to note that CobB, which is a member of the sirtuins family in bacteria, has also efficient desuccinylase as well as a deacetylase activities (Colak et al., 2013).

CBP/p300 have been shown to possess an *in vitro* succinyl-transferase activity (Hu et al., 2014), but acyltransferases capable of catalyzing succinylation, malonylation, and glutarylation *in vivo* remain to be identified.

The hypothesis that the levels of acyl-CoAs metabolic intermediates are the main drivers of lysine acylation obtained a strong support from the observed alteration of malonylation, butyrylation, and propionylation upon changes in the levels of the corresponding acyl-CoAs. Indeed, an increased lysine malonylation was reported in metabolic disorders i.e malonyl-CoA decarboxylase deficiency (Pougovkina et al., 2014).

A detailed study of succinylation in prokaryotes and eukaryotes showed its conserved nature in four tested organisms including Hela cells, yeast, bacteria, and mouse liver tissue. Additionally, succinylation can target the same sites as acetylation in mouse, yeast, and bacteria, since most of the succinylation sites in the mentioned species were also marked with acetylation (Weinert et al., 2013).

The screen of human bromodomains did not reveal any proteins from the BET family with the ability to bind the succinyl mark (Flynn et al., 2015). Therefore further work is required to identify the protein module responsible for the recognition of this PTM.

3-2-6- Propionylation

Propionyl-CoA is structurally similar to acetyl-CoA with one more CH₂ group. It is also the donor molecule in the propionylation reaction (Chen et al., 2007). Like acetylation, propionylation neutralizes the positive charge of a lysine and can be recognized by acetyl-lysine readers (bromodomain-containing proteins) (Flynn et al., 2015). Lysine propionylation has been found in histone and non-histone proteins such as CBP, p300, and p53. CBP/p300

can ensure a propionyltransferase activity *in vitro* and human homolog of yeast Sir2 is identified as the eraser (Liu et al., 2009). Overexpression of CBP/p300 in H1299 cells showed an increase in p53 propionylation levels, indicating their involvement in modifying p53 by propionyl group *in vivo* (Cheng et al., 2009).

Another work reported the presence of propionylation at H3K23 in the leukemia cell line U937, whose abundance decreased upon monocytic differentiation. H3K23pro is conserved from yeast to human (Liu et al., 2009; Zhang et al., 2009).

4- Mammalian spermatogenesis

4-1- Testis organization

From Books (De Jonge and Barratt, 2006, Cheng, 2008).

Spermatogenesis is the process that produces spermatozoa from undifferentiated male progenitor germ cells. The mouse male reproductive system is composed of a number of organs and glands including the testes, the extra-testicular ducts and the accessory sex glands. Spermatogenesis initiates in the male testis (from the Greek word *orchis*) at the beginning of puberty. The male gonads, the testes, display an oval structure and are localized in and protected by the scrotum. At the histological level, the testis is surrounded by the tunica albuginea and it consists of two main parts which are the interstitial (or intertubular) tissue and the seminiferous tubules. The interstitial compartment, which is between seminiferous tubules, contains Leydig cells, blood, macrophages and lymphatic vessels (Jan et al., 2012). The Leydig cells, upon stimulation by the pituitary luteinizing hormone (LH), produce testosterone as well as other steroids. Spermatogenesis, the differentiation of male germinal cells, takes place in the walls of the seminiferous tubules, from its periphery to its central lumen, where matures spermatozoa are released. Spermatozoa produced in testis are not immediately ready for fertilization, but it will terminate its maturation while passing through the epididymis, a long tube that connects each testis to the vas deferens, where the sperm cells are concentrated and stored. Epididymis is organized in three parts: 1) Caput (head) 2) Corpus (body) 3) Cauda (tail). Immature sperm cells enter the caput (through efferent ducts of the testis) and finally become mature in the cauda, before it moves towards vas deferens (Figure 15).

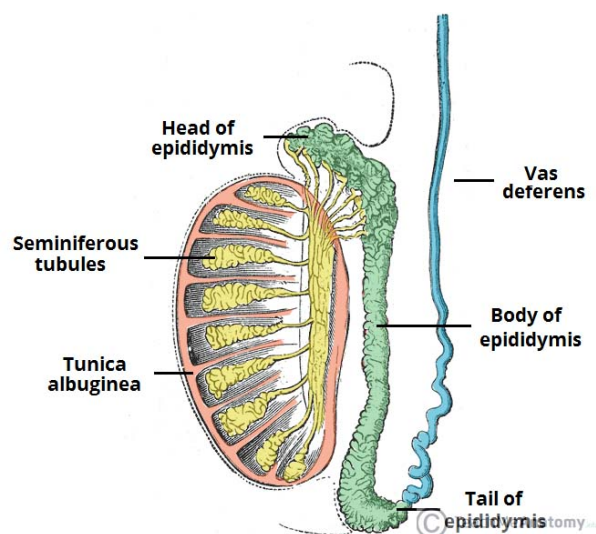


Figure 15: The structure of testis and epididymis. Taken from (teachmeanatomy.info).

4-2- Phases of spermatogenesis

The epithelium of the seminiferous tubules consists of a single type of somatic cell, named sertoli cell, and of germinal cells at different stages of their maturation. Sertoli cells form one layer that extend from the basal lamina (at the periphery of the tubule) towards the lumen part of seminiferous tubule with the function of nurturing and protecting germ cells ([Figure 16 and \(Cheng, 2008\)](#)). During their maturation process, the male germ cells progress between sertoli cells from the periphery (immature cells, spermatogonia) to the lumen of the tubule, where spermatozoa are finally released.

The process of spermatogenesis can be roughly subdivided to three different phases: the mitotic (proliferative phase) during which cells divide rapidly, the meiotic phase, which consists of two divisions following one replication and involve genetic recombinations, leading to the production of haploid cells, and finally, the post-meiotic phase also named spermiogenesis. During this final phase round haploid cells, named spermatids, progressively differentiate into specialized cells called spermatozoa, which are unique cells with the extraordinary ability to transport the male genome out of the body and deliver it to the female gamete. Progenitor adult stem cells (spermatogonia) proliferate mitotically and differentiate into primary spermatocytes, which then undergo two meiotic divisions. During the prophase of the first meiotic division, genetic recombinations ensure genomic shuffling. The production of secondary spermatocytes is rapidly followed by the generation of haploid round spermatids, which marks the initiation of spermiogenesis. In the post-meiotic phase, round haploid cells differentiate into elongating/condensing spermatids. The latter undertake a real metamorphosis to become mature spermatozoa which are then released into the lumen of seminiferous tubules ([Figure 16 and \(Gaucher et al., 2010; Rathke et al., 2013\)](#)).

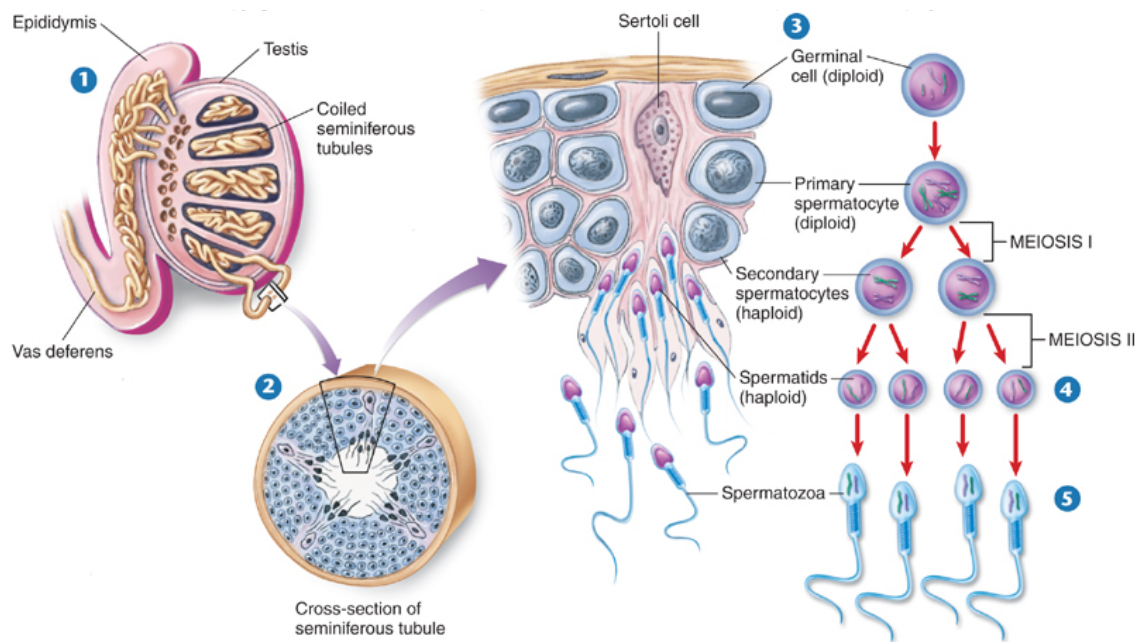


Figure 16: Testis morphology. 1) A cross section through testis illustrating seminiferous tubules, epididymis, and vas deferens. 2,3) A cross-section of seminiferous tubule showing germ cells. Taken from (bio.utexas.edu).

4-2-1- Mitotic phase

The mitotic phase of spermatogenesis ensures the maintenance of the spermatogonial cell pools (type A) on the one hand, and the supply of cells which will undergo meiosis (type B), on the other hand. Indeed, type-B spermatogonia mitotically divide to generate primary spermatocytes (de Rooij, 1998).

4-2-2- Meiotic phase

After mitotic division, the diploid type B spermatogonia forms preleptotene spermatocytes (the first meiotic cells) that could be distinguished from spermatogonia B by light microscopy

based on their size. These cells enter meiosis, a critical step of spermatogenesis whose successful completion will generate four haploid cells from each diploid cell. Haploid spermatids production is accomplished by one DNA duplication followed by two subsequent nuclear divisions. During the prophase of meiosis I, the paternal and maternal sister chromatids pair and move towards the pole of the cells which ultimately segregate into haploid cells after meiosis II.

The long prophase of the first meiotic division takes almost three weeks. It can be divided in four sub-stages with specific characteristics: leptonema (originated from the Greek words meaning “thin threads”), zygonema, pachynema and diplonema. As shown in [Figure 17](#), the most important events in meiotic prophase I include meiotic recombinations, which involve DNA double-strand breaks formation (DSBs) and the sex body formation, which is associated with global silencing of the sex chromosomes, named MSCI (this phenomenon is discussed more in details in section 7-1-1), etc.

Meiosis I initiates with leptotene stage in which chromatin becomes condensed and associates with critical proteins such as cohesins and synaptonemal complex (SYCP2, SYCP3) required for holding sister chromatids together and ensuring homologous chromosomes synapsis, respectively ([Page and Hawley, 2004](#)). In zygotene, homologous chromosomes progressively achieve full synapsis (following the synaptonemal complex assembly) and meiotic recombinations start to occur, leading to chromosomal crossing over formation at the pachytene stage. In contrast with autosomal chromosomes, the sex chromosomes, which share homology only in a small region, stay unsynapsed in non-homologous regions and form a structure known as the sex body.

The sites of genetic recombination between non-sister chromatids form the chiasmata. At the diplotene phase, the synaptonemal complexes disappear and the chiasmata become clear. After the completion of prophase I, homologous paired chromosomes line up their chiasmata on metaphasis plate (kinetochore microtubule) in metaphase I, and homologous chromosomes are pulled apart towards opposite poles during anaphase I, which leads to the generation of two daughter cells, each one containing half of the chromosomes (each chromosome has two sister chromatids). Shortly after cells enter meiotic II division, during which the sister chromatids of each chromosome are pulled apart (in a way that is very similar to mitosis) resulting in the generation of four haploid daughter cells.

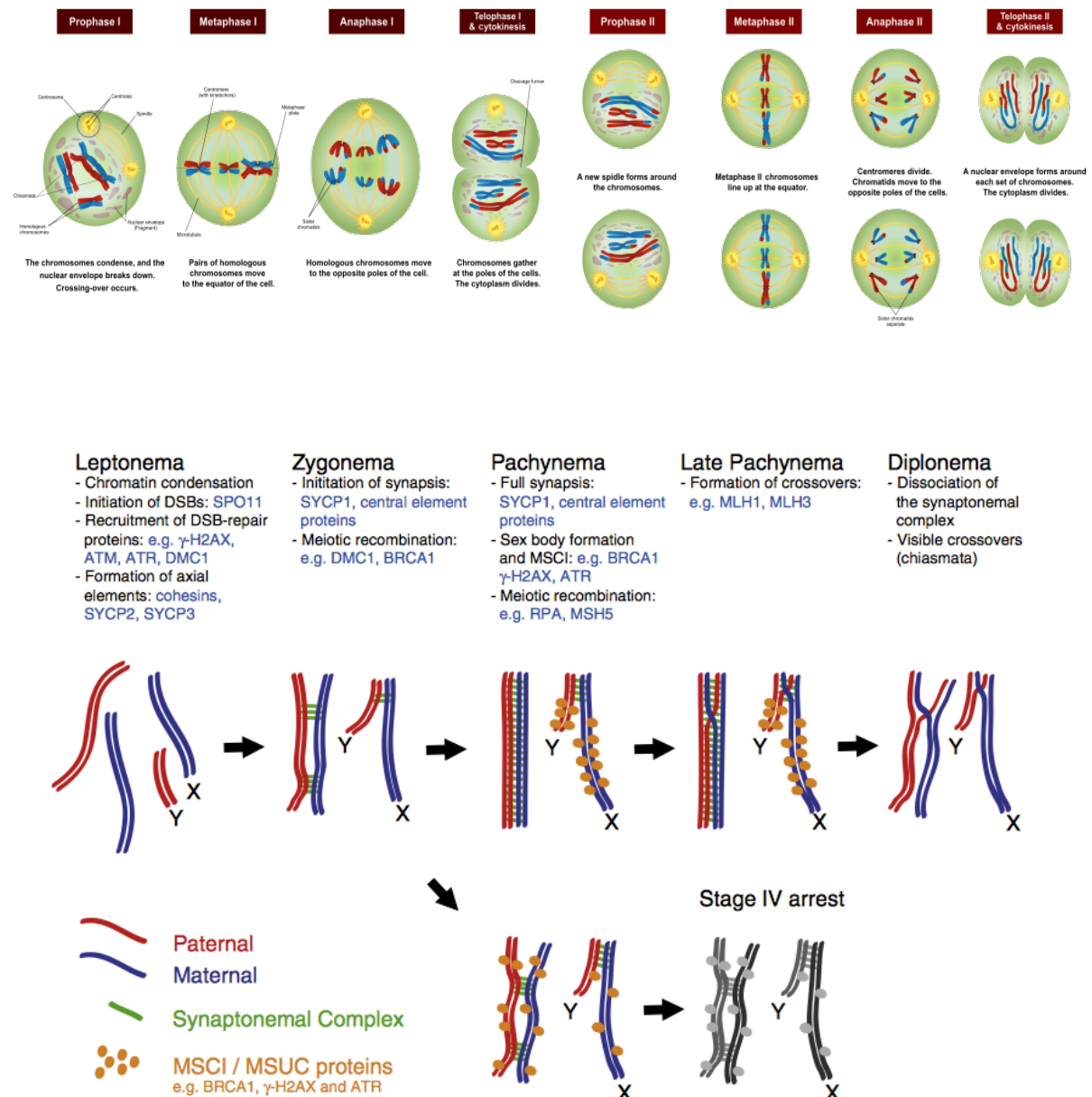


Figure 17: Meiotic phase of spermatogenesis. On the top panel, all the stages of meiosis I and meiosis II are illustrated. Taken from (https://commons.wikimedia.org/wiki/File:Meiosis_Stages.svg). On the lower panel, illustration of meiosis prophase I sub-stages: Leptonema, Zygotene, Pachytene, Late pachytene, Diplotene, together with important events including homologous chromosomal pairing, chromosomal crossover and meiotic sex chromosome inactivation (MSCI). The improper pairing and synapsis of homologous chromosomes prevents MSCI and stops the progression of meiosis at stage IV of mouse spermatogenesis. Taken from (Jan et al., 2012).

4-2-3- Post-Meiotic phase

After the successive meiotic divisions, the haploid round spermatids enter the post-meiotic phase and undergo extensive morphological changes during which the spherical round spermatids will progressively change shape and undergo morphological changes, becoming elongating and condensing spermatids, then mature spermatozoa. This transformation involves four phases: Golgi phase, Cap/Acrosome phase, tail formation, and maturation phases.

The Golgi apparatus is involved in the acrosome formation by producing vesicles and granules. At the beginning of spermiogenesis, haploid spermatids have only a perinuclear Golgi apparatus and no acrosome. As cells progress, the Golgi apparatus produces proacrosomal vesicles which contain proacrosomal granules. Then, proacrosomal vesicles combine to form the acrosomal vesicle with acrosomal granules inside it. Capping refers to the process during which the acrosomal vesicle (containing acrosomal granule) reaches the nuclear envelope and becomes flat over the nucleus surface to create the acrosome. The acrosome, which is placed at the sperm's head contains proteolytic enzymes that will be released upon the acrosome reaction, which occurs during the progression of the spermatozoa through the zona pellucidae surrounding the oocyte, and the following proteolysis of the protective "coat" of the egg, allowing gamete fusion to occur. The flagellum (or tail) development is initiated by the formation of an axoneme from one of the two centrioles migrated to the cell surface. The axoneme with microtubular structure causes the plasma membrane of the spermatid to extend out of the cell and to form flagellum that receives additional components afterwards. The formed tail can be divided into three parts including the middle, principle, and end pieces. The flagellum is necessary for sperm motility, which is required to accomplish its mission to travel out of male body and deliver the male genome to the female oocyte. During the final maturation phase, flagellum formation is accompanied by the removal of excess cytoplasm, which form "residual bodies". Towards the end of their maturation, the post-meiotic cells face two important events including an arrest of gene transcription and histone exchanges by transition proteins and protamines. The replacement of histones initiates the genome compaction process leading to the generation of a condensed and compact nucleus. Finally, condensed spermatids become spermatozoa, which release in the lumen of the seminiferous tubules and then will reach the epididymis where they achieve their final maturation and acquire motility.

4-3- Stages of spermatogenesis

Mouse germ cells differentiation is classified into 12 stages that correspond to one cycle of maturation of spermatogonia to spermatozoa (Figure 18). This classification is based on the cell associations in each section of the seminiferous epithelium of the tubular testis. This classification was established by Oakberg using periodic acid-Schiff technique (PAS-staining) of testis paraffin cross-sections. In general, the acrosome development in round spermatids and the nucleus morphology changes of spermatids during the elongation and condensation process, are among the most important criteria to determine the stages.

Stage I-VIII:

Round/elongating spermatids and spermatocytes form the cell populations of the stage I-VIII. In stage I, round spermatids look like spermatocytes except that the formers are smaller in size with small nucleus and also have a nucleolus. The acrosome is not formed in this stage, the Golgi is small and without granular material and the spermatid heads are shorter compared to those of stage 12.

Stage II-III:

The Golgi starts to show proacrosomal granules in its center, which forms an acrosomal granule at stage III. The very dense chromatin is visible in pachytene spermatocytes.

Stage IV:

The acrosome granule starts to flatten over the nucleus surface.

Stage V-VIII:

The acrosome length grows until it covers half of the nucleus surface of spermatids at stage VIII. Pachytene spermatocytes show the sex vesicle and become larger. Pre-leptotene spermatocytes appear at stages VII and VIII.

Stage IX-XII:

The spermatids nucleus begin to elongate, so they become elongating spermatids. Leptotene and zygotene spermatocytes appear during stages IX-X and XI-XII, respectively. The presence of secondary spermatocytes is another important feature of stage XII.

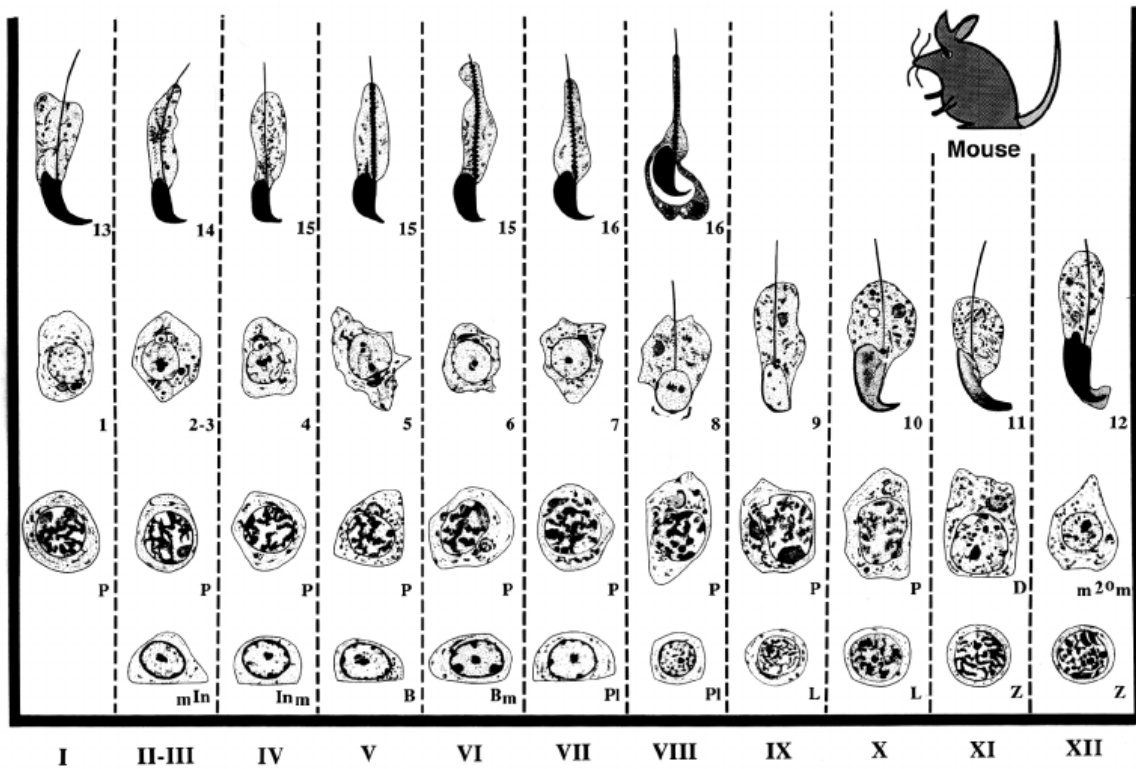
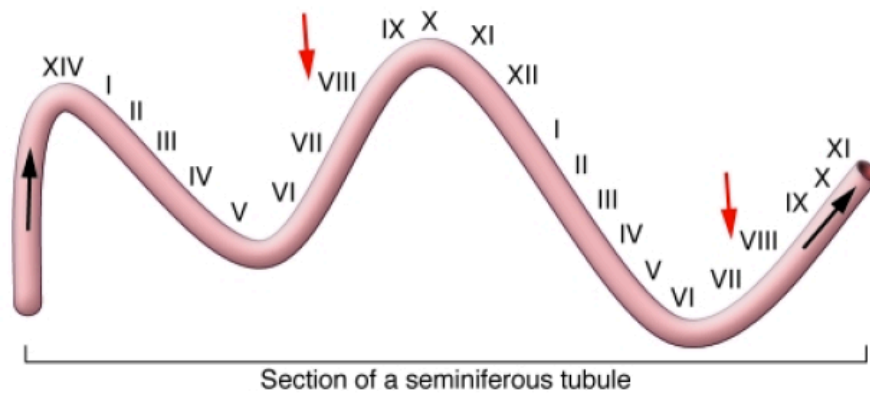


Figure 18: The cycle of the seminiferous epithelium and stages in mouse spermatogenesis. The top panel shows stages are shown along the length of a seminiferous tubule. Taken from (Hogarth and Griswold, 2010). The bottom panel illustrates the twelve stages of spermatogenesis (I-XII) and associated germ cells within each stage. Taken from (Russell et al., 1990).

5- Chromatin structure and gene expression during spermatogenesis

Germ cells undertake the transmission of the parental genetic information from generation to generation. Male germ cells differentiation relies on a precise and highly regulated gene expression profile, which is essential for driving and controlling the complex cellular processes occurring during the three different phases of spermatogenesis (mitotic, meiotic, post-meiotic) (Shima et al., 2004). A number of specialized regulatory pathways makes germ cells different from somatic cells, even though they both have the same DNA sequence (Kimmins and Sassone-Corsi, 2005; Ooi and Henikoff, 2007). The germ line-specific gene expression program, which controls spermatogenesis progression encompasses genes specifically expressed in spermatogonia, spermatocytes and spermatids during meiotic and post-meiotic phases, respectively (Montellier et al., 2012).

It is of note that a transcription shut down occurs at a defined point, approximately in the middle of the post-meiotic maturation process in the elongating spermatids which is before histone removal and replacement. The completion of spermiogenesis hence depends on the storage of mRNAs of genes that need to be expressed in post-meiotic phases and on a translational repression, which prevents untimely expression and production of the corresponding proteins (Figure 19 and (Rathke et al., 2013)).

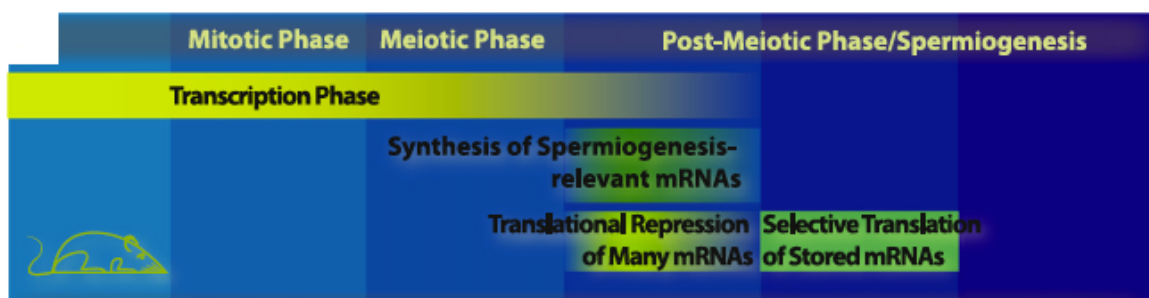


Figure 19: Regulation of gene expression in mouse spermatogenesis. Schematic illustration of the arrest of transcription occurring in transition from round spermatids to elongating spermatids and also mRNAs synthesis and storage of genes required to be expressed in post-meiotic cells, which will become transcriptionally silent. During later spermatogenesis. Adapted from (Rathke et al., 2013).

5-1- Gene expression regulation

Spermatogenic cell gene expression program is controlled through intrinsic genetic program and also extrinsic regulators. Testosterone and follicle stimulating hormone (FSH) are among the extrinsic factors having indirect effect on spermatogenesis.

Gene expression is regulated at three levels: transcription, translation and post-translation via intrinsic regulators. In general, transcription initiates upon the binding of transcription factors through their DNA binding domains (DBDs) to the promoter of the transcribed gene DNA sequence, which can regulate the activity of the transcriptional machinery. According to their expression and distribution in tissues, transcription factors belong to two groups: 1) a group of transcription factors ubiquitously expressed that are capable of regulating gene expression in many tissues. 2) tissue-specific transcription factors that are in charge of the expression of genes only in specific tissues.

Germ cell's transcription factors possess some differences in terms of their amounts and specificity compared to somatic cells. For instance, early haploid cells have higher levels of TBP (TATA-binding protein), TFIIB, and RNA polymerase II than somatic cells ([Sassone-Corsi, 2002](#)). Transcription factors specific to spermatogenic cells are SPRM1, TAK-1, ZFY-Z, OCT-2. Different isoforms of two transcription factors such as cAMP-responsive element binding protein (CREB) and cAMP-responsive element modulator (CREM) are present in testis, with the function of gene expression activation or repression during specific phases of spermatogenesis.

Promoter-binding proteins in the nuclei of germ cells regulate gene expression through their ability to bind to DNA promoter sequence motifs ([Eddy and O'Brien, 1998](#)). Histone H1t TE-element binding proteins ([vanWert et al., 1996](#)) and the phosphoglycerate kinase 2 promoter-binding proteins ([Gebara and McCarrey, 1992](#)) are considered as positive effectors of gene expression during spermatogenesis. C-mos is a protein expressed in germ cells but not in somatic cells. The tissue-specific expression of c-mos is due to the binding of a protein (specific to somatic cells) to Negative regulatory element (NRE) of c-mos promoter, thereby repressing c-mos expression in somatic cells ([Xu and Cooper, 1995](#)). In fact, several proteins such as protamine 2, phosphoglycerate kinase 2, cytochrome c_T and heat shock protein Hst 70, with germ cells specific expression have shown similar sequence to the NRE of *c-mos* in their

genes promoter which could be a reason why they are silent in somatic cells. Late haploid spermatids lose their transcriptional ability upon histone removal and nucleus condensation taking place in steps 9-12 of spermatogenesis. Germ cells strategy to keep essential proteins expressed once transcriptional shut down has occurred is through storing synthesized mRNA in an inactive form with the ability to become active at anytime when they are needed. Thus gene expression regulation at the mRNA level seems to be carried out by mRNA binding proteins that can interact and delay their translation (Kleene, 1996). For instance, protamine 1 expression in spermatids is regulated by the association of several proteins to the 3' untranslated regions (UTR) of protamine 1 mRNA (Braun et al., 1989). In addition, different isoforms of proteins could be expressed in spermatogenic cells via alternative transcripts generation.

5-1-1- Meiotic sex chromosome inactivation

An interesting event in the course of spermatogenesis is the meiotic sex chromosome inactivation, particularly known as “meiotic sex chromosome inactivation” (MSCI), which is essential for germ cells differentiation. Indeed, it has been shown that spermatocytes incapable of silencing their sex chromosomes die through apoptosis (Turner et al., 2004; Turner et al., 2006).

Sex chromosome transcriptional silencing contributes to germ line-specific gene expression program. During meiotic phase of spermatogenesis, homologous chromosome synapsis (between autosomal chromosomes) in zygotene spermatocytes is associated with the formation of the synaptonemal complex which also assists the crossing overs between them. In this context, the pairing of the X and Y non-homologous chromosomes (that tend to localize in the peripheral nucleus region) occurs only in the “pseudoautosomal region” of these chromosomes, which ultimately form the sex body in pachytene spermatocytes (Hoyer-Fender, 2003; Handel, 2004; Yan and McCarrey, 2009). The sex body, a specific chromatin region made up of transcriptionally inactive sex chromosomes, is characterized by massive histone exchanges, incorporation of specific histone variants and specific changes in histone modifications (PTMs). MSCI starts with phosphorylation and ubiquitylation of histone variant H2A.X and histone H2A respectively. Histone macroH2A incorporation, the removal and replacement of histone H3.1 and H3.2 by H3.3 follow, accompanied by the accumulation of

H3k9me and recruitment of HP1 isoforms. In post-meiotic round spermatid cells, despite the proximity of sex chromosomes to the repressive environment consisting of the chromocentre region (formed by the accumulation of chromosomes centromere) in the nucleus, a group of previously repressed sex-linked genes are re-activated.

Hrb6 and Sly are among the factors that could contribute to the sex chromosome gene repression via the maintenance of repressive histone marks either on the chromocentre or on the sex chromosomes. Indeed, the chromosome-wide transcriptional silencing could inhibit the expression of genes that are essential for the male germ cells differentiation progression, thus it is critical to reactivate these fractions of genes. The strategy of post-meiotic germ cells to overcome this issue is through the gain of active histone marks, for instance H3K4me3 and H4K8ac/K12ac and the formation of unstable nucleosomes by H3.3 and H2A.Z assembly (Figure 20 and (Montellier et al., 2012)).

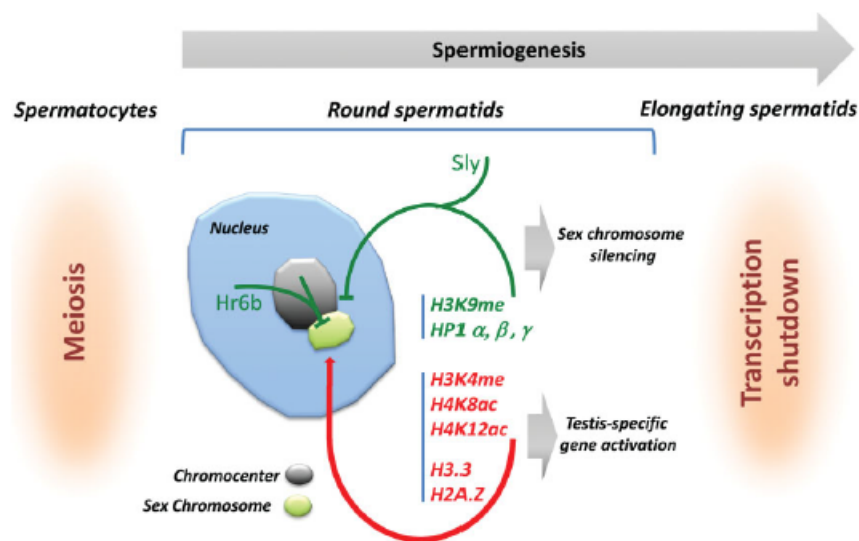


Figure 20: Transcriptional silencing of sex chromosomes during meiotic phase of spermatogenesis. Taken from (Montellier et al., 2012).

5-2- Chromatin remodeling and genome wide histone removal

The reorganization of chromatin is a remarkable multistep process that reshape the structure of haploid spermatids nucleus from round to elongated and condensed shapes. This chromatin remodeling and the associated morphological changes are essential to the life cycle ([Kimmins and Sassone-Corsi, 2005](#)). Over the past two decades, some of the critical elements that accompany chromatin remodeling in late spermatogenesis have been identified. Histone modifications, histone variants, readers of histone modifications including Brdt, and transient DNA strand breaks are the inducers of these genome reorganizations, yet the mechanisms/regulation and contribution of these elements in packing the male genome, still need to be unraveled.

During spermatogenesis, nearly all of the nucleosomal histones are removed and replaced by small basic proteins specific to germ cells in a step-wise manner, first transition proteins (TPs), then protamines (Prms) ([Figure 21A and \(Boussouar et al., 2008; Gaucher et al., 2010\)](#)). Classically, transition proteins (an important protein of haploid spermatids) assemble into chromatin after histone removal has occurred, and it is thought that they might have a role in preparing DNA for protamine incorporation ([Meistrich et al., 2003](#)). Still many investigations should be carried out to understand the molecular basis underlying histone-to-Prm exchange. Questions such as “how the removed histones could be degraded” and “what is the importance of remaining histones in mature spermatozoa” are still awaiting answers ([Goudarzi et al., 2014](#)).

All the factors indispensable for this massive histone removal and their replacement with non-histone proteins need to be expressed in a right timing ([Figure 21B](#)).

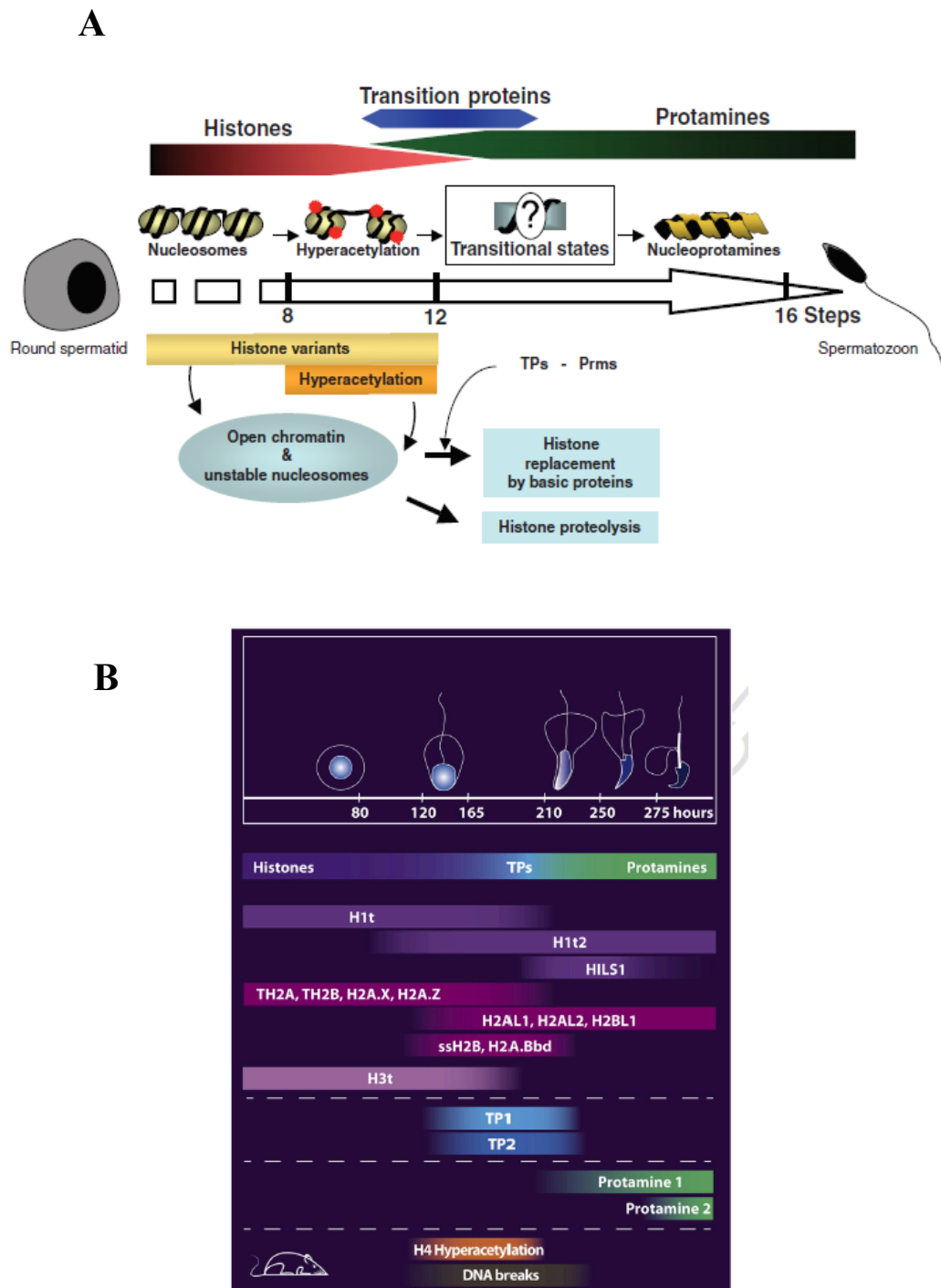


Figure 21: Main features of mouse spermiogenesis. **A)** Model of the chromatin reorganization in elongating spermatids accompanied by massive histone hyperacetylation, incorporation of histone variants, and recruitment of Brdt to acetylated histone (not depicted). An almost complete removal of histones occurs and shifts chromatin structure from nucleosomal-based to protamine-based structures. Taken from (Gaucher et al., 2010). **B)** Schematic representation of the timing of the expression of critical chromatin components that contribute to spermatid differentiation in mice. Taken from (Rathke et al., 2013).

5-2-1- Histone modifications

The drastic histone displacement by TPs and Prms requires mechanisms with the ability to relax the chromatin structure. Open chromatin could allow critical machineries to reach and further open chromatin domains in order to accomplish the eviction of histones (Voigt and Reinberg, 2011). Histone modifications are known to alter chromatin configuration and contribute to either open or close it. Indeed, depending on the nature of the acyl-moiety that histones are modified by, and also on the subsequent events through the function of proteins that recognize these modifications, chromatin can undergo important structural transitions. Among the common and known chromatin modifications, acetylation, methylation, ubiquitination and phosphorylation are reported to occur on histones and might have important roles in male genome reorganization. Phosphorylation of H2A.X and poly(ADP) ribosylation of proteins observed in the nucleus of elongating spermatids are the responses to the DNA strand breaks during spermatid elongation (Laberge and Boissonneault, 2005; Meyer-Ficca et al., 2005; Leduc et al., 2008). Following the discovery of 67 novel histone modifications (Tan et al., 2011), we still have a long way to go to clarify how these PTMs could contribute to the disappearance and replacement of histones. In various species, a genome-wide histone hyperacetylation precede histone eviction and replacement by TPs (Hazzouri et al., 2000; Govin et al., 2004; Boussouar et al., 2008). Histone acetylation, more particularly that of histone H4 associates with histone removal through nucleosome destabilization (altering histone-DNA and histone-histone interaction) and recruitment of chromatin binders capable of coordinating histone-to-Prm exchange.

In addition to acetylation, other modifications such as methylation and ubiquitination could also be considered to have a role in assisting chromatin reorganization during spermiogenesis. H3K4/K9 mono-, di-, and trimethylation as well as H3K27 di- and trimethylation have been observed in elongating spermatids of mammals and flies (Godmann et al., 2007; Rathke et al., 2007). H2A/H2B ubiquitination regulated by RNF8 is enhanced during meiosis and in elongating spermatids in mice (Chen et al., 1998; Baarends et al., 1999). RNF8-deficient mice, lacks H2A/H2B ubiquitination, and thus fails in histone replacement (Lu et al., 2010). However this observation is controversial since Sin et al showed that histone replacement occurs in RNF8-deficient mice (Sin et al., 2012).

It is speculated that the timing of histone H4 acetylation is under the control of H4S1 phosphorylation, since a decrease in H4S1 phosphorylation promotes acetylation of histone H4. This observation suggests the existence of histone PTMs cross-talks during spermiogenesis (Govin et al., 2010). Genome-wide hyperacetylation involves Lysine5, Lysine 8, Lysine 12 (K5, K8, K12,) of histone H4 in elongating spermatids. H4K16 appears already acetylated before the wave of acetylation.

The cause of histone hyperacetylation in elongating spermatids is still unclear. Hyperacetylation is likely to correlate to the observed decrease in the level of HDAC1 and HDAC2 in elongating spermatids (Caron et al., 2003). The possible involvement of HDACs in the occurrence of histone hyperacetylation in late spermatogenic cells was also supported by the observed histone hyperacetylation induced by treating round spermatids by the HDAC inhibitor TSA (Hazzouri et al., 2000). Although histone acetyltransferases could also have a role in the establishment of this histone hyperacetylation, the investigation of several HATs such as CBP/p300, Cdy1, and Mof has not provided any proof yet (Caron et al., 2003; Sin et al., 2012; Boussouar et al., 2014).

The recent investigation of some of the new histone modifications in the context of male germ cell differentiation, and more specifically with respect to chromatin reorganization, provides insights that, besides acetylation, many other histone PTMs, including crotonylation, butyrylation and 2-hydroxyisobutyrylation, occur on chromatin right before histone replacement (Tan et al., 2011; Dai et al., 2014; Goudarzi et al., 2016).

5-2-2- Brdt mediates histone removal

Histone hyperacetylation is a sign of histone removal but the molecular basis underlying this phenomenon is unknown. In 1999, Dhalluin et al showed that bromodomains can recognize and bind acetylated histone tails (Dhalluin et al., 1999). Then, an in silico study in search for testis-specific bromodomain-containing proteins identified Brdt, a factor with two bromodomains belonging to the BET family as a good candidate for binding acetylated histones and remodelling hyperacetylated chromatin (Pivot-Pajot et al., 2003). The multifunctional Brdt gene is expressed once spermatogenic cells enter the meiosis phase. Brdt is driving the specific male gene expression program as well as histone removal in post-

meiotic phase. Several observations support the speculation that the first bromodomain of Brdt targets acetylated histone H4 and mediates the replacement of histone to TP (Moriniere et al., 2009; Sasaki et al., 2009; Gaucher et al., 2012). The structural studies of Brdt's two bromodomains revealed that acetylation of both lysines 5 and 8 of histone H4 is indispensable for the interaction of Brdt's first bromodomain with acetylated histone (Moriniere et al., 2009). Further investigation by FRET analysis in living cells, provided another support to the structural studies of Brdt (Sasaki et al., 2009). A study on knock-out and BD1-less Brdt mice demonstrated that Brdt is a critical factor in male germ cell differentiation. Brdt^{ΔBD1/ΔBD1} male mice were sterile with normal hyperacetylation wave but impaired histone removal. Interestingly, Brdt-deficient mice cannot fully complete meiosis I and show spermatogenic arrest in late meiosis phase. Misregulation of many genes in the absence of Brdt and the observed interaction of the P-TEFb complex with the C-terminal domain of Brdt support Brdt's involvement in driving the highly specific male germ cells gene expression program (Gaucher et al., 2012). Taken together, these observations strongly point to Brdt as an important factor capable of regulating transcription and remodeling chromatin in the course of spermatogenesis.

5-2-3- Histone variants

Male germ cells express nearly all of the variants of the histones H2A, H2B, H3 and H1 (Figure 21B). Mouse testis-specific H2A and H2B variants are H2AL1, H2AL2, H2AL3, H2BL1, H2BL2, and TH2B (Govin et al., 2007). Structural and functional studies of the testis specific histone variants H2AL2, TH2B, H3T, and H2ABbd revealed that the incorporation of these variants to nucleosomes leads to nucleosome instability compared to nucleosomes containing canonical histones (Gautier et al., 2004; Govin et al., 2007; Li et al., 2005; Tachiwana et al., 2010).

It is very tempting to propose that H2AL1 and H2AL2 prefer to incorporate in nucleosome-containing TH2B than H2B to generate unstable nucleosomes in spermatogenic cells (Govin et al., 2007). It is also of note that, H2A.Bbd, which is the human homolog of H2AL2, has the capacity to destabilize nucleosomes (Gonzalez-Romero et al., 2008). Additionally, somatic-types of histone variants such as H2A.Z and H3.3 can affect the stability of nucleosomes in spermatogenic cells (Boussouar et al., 2008; Jin et al., 2009).

5-2-4- Transition proteins

Transition nuclear proteins (non-histone DNA-packaging proteins) are considered as intermediate chromatin components (about 90% of the basic chromatin components) between histone displacement and protamine deposition (Meistrich et al., 2003). Two prominent transition proteins in the spermatids of rodents are TP1 and TP2 (Grimes et al., 1977) and constitute approximately the same percentage of total nuclear proteins in spermatids which is 55% of TP1 and 40% of TP2.

TP1, a 6.2 k Da protein, contains 20% arginine and 20% lysine, without cysteine and is conserved in different mammals (Kistler et al., 1975; Kremling et al., 1989), whereas TP2, a 13 k Da protein, with 10% arginine, 10% lysine, and 5% cysteine is not well conserved. TP1 and TP2 appear to be expressed at different time (earlier) than Prm1 and Prm2, but their transcription occurs at the same time as protamines, which is at stage 7 of mouse spermatogenesis (Mali et al., 1989). TPs and Prms start to appear at stages 10-11 and 12-13, respectively (however this timing is not clear yet).

Some *in vitro* studies provided useful information on the mechanisms underlying TP actions, such as TP1's ability to decrease the melting temperature of DNA. TP1 could thus induce DNA relaxation and facilitate the function of topoisomerase I (Singh and Rao, 1988; Akama et al., 1998; Akama et al., 1999). Others reported that none of the two TPs, TP1 and TP2 could exert topological changes on supercoiled DNA (Levesque et al., 1998). It is also proposed that TP1 stimulates single-strand DNA breaks and ligation (Caron et al., 2001).

Further analysis of TPs *in vivo* in mice expressing only one of the TPs, either TP1 or TP2, revealed that first TPs are not involved in histone removal since the histone exchange and protamine assembly occurs normally in these mice, second these two proteins have redundant functions. Most of the TP1-deleted mice are sterile (about 60%), with sperm mobility defects, and increased levels of TP2 (Levesque et al., 1998; Yu et al., 2000). In contrast, TP2 knock out in mice does not affect fertility and leads to less severe abnormalities (Zhao et al., 2001). In the absence of both TPs, mice are sterile with severe chromatin condensation defects but normal histone removal and protamine deposition (Zhao et al., 2004a; Zhao et al., 2004b; Rathke et al., 2013).

5-2-5- Protamines

Chromatin protamination follows the sequential exchange of histones by small basic nuclear proteins: TPs and Prms, which drives chromatin remodeling during spermiogenesis. First destabilized histones (facilitated by the occurrence of post-translational modifications and histone variants incorporation into chromatin) are replaced by transition proteins that are considered as intermediate components in histone replacement. Then protamine pack the genome (Figure 22).

Protamines are thought to have important roles in (i) compacting the genome, (ii) protecting the male genome, (iii) ensuring the transcription suppression, (iv) eventually ensuring successful reproduction (Takeda et al., 2016). Humans and mice express two types of protamines: protamine 1 (PRM1 or P1) and protamine 2 (PRM2 or P2) (Braun, 2001).

Whereas protamine 1 is expressed in all the mammals, protamine 2 possess restricted expression only in some species (Domenjoud et al., 1991). Protamines are small, highly basic proteins enriched in arginine and cysteine. Although protamine 1 is expressed as mature protein (51 amino acids), protamine 2 is synthesized as a precursor protein (102 amino acids) and becomes mature by proteolytic cleavage of its N-terminal (57 amino acids) (Balhorn et al., 1999).

The basic nature of protamines allows their tight interaction with negatively-charged DNA molecule. Additionally, disulfide bridges formed between cysteine residues link protamines together, which is critical for the tight genome packaging.

Phosphorylation of protamines facilitates their interaction with DNA. Prm1 phosphorylation occurs right after its translation through the activity of serine/arginine protein-specific kinase 1, while Ca^{2+} /calmoduline-dependent protein kinase IV (Camk4) is the main regulator of phosphorylation of Prm2 (Green et al., 1994; Papoutsopoulou et al., 1999). Disruption of Camk4 impairs TP-to-Prm exchange and spermiogenesis, leads to the loss of protamine 2 and longer persistence of TP2 in late spermatids. Camk4^{-/-} male mice are infertile (Wu et al., 2000; Braun, 2001). Prm2 phosphorylation is not only essential for its interaction with DNA but it is also required for the processing of Prm2 through proteolysis.

One of the interesting features of protamine biology is the mechanism of translational regulation, the delay between transcription and translation, that allows the protamine proteins to appear in the right timing. To accomplish this, protamine mRNAs synthesized in round spermatids do not translate until spermatid elongation occurs. Protamine transcribed mRNAs are stored translationally inactive, as messenger ribonucleoprotein (mRNP) particles in chromatoid bodies (Steger, 2001; Steger et al., 2001). For instance, protamine 1 mRNA translation is repressed by the binding of protamine 1 RNA binding protein (PRBP), which is encoded by the gene *Tarbp2*, to the 3'-UTR of protamine 1 mRNA. Mice with a disrupted *Tarbp2* gene are sterile because of the delay in transition proteins replacement (Zhong et al., 1999). In elongating spermatids, translational derepression of Prm1 and Prm2 depends on the shortening process by poly A-deadenylation (Steger, 1999).

Mutation in one allele of either Prm1 or Prm2 severely interferes with normal sperm production, protamine 2 processing, and results in the failure of genetic transmission of both mutant and wild-type alleles. In contrast to TPs, protamines cannot be considered as functionally overlapping proteins. A decrease in Prm2 reduces Prm1 and Prm2 levels in the sperm which is not the case for Prm1 chimeras. These findings suggest that Prm1 and Prm2 are critical elements for producing intact sperm (Cho et al., 2001).

6- Fertilization and zygote formation

Fertilization is the fusion of sperm and oocyte, first producing a zygote which then develops into an embryo. Sperm should pass through the different layers surrounding the oocyte, the cumulus cell barrier and the zona pellucida. Sperm penetration into zona pellucida occurs through acrosome reaction, a process, which results in the secretion of the acrosomal enzymes that are required for dissolving zona pellucida. Serine protease and acrosin are some of these enzymes. The microvilli in the oocyte surface are also important in the step before sperm-oocyte fusion takes place (Figure 23).

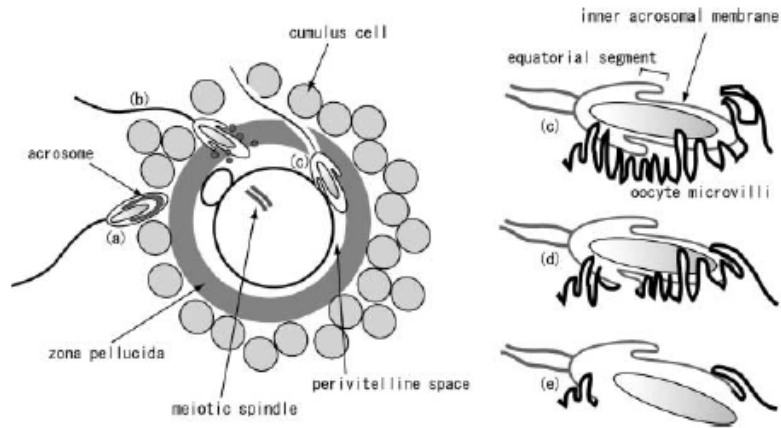


Figure 23: The process of fertilization, sperm and egg fusion. On the left, different layers surrounding oocyte and the steps in sperm penetration are shown. On the right, illustration of oocyte microvilli in anchoring sperm to the egg. Taken from (Kaji and Kudo, 2004)).

Thesis objectives

As described in the introduction, Spermatogenesis is a unique process, which occurs in three different phases and is characterized by stage-specific gene expression programs. Male genome reorganization in the post meiotic phase of spermatogenesis involves an almost complete replacement of histones, first by transition proteins (TPs) and then by protamines (Prms). Two events precede histone replacement by TPs and protamines: first the replacement of canonical histones by specific variants and second a global increase of their acetylation levels. The histone hyperacetylation wave occurs in early elongating spermatids after a general shut down of transcription. Considering the importance of acetylation in male genome reprogramming and the possible role of HDACs (reduction in the level of HDAC 1 and HDAC 2 in elongating spermatids) in its establishment, it is not still clear whether histone acetyltransferases (HATs) could also have a role or not. Following a systematic search for bromodomain containing proteins highly or specifically expressed in testis, my host laboratory identified Brdt, the testis specific member of the double bromodomain containing proteins of the BET family. One of the main outcome of our laboratory's work on the structure and function of Brdt's bromodomains is that Brdt binds a specific combination of acetylated histone tails, and that its first bromodomomain, BD1, specifically recognizes the hyperacetylated form of histone H4 ([Moriniere et al., 2009](#)). Additionally, several studies have shown that histone post-translational modifications (PTMs) play an important role in the regulation of chromatin structure and function. Indeed, recent work from Yingming Zhao's laboratory, in collaboration with our team, revealed that many other histone post-translational modifications (PTMs) exist and are potentially involved in genome programming ([Tan et al., 2011](#)). My host laboratory performed detailed investigations of crotonylation and 2-hydroxybutyrylation during spermatogenesis ([Tan et al., 2011](#); [Dai et al., 2014](#)). In general, my PhD project aimed at a better understanding of the role of histone post-translational modifications (PTMs) in male genome reprogramming.

More specifically, the objectives of my PhD work was to address the following questions:

- 1) Role of the Histone Acetyl Transferases (HAT) CBP and p300 in controlling the male germ cells gene expression program and the histone hyperacetylation wave preceding histone to protamine replacement.**
- 2) Functional significance of butyrylation as an unknown histone post translational modification**
- 3) The mechanisms by which p300 catalyze lysine acylations**

Results

1- Role of CBP and p300 in male gene expression program and histone hyperacetylation wave.

Despite the important role of the hyperacetylation wave in post-meiotic phase of spermatogenesis, nothing is known about the enzymatic machineries, which control this phenomenon. CBP and p300 are ubiquitously expressed histone acetyltransferases (HATs). Therefore, in order to investigate the role of CBP and p300 in this hyperacetylation wave and in the gene expression program during post-meiotic phase of spermatogenesis, we generated conditional double knockout mice $cbp^{\text{flox/flox}}$ and $p300^{\text{flox/flox}}$ which were crossed with Prm1-Cre mice, expressing Cre recombinase under the control of the protamine promoter, aiming to delete these genes specifically in haploid germ cells. In this part of my thesis, I was interested to investigate the function of two HATs CBP/p300 during post-meiotic male genome programming.

Main results obtained

1-1- Expression of CBP/p300 during spermatogenesis

Considering the fact that each stage of spermatogenesis occurs in a precise day post-partum, testes of mice were collected at different days post-partum (dpp): pre-leptoten and leptoten spermatocytes appear at 10 days post-partum, round spermatids accumulate at 20 dpp, which finally give rise to mature spermatozoa, the above-mentioned developing testes and also fractionated germ cells were subjected to RNA extraction and RT-qPCR. We observed that both CBP and p300 are highly expressed in spermatogonia and early spermatocytes but their expression levels decrease in post-meiotic cells. This result was confirmed by immunohistochemistry approach on testis paraffin section.

1-2- Role of the HATs CBP/p300 in the control of the histone hyperacetylation wave preceding histone removal

Cbp and *p300* genes deletion in late spermatogenic cells was achieved by crossing double knockout mice $cbp^{\text{flox/flox}}$ and $p300^{\text{flox/flox}}$ with mice expressing the Cre recombinase under the

control of the protamine promoter, expressed only in elongating spermatids, where H4 hyperacetylation takes place.

Despite an efficient deletion of the considered genes, due to the relative stability of CBP/p300 mRNAs and probably of the encoded proteins, we observed only a very limited decrease of the mRNAs and hence we could not determine if CBP and p300 actually play a role in the establishment of the hyperacetylation wave in elongating spermatids.

1-3- Role the the HATs CBP/p300 in controlling male germ cells gene expression programs

However we performed a comparative transcriptomic analysis of round and elongating spermatids, between mutated and wild type mice.

We hypothesized that the slight decrease in CBP and p300 mRNA might affect the genes whose transcription are the most sensitive to p300 activity. Our results show that indeed such a category of genes exists.

This analysis revealed that CBP and p300 control a late male gene germ cells expression program, which includes a subset of genes activated immediately prior to the global transcriptional shut-down. Interestingly the investigation of the functional annotation of these CBP and p300-dependent genes demonstrated that they are mainly involved in metabolic remodelling characterizing the late phase of spermatids maturation.

The results of this study are published in *Andrology* in 2013. For this work I performed all the experiments except the transcriptomic and statistical analyses, and the generation of CBP/p300 knock-out mice.

ORIGINAL ARTICLE

Correspondence:

Fayçal Boussouar, Saadi Khochbin and Sophie Rousseaux, INSERM, U823, Institut Albert Bonniot, Université Joseph Fourier - Grenoble 1, Grenoble, 38700, France. E-mails: faycal.boussouar@ujf-grenoble.fr, khochbin@ujf-grenoble.fr, sophie.rousseau@ujf-grenoble.fr

Keywords:

animal models, cell differentiation, chromatin, epigenetics, gene expression, histone modifications, immunohistochemistry, mouse, seminiferous tubules, spermatids

Received: 22-Oct-2013

Revised: 27-Dec-2013

Accepted: 30-Dec-2013

doi: 10.1111/j.2047-2927.2014.00184.x

A specific CBP/p300-dependent gene expression programme drives the metabolic remodelling in late stages of spermatogenesis

¹F. Boussouar, ¹A. Goudarzi, ¹T. Buchou, ¹H. Shiota, ¹S. Barral, ¹A. Debernardi, ²P. Guardiola, ³P. Brindle, ⁴G. Martinez, ⁴C. Arnoult, ¹S. Khochbin and ¹S. Rousseaux

¹INSERM, U823, Institut Albert Bonniot, Université Joseph Fourier - Grenoble 1, Grenoble, ²INSERM, U892, Centre de Recherche sur le Cancer Nantes Angers et UMR_S 892, Université d'Angers, Plateforme SNP, Transcriptome & Epigénomique, Centre Hospitalier Universitaire d'Angers, Angers, France, ³Department of Biochemistry, St Jude Children's Research Hospital, Memphis, TN, USA, and ⁴Laboratoire AGIM, CNRS FRE3405, Université Joseph Fourier, Grenoble, France

SUMMARY

Histone hyperacetylation is thought to drive the replacement of histones by transition proteins that occur in elongating spermatids (ELS) after a general shut down of transcription. The molecular machineries underlying this histone hyperacetylation remain still undefined. Here, we focused our attention on the role of Cbp and p300 in histone hyperacetylation and in the preceding late-gene transcriptional activity in ELS. A strategy was designed to partially deplete *Cbp* and *p300* in ELS. These cells progressed normally through spermiogenesis and showed normal histone hyperacetylation and removal. However, a genome-wide transcriptomic analysis, performed in the round spermatids (RS) and ELS, revealed the existence of a gene regulatory circuit encompassing genes presenting high expression levels in pre-meiotic cells, undergoing a repressed state in spermatocytes and early post-meiotic cells, but becoming reactivated in ELS, just prior to the global shutdown of transcription. Interestingly, this group of genes was over-represented within the genes affected by Cbp/p300 knock down and were all involved in metabolic remodelling. This study revealed the occurrence of a tightly regulated Cbp/p300-dependent gene expression programme that drives a specific metabolic state both in progenitor spermatogenic cells and in late transcriptionally active spermatids and confirmed a special link between Cbp/p300 and cell metabolism programming previously shown in somatic cells.

INTRODUCTION

After meiosis, haploid germ cells turn on a gene expression programme (Sassone-Corsi, 2002) that is essential for driving very unique events specific to spermatid differentiation (Rousseaux *et al.*, 2008), including a stepwise global genome reorganization (Gaucher *et al.*, 2010, 2012; Montellier *et al.*, 2013), a drastic metabolic adaptation (Boussouar & Benahmed, 2004; Alves *et al.*, 2013) and the generation of highly specialized structures such as the flagellum. Although the molecular mechanisms underlying most of these post-meiotic events have remained obscure (Rousseaux *et al.*, 2011), it is known that critical determinants rely on genes that are mostly activated in the haploid genome (Montellier *et al.*, 2012). In contrast, many genes that are active in somatic cells or in preceding spermatogenic stages become silent in post-meiotic cells (Namekawa

et al., 2006; Gaucher *et al.*, 2012). Therefore, a fine balance of specific gene activations and repressions is essential for the controlled differentiation of haploid cells. With this respect, the best understood mechanism of gene silencing in male germ cells is the meiotic sex chromosome transcriptional inactivation (MSCI). Indeed, in spermatocytes, at the time of homologous chromosome pairing, both sex chromosomes X and Y undergo a chromosome-wide transcriptional repression (Turner, 2007; Montellier *et al.*, 2012). This repressive state is inherited by the haploid cells after the completion of meiosis. However, as both X and Y bear important post-meiotic genes, the sex chromosomes undergo a partial gene transcriptional reactivation in early (round) spermatids, immediately after the meiotic division (Namekawa *et al.*, 2006; Mueller *et al.*, 2008). In addition, in these haploid cells, many autosomal genes, which are active in

spermatogonia and/or spermatocytes, undergo silencing before the general transcription arrest, through a yet unknown mechanism.

Recently, we demonstrated that histone lysine modifications by acetylation and by another acyl group, crotonylation, guide the expression of male germ cell genes at different stages of spermatogenesis (Tan *et al.*, 2011; Gaucher *et al.*, 2012). Histone acetylation gives a signal for an acetylation-guided action of Brdt, which, through its binding to genes bearing acetylated histones at their transcriptional start sites (TSS), successively induces meiotic and post-meiotic genes (Gaucher *et al.*, 2012). We also observed that histone crotonylation, in post-meiotic cells, is highly enriched in the TSS of genes that are active on autosomes as well as of the TSS of X-linked genes that escape the sex chromosome general silencing (Tan *et al.*, 2011). Therefore, histone crotonylation appears to mark a specific category of testis-specific genes that are active in post-meiotic cells, including a proportion of sex chromosome-linked genes, which escape silencing (Montellier *et al.*, 2012). At later stages, at the time of general transcriptional shut-down, the occurrence of a genome-wide histone hyperacetylation also prompts Brdt to induce a bromodomain-dependent histone-to-transition protein replacement (Gaucher *et al.*, 2012).

Despite the essential role of histone acetylation in driving stage-specific male germ cell gene expression, almost nothing is known on the enzymatic machineries that control this dynamic histone acetylation at critical loci. CBP/p300 are among the most critical acetyltransferases involved in a large number of biological processes (Goodman & Smolik, 2000; Bedford *et al.*, 2010). However, up to now, nearly nothing is known on the role of these enzymes in the control of male germ cell-specific gene expression.

Here, we generated conditional double KO for *Cbp* and *p300* with a post-meiotic deletion of the corresponding genes, aiming to investigate more specifically their role in haploid cells. This approach enabled us to discover a group of genes that are highly sensitive to the variation in *Cbp* and *p300* expression and revealed a post-meiotic genes regulatory circuit which specifically acts in late spermatogenic cells, just prior to the general transcriptional shutdown, and is predominantly controlling global metabolic remodelling of elongating spermatids (ELS).

MATERIALS AND METHODS

Mice

Cbp^{flox} and *p300^{flox}* mice were described previously and validated to generate null alleles after Cre-mediated recombination (Kang-Decker *et al.*, 2004; Kasper *et al.*, 2006). *Prm1-Cre* transgenic mice that express Cre specifically in post-meiotic germ cells were from AH. Peters (Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland). All experiments described in this study used mice on C57/BL6 and 129Sv mixed background. Ad hoc committees approved all experiments using mice in this study.

Genotyping

Semiquantitative PCR genotyping for *Cbp^{flox}* and *p300^{flox}* was described previously and was used routinely in this study to confirm gene recombination in post-meiotic germ cells (Kasper

et al. 2006). For *p300*, recombination of the LoxP sites flanking *p300* exon 9 in germ cells expressing Cre was determined by semiquantitative PCR using primers p4, p5, and p6, producing a 720-bp *p300^{flox}* band and a 1.1 kb *p300^{Δflox}* (Fig. 2B). For *Cbp*, the floxed allele gives a 1.3 kb band and the δ flox allele gives a 800 bp band.

Histology and immunohistochemistry

Testes were AFA-fixed overnight and paraffin embedded. Slides were either counterstained with eosin/haematoxylin or submitted to immunohistochemistry as described in Gaucher *et al.* (2012). Polyclonal anti-Cbp antibody A-22 (sc-369; Santa Cruz Biotechnology, Heidelberg, Germany) and polyclonal anti-p300 antibody N-15 (Santa Cruz Biotechnology) were used.

Sperm count

Spermatozoa were recovered from the cauda epididymis, suspended in an appropriated volume, and sperm cells were counted using a Malassez cell counting chamber.

Sperm motility assessment

Experiments were performed on a CASA CEROS v.12 (Hamilton ThorneBiosciences, Beverly, MA, USA) using Leja double chamber slide for standard count with 100 μ m depth. At least 150 cells by mouse of each genotype were analysed and recorded with the following parameters:

Image capture	Frames per sec. 60 Hz	No. of frames 100
Cell detection	Minimum contrast 30	Minimum cell size 4 pix
Progressive cells	Path velocity (VAP) 50 μ m/s	Straightness (STR) 50%
Defaults if <5 motile cells	Cell size 13 pix	Cell intensity 75
Slow cells	VAP cut-off 10 μ m/s	VSL cut-off 0 μ m/s
Static intensity gates	Min 0.10	Max 1.52
Static size gates	Min 0.13	Max 2.43
Static elongation gates	Min 5	Max 100

The measured parameters were as follows: VAP = speed according to average trajectory (μ m/sec), VCL = curvilinear velocity (μ m/sec), VSL = linear speed of progression (μ m/sec), ALH = amplitude of head movements (μ m), BCF = beat frequency of the flagellum (beats/sec). Using these data, two additional parameters were calculated: LIN = linearity (VSL/VCL) and STR = straightness (VSL/VAP). These parameters are illustrated in Fig. S2C.

Male germ cells fractionation by sedimentation on a BSA gradient

Male germ cells at different stages of spermatogenesis, spermatocytes, round spermatids (RS) and ELS were enriched by using sedimentation on a BSA gradient as described previously (Govin *et al.*, 2007; Tan *et al.*, 2011; Gaucher *et al.*, 2012; Montellier *et al.*, 2013) following a protocol originally described by Bellvé (1993). Briefly, for each experiment, four testes were removed from euthanized male mice, the albunine removed and male germ cells recovered by incubation of seminiferous tubules in collagenase (1 mg/mL) at 35 °C for 10 min and rapid pipetting. After washes, germ cells were suspended in a 0.5% BSA solution in DMEM/F12 at 4 °C, and allowed to sediment on a 2–4% BSA gradient for 75 min at 4 °C. Ten-mL cell fractions were then collected in 15-mL Falcon tubes, identified by examination of an aliquot under phase contrast microscopy, and pooled into three main

fractions, respectively, enriched in spermatocytes (>80% purity), RS (>95%) and elongating/condensing spermatids (>85%). For the transcriptomic analyses, we paid special attention to have strictly comparable fractions in the terms of proportions of cell types when comparing wild-type and double KO spermatogenic cells.

RNA preparation and transcriptomic and statistical analyses

RNA from testes taken at various time after birth or from fractionated cells was extracted using Qiagen (Courtaboeuf, Cedex, France) RNeasy mini kit. RNA was used for RT-qPCR measurement of specific gene expression or analysed on Illumina whole-genome chip as performed in Gaucher *et al.* (2012). The global gene expression level in RS and EIS wild types or double KO for *Cbp* and *p300* was compared as described in the above references.

Immunofluorescence, immunohistochemistry and microscopy

Protein visualization by immunofluorescence was carried out as previously described (Govin *et al.*, 2007).

RESULTS

Cbp and *p300* are highly expressed in spermatogonia and early spermatocytes, but show a decreased expression in post-meiotic cells

To have a first insight into the function of *Cbp* and *p300* in spermatogenic cells, we investigated the expression profile of both genes during mouse spermatogenesis. Developing testes were harvested at time points corresponding to defined stages of spermatogenesis: meiotic divisions start at 10 days post-partum (dpp) when pre-leptotene and leptotene spermatocytes first appear, spermatocytes undergo meiotic divisions and produce RS, which start to accumulate at 20 dpp, and later differentiate to finally generate mature spermatozoa.

A RT-qPCR using a *Cbp*-specific primer set revealed a high expression level of *Cbp* in spermatogonia as well as in pre-leptotene and leptotene spermatocytes. The amount of *Cbp* mRNA then continually declines to reach stable residual levels when post-meiotic cells appear (Fig. 1A: left panel). The same analysis performed for *p300* gene expression in the same samples showed a slightly different profile. While, as *Cbp*, *p300* expression raised at the onset of meiosis, its decrease started later, after the zygotene stage (Fig. 1A: right panel). In fractionated adult male germ cells, which included spermatocytes and EIS, a RT-qPCR detection confirmed the observation reported above and revealed a significant decrease in the expression of both genes in EIS compared with a pool of spermatocytes (Fig. 1B).

The use of specific antibodies against Cbp and p300 nicely confirmed the mRNA expression profile observed by RT-qPCR. Indeed, in both cases the most intense signals were observed in spermatogonia and in early spermatocytes. At later stages, the detection of both HATs showed a marked decrease, with signals which remain, however, significantly above the background (Fig. 1C). Finally, a more sensitive approach by immunofluorescence clearly confirmed the presence of significant levels of Cbp and p300 in post-meiotic cells: Cbp and p300 were both clearly detected in spermatids, including EIS, but were undetectable in condensing spermatids (Fig. 1D).

Generation of mice with a *Cbp*^{flox}; *p300*^{flox}; *Prm1-Cre* conditional knockout allele in post-meiotic germ cells

Despite a significant decrease in the amounts of both *Cbp* and *p300* mRNAs and proteins in post-meiotic cells, we hypothesized that these residual amounts of enzymes could sustain critical functions, mainly specific post-meiotic gene expression programmes.

To verify this hypothesis, we setup a strategy to specifically knockout both genes in post-meiotic cells. *Cbp*^{flox/flox} mice were crossed to *p300*^{flox/flox} mice and homozygous mice for both *Cbp*^{flox} and *p300*^{flox} alleles were selected and bred. These mice were then crossed with *Prm1-Cre* mice expressing the Cre recombinase under the control of the protamine 1 promoter solely in post-meiotic cells (Fig. 2A).

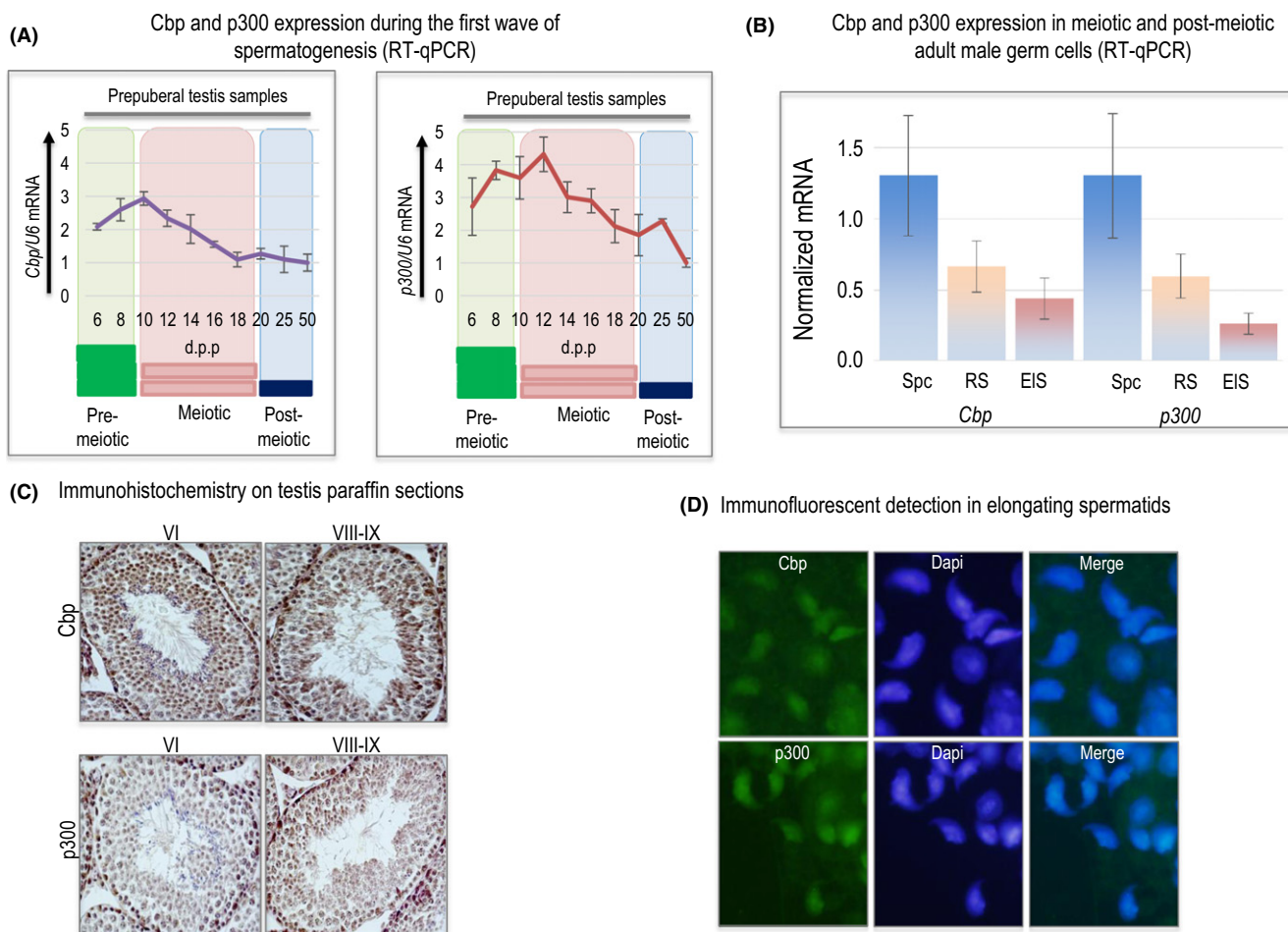
To test the efficiency of Cre recombination and the deletion of both *Cbp* and *p300* alleles, we fractionated spermatogenic cells from *Cbp*^{flox}; *p300*^{flox}; *Prm1-Cre* mice and the occurrence of the deletion was monitored in each cell fraction by semiquantitative PCR using an appropriate set of primers for both *Cbp*^{flox} and *p300*^{flox} alleles (Fig. 2B). Nearly, 90–95% deletion was achieved in EIS for the *Cbp*^{flox} allele which appeared to be complete in spermatozoa (left panel). The deletion of *p300*^{flox} was less efficient in elongating cells, but reached about 90% in spermatozoa (right panel).

We then analysed the state of *Cbp* and *p300* expression in fractionated post-meiotic cells from wild-type testes or the Cre-expressing testes described above. RNAs were prepared from fractionated cells highly enriched in EIS, where a clear deletion of both *Cbp* and *p300* alleles was observed. In agreement with the partial deletion of *Cbp* and *p300* genes, the RT-qPCR data showed that the levels of both *Cbp* and *p300* mRNAs were partially decreased, to approximately 70–60%, respectively, compared with their levels in wild-type male germ cells (Fig. 2C).

Male mice having a partial depletion of Cbp and p300 were fertile and did not show any defects in fertility. The analysis of testis histology from mice showing the down-regulation of *Cbp* and *p300* expression compared with control testes did not reveal any noticeable phenotype (Fig. S1). In addition, the fertility of male CBP/P300 double knockout mice was assessed by analysing spermatozoa harvested from cauda epididymes. As shown in Fig. S2A, B, the mutant mice showed normal sperm counts and testes weight. Sperm motility was also evaluated in a double-blind experiment on 150 sperm cells from wild type and mutant animals using CASA CEROS v.12 (Hamilton ThorneBiosciences; described in Methods) and did not show any difference (see Fig. S2C, Table S1, movie clips S1 and S2). Finally, three different mutant males were mated to control CBP flox/P300 flox females, and live pups were obtained (not shown), although the litter sizes were reduced because of the embryonic lethality of the double heterozygous *Cbp* and *p300* KO embryos.

To refine the analysis of a possible phenotype caused by the reduced expression of *Cbp* and *p300*, we have also monitored the wave of histone hyperacetylation that occurs in the late EIS at the time of histone-to-transition protein replacement. Figure S3 shows that, for all the histone acetylation marks tested, H3K9-ac, H3K14-ac, H4K5-ac and H4K8-ac, no particular changes can be observed in cell-expressing reduced levels of *Cbp* and *p300*.

Figure 1 *Cbp* and *p300* expression in the mouse testis. (A) Testes were obtained from mice at various ages after birth during the first wave of spermatogenesis, at 6 and 8 days post-partum (dpp), corresponding to stages with only pre-meiotic cells, then at 10, 12, 14, 16 and 18 dpp, when spermatocytes appear at different stages of meiotic 1 prophase, then at 20 and 25 dpp, when round and elongating post-meiotic cells successively appear, and finally in adult 50 dpp testes. *Cbp* (left panel) and *p300* (right panel) expressions were monitored by quantitative RT-PCR. Each point represents pooled RNA from two animals (mean $\pm$ standard deviation). (B) Expression of *Cbp* and *p300* analysed by RT-QPCR in meiotic (Spc = Spermatocytes) and post-meiotic round (RS) and elongating spermatids (EIS) obtained by fractionation of adult male germ cells (as described in the Methods section). Each bar represents the mean expression level (and standard deviation) of the gene obtained with five replicates from two independent experiments and mice. The expressions of *Actin* and *U6* were used as controls. (C) Immunolocalization of *Cbp* (top panels) and *p300* (bottom panels) using paraffin-embedded sections of adult testes. Roman numerals indicate the stages of the seminiferous epithelium. Two representative tubule sections are shown for each antibody. (D) Immunofluorescent detection of *Cbp* (top panels) and *p300* (bottom panels) in EIS (from control *Cbp*^{flx/flx}; *p300*^{flx/flx} mice). Seminiferous tubules were stained with DAPI to visualize nuclei and with antibodies specific for *Cbp* and *p300*.



Transcriptional control by *Cbp* and *p300* in EIS

Despite the partial decrease in *Cbp* and *p300* mRNA content in EIS, we decided to perform a genome-wide transcriptomic analysis of the corresponding cells to see if a specific gene expression programme, depending on these HATs, could be affected in these late post-meiotic cells, which are executing the final stage of the spermatogenesis gene expression programme. In addition, to identify the genes specifically involved in this final expression programme, we looked for genes differentially expressed between early (RS) and late (EIS) post-meiotic cells.

Spermatogenic cell suspensions from control or double partial *Cbp/p300* KO littermates (as described above) were fractionated and RNA was purified from fractions highly enriched for RS and EIS.

This whole-genome gene expression analysis shows that RS from control (wt) and from *Cbp*^{flx}; *p300*^{flx}; *Cre*-expressing mice

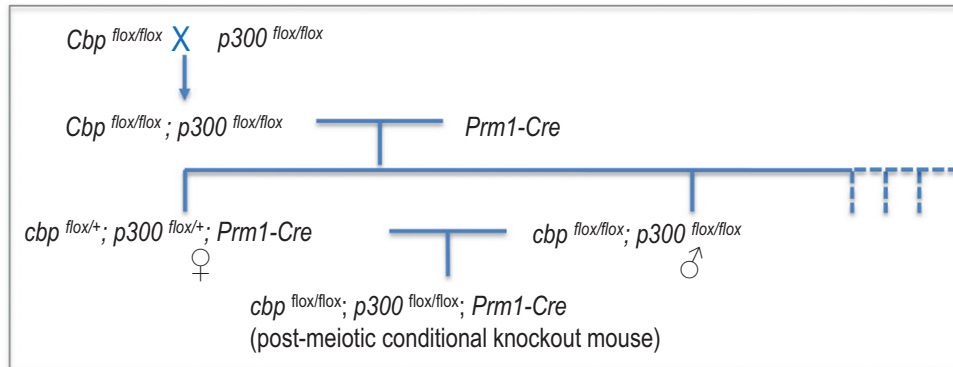
(mut) displays very similar gene expression patterns (Fig. 3A: RS Ctrl vs. dKO). In contrast, in EIS, a group of genes are significantly affected by the partial inactivation of CBP and p300 (Fig. 3A: EIS Ctrl vs. dKO), which suggests the involvement of the two HATs in the control of these genes and prompted us to know more about them.

The identification of a novel *Cbp*/*p300*-dependent gene regulatory circuit in EIS

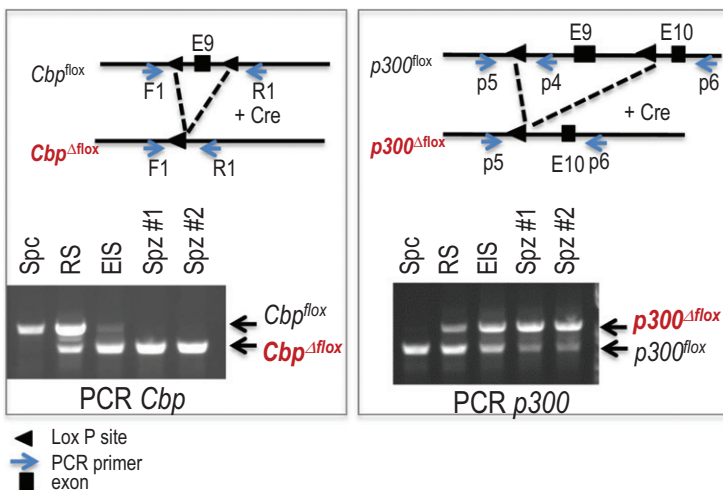
The partial depletion of *Cbp* and *p300* in EIS leads to a significant alteration in the expression of 114 genes (listed in Table S2): the expression of 64 genes is up-regulated and the expression of 50 other genes is down-regulated in mutant compared with wild-type EIS (Fig. 3A: red and green bars respectively). These genes should be those whose expression is most sensitive to variations in *Cbp* and *p300* transcripts and therefore we decided to

Figure 2 Generation of a conditional knockout mouse model to ablate *Cbp* and *p300* in post-meiotic germ cells leading to the partial depletion of *Cbp* and *p300* in late post-meiotic cells. (A) Strategy: *Cbp*^{flox/flox} were crossed to *p300*^{flox/flox} mice to generate double homozygote animals for *Cbp* and *p300* flox alleles that were in turn crossed to transgenic mice expressing Cre under the control of the protamine 1 promoter (*Prm1-Cre*). For all experiments of this study, the males were obtained after crossing *Cbp*^{flox/flox}; *p300*^{flox/flox} males to *Cbp*^{flox/flox}; *p300*^{flox/flox}; *Prm1-Cre* females. (B) Partial deletion of *Cbp* and *p300* in post-meiotic germ cells: Semiquantitative PCR on germ cell DNA showing detection of flox and Δ flox alleles of *Cbp* (left panel) and *p300* (right panel). DNA from spermatocytes (Spc), round (RS) and elongating spermatids (EIS), and spermatozoa (spz) was prepared and amplified. Spz #1 and #2 were obtained from separate *Cbp*^{flox/flox}; *p300*^{flox/flox}; *Prm1-Cre* mice showing high deletion efficiency in the final mature spermatozoa. (C) Expression of *Cbp* and *p300* in EIS from control (Ctrl) and double knockout (dKO) mice: Quantitative RT-PCR of control (Ctrl: *Cbp*^{flox/flox}; *p300*^{flox/flox}) and mutant (dKO: *Cbp*^{flox/flox}; *p300*^{flox/flox}; *Prm1-Cre*) EIS to detect *Cbp* (left panel) and *p300* (right panel) mRNA (mean $\pm$ standard deviation). Each bar represents the mean expression level (and standard deviation) of the gene obtained with four replicates from two independent experiments and mice. The expressions of *Actin* and *U6* were used as controls.

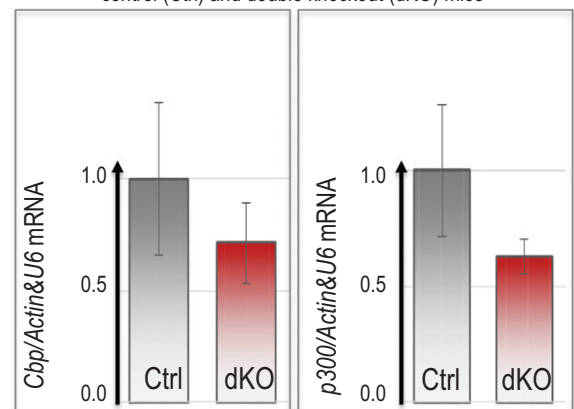
(A) Strategy for the generation of a conditional knockout mouse model to ablate *Cbp* and *p300* in post-meiotic germ cells



(B) Partial deletion of *Cbp* and *p300* in post-meiotic germ cells



(C) Expression of *Cbp* and *p300* in elongating spermatids from control (Ctrl) and double knockout (dKO) mice



investigate their specific features. Using publicly available transcriptomic data from staged spermatogenic cells (GSE21749 and GSE4193), we investigated the normal expression pattern of the genes whose expression in EIS was affected by *Cbp* and *p300* down-regulation.

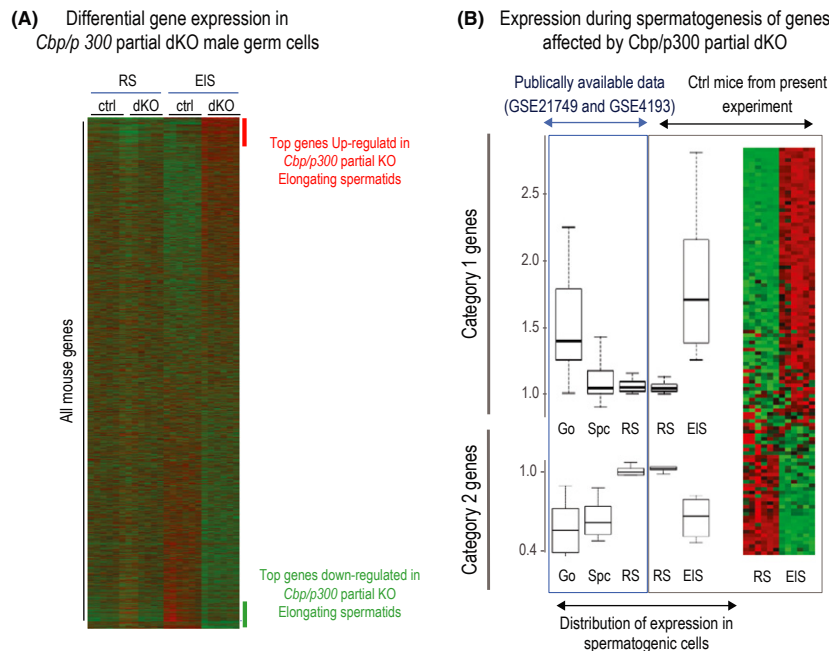
This analysis revealed that the partial depletion of *Cbp* and *p300* in EIS affects the late post-meiotic expression of a specific category of genes (defined as category 1), which are normally highly active in spermatogonia, and that become silent during meiosis in spermatocytes and in early RS (Fig. 3B: box plots from 'publicly available data'; category 1 genes).

In contrast, approximately a third of the genes (defined as category 2) whose expression was affected in EIS with partial depletion of *Cbp* and *p300* are normally highly expressed in early haploid cells (Fig. 3B: box plots from 'publicly available data'; category 2 genes) suggesting that these HATs also play a role in the induction of post-meiotic gene expression.

We checked the normal state of expression of these two categories of genes in our own transcriptome comparing control round and EIS (Fig. 3B: 'present experiment': box plots and heat map). Interestingly, approximately two-third of them also correspond to genes with an increased expression between round and EIS.

In summary, Fig. 3B shows that the genes affected by *Cbp* and *p300* in EIS belong to two main categories, which are specifically regulated during normal spermatogenesis. The first category encompasses genes that are normally highly expressed in spermatogonia, are down-regulated in meiotic cells and remain silent in early post-meiotic cells (RS), but finally show an increased expression in EIS (Fig. 3B: category 1 genes). Interestingly, the expression of most of the genes belonging to this category is even more stimulated if a decrease in the amount of *Cbp* and *p300* occurs (not shown), but some genes of this category are dependent on *Cbp/p300*

Figure 3 The genes whose expression is affected by *cbp/p300* partial deletion in elongating spermatids (EIS) are mostly highly expressed both in pre-meiotic cells and in EIS ("category 1" genes). (A) Heat map showing the differential gene expression in *Cbp/p300* partial KO (dKO) compared with control (Ctrl: *Cbp^{fllox/flox}; p300^{fllox/flox}*), round spermatids (RS) and EIS. The colour scale is from green to red corresponding, respectively, to low and high expression values. (B) Expression during spermatogenesis of genes affected by *Cbp/p300* partial dKO: box plots (left) showing the distribution of expression values of the genes whose expression in EIS is affected by *Cbp/p300* partial deletion. The expression values in spermatogenic cells from publically available data (GSE21749 and GSE4193: Go = spermatogonia, Spc = spermatocytes, RS = round spermatids) and from our experimental data (black rectangle: RS and EIS from control mice) are shown. To be comparable, the values in each experiment were normalized using the value of the first quartile in RS. Two categories of genes (1 and 2) were defined by their expression pattern in male germ cells (please see text).



to remain fully active. The second category of genes affected by the partial depletion of *Cbp/P300* (mostly corresponding to the genes down-regulated in mutant cells) is highly expressed in early post-meiotic cells (RS), but is normally down-regulated in EIS (Fig. 3B: category 2 genes).

Characterization of a last wave of gene expression before the general shutdown of transcription

By looking at the whole transcriptomic data between elongating and RS, we observed that, whereas most genes show a down-regulation or no significant variation in their expression levels between round and EIS, a relatively small group of 1546 genes (representing approximately 5% of the whole transcriptome) displayed a significant and marked increase in their expression between round and EIS, corresponding to a last post-meiotic wave of gene activations during spermatogenesis (Fig. 4A: grey rectangle). Considering the state of expression of these genes during the other spermatogenic stages, using the available data on staged spermatogenic cells, we found that 435 of these genes (approximately a third) are highly expressed in spermatogonia (Fig. 4B), but their expression decreases in meiotic (spermatocytes) and early post-meiotic cells (RS) until the start of haploid cell elongation when they will become reactivated (as shown Fig. 4A).

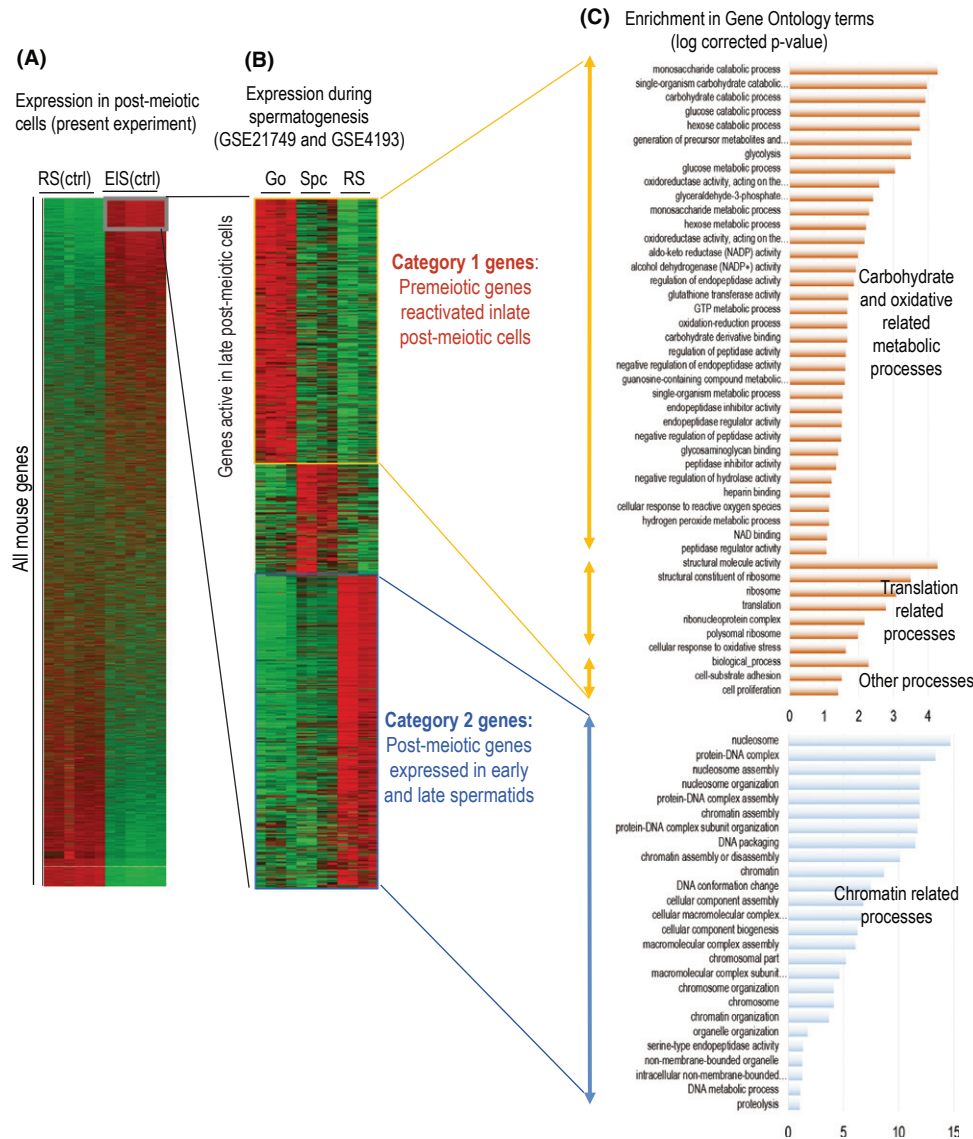
Interestingly, this category of genes appears to be the most sensitive to the decrease in *Cbp* and *p300* expression. Indeed, as shown above, the majority of the genes whose expression is affected by *Cbp* and *p300* down-regulation belongs to this category (Fig. 3B).

A specific *Cbp/p300*-dependent gene expression programme drives metabolic reprogramming in EIS

To better characterize genes that are activated during late spermatogenesis in EIS, we first carried out an analysis of the potential functions enriched in the genes whose expression is increased between control elongating and RS using Gene Sets Enrichment Analysis (GSEA) tool and the collections of gene sets made available by the Broad Institute, as well as a gene ontology (GO) terms enrichment analysis. From this analysis we found that most of the genes up-regulated in late transcriptionally active spermatids control processes involving translation, chromatin remodelling and nuclear reorganizations. We then paid specific attention to the subset of genes formed by the 435 genes, encompassing the *Cbp/p300* responsive genes, which are highly expressed in spermatogonia, repressed in meiotic and early post-meiotic cells and later reactivated in EIS. Interestingly, the GO terms enrichment analysis on this subgroup of genes showed that most of the highly enriched terms are related to carbohydrate metabolism and oxidoreductase activities (Fig. 4C: top panel), in contrast to the GO terms enriched in the other post-meiotically active genes, which mostly involve chromatin-related activities (Fig. 4C: lower panel).

Overall, this transcriptomic analysis revealed a late wave of gene activations in EIS, largely represented by genes encoding factors driving the dramatic genome reorganizations that occur in the subsequent stages. However, this late burst of gene activations also concerns another category of genes, which are reactivated in EIS after a long-lasting meiotic and early post-meiotic

Figure 4 The 'category 1' genes, which are highly expressed in pre-meiotic cells and in elongating spermatids (EIS), are mainly involved in carbohydrate and oxidative-related metabolic processes, whereas the 'category 2' genes expressed in post-meiotic round and EIS are involved in chromatin-related processes. (A) Heat map comparing the expression of all mouse genes between round spermatids (RS) and EIS from control mice (Ctrl) in our experiment. (B) Heat map showing the expression pattern, during spermatogenesis (from publically available data GSE21749 and GSE4193), of genes whose expression level is higher in EIS than in RS. Two major expression patterns are observed for these genes: (i) the first pattern correspond to genes highly expressed in spermatogonia, with low expression in spermatocytes (Spc) and RS, and reactivated in EIS. (ii) The second pattern correspond to post-meiotic genes with an increased expression in round and EIS. The 'category 1' genes mostly affected by the partial deletion of *Cbp/p300* belong to the first group. (C) Gene ontology (GO) terms enrichment of the two groups of genes defined in 4B; the bars represent the $-\log$ (corrected p -value), the p -value was calculated using the hypergeometric probability. The GO terms corresponding to the 'category 1' genes (including the genes affected by the partial deletion of *Cbp/p300*) are in the yellow rectangle. The GO terms corresponding to the 'category 2' genes are in the blue rectangle.



down-regulation, and are driving a 'metabolic remodelling'. This latter gene group includes a high proportion of the genes affected by our *Cbp/p300* post-meiotic KO, suggesting that *Cbp* and *p300* are master players in the regulation of their expression.

DISCUSSION

Our approach allowing a partial down-regulation of *Cbp* and *p300*, at a specific time just prior to the general shutdown of transcription at the transition from round to EIS, revealed the existence of stage-specific tightly regulated distinct gene expression programmes. Indeed, our experimental data in control mice

revealed that approximately 5% of the genes are activated between round and EIS, meaning that their RNA levels are still high immediately prior to the general shut down of gene expression, known to occur during the transition from elongating to condensing spermatids. Among these genes we identified two main categories of genes as follows. One category of genes, called here category 2, are silent in pre-meiotic and meiotic cells, but become expressed in early post-meiotic cells, RS, and their RNA levels continue to increase in EIS. A GO enrichment analysis found that a large proportion of these genes encode essential elements in the final programming and the reorganization of the male genome, by directly acting on chromatin and the genome.

Taking into account the fact that dramatic transitions occur right after this wave of gene expression, it is indeed expected that determinant driving elements should be expressed during the preceding period. However, this analysis also revealed another category of genes, called here category 1, which are highly active in progenitor cells, spermatogonia, but undergo silencing during meiotic divisions and remain silent in early post-meiotic stages in RS. Interestingly, our transcriptomic analysis comparing round and EIS showed that these genes undergo a late reactivation, in EIS, during a last wave of gene expression just prior to the general shutdown of transcription. Interestingly, the results of our GO analysis suggest that the genes of category 1 are nearly all involved in driving the cell metabolism. This is a remarkable observation in that one of the last command of the male genome over the spermatogenic differentiation encompasses parallel actions on both the genome itself and the metabolism. Interestingly, most of the genes affected by a late depletion in *Cbp/p300* expression are among those normally activated between round and EIS. The *Cbp/p300* partial post-meiotic depletion seems to be predominantly affecting category 1 genes, involved in the metabolic remodelling. As some category 2 genes, involved in chromatin reorganization, are also affected by the *Cbp/p300* partial depletion in EIS, it is likely that these HATs also partly and indirectly control the histone-to-protamine exchange. However, as the depletion here was achieved partially in late but transcriptionally active post-meiotic cells (elongating stage) and did not affect the following wave of histone acetylation occurring at the time of histone removal and exchange, the data presented here do not support any direct effect of the decreased expression levels of Cbp and p300 on histone removal and exchange by transition proteins and protamines.

A potential explanation of the rather mild effect of the post-meiotic deletion of *Cbp/p300* is the partial nature of *cbp/p300* knock down leading to a significant but slight alteration of the gene expression patterns, therefore only partly changing the expression of the subgroup of genes which are highly sensitive to the down-regulation of *cbp/p300*. In addition, the mild changes in gene expression reported here after a reduction almost by half of *Cbp* and *p300* mRNAs could also be explained by the long half-life of the encoded proteins that could result in even a smaller decrease in the proteins' amount.

The activity of the remaining *Cbp/p300* could explain that the deletion does not result in a more striking phenotype. However, the consistency of our results for the affected genes (whose levels of expression consistently varied out of the normal distribution range – above or below the mean expression level in the six corresponding control testes ± 2 standard deviations – in at least four of six of the mutated EIS) suggests that this work reveals the most sensitive among a larger number of genes controlled by *cbp/p300*.

In the recent years, our understanding of epigenetically controlled male germ cell-specific gene expression has been significantly increased (Tan *et al.*, 2011; Gaucher *et al.*, 2012). These investigations revealed a particular role for histone acetylation in directing Brdt and its associated P-TEFb to particular sets of meiotic and post-meiotic genes (Gaucher *et al.*, 2012). However, almost nothing is known on the role of the enzymatic machinery that controls a programmed establishment of the acetylation marks involving specific categories of gene expression programmes. As Cbp and p300 are among the most prominent and

multifunctional cellular HATs, we became interested in their role in the control of gene expression in male germ cells. We found that, following the commitment of spermatogonia in the meiotic divisions, a dramatic down-regulation of *Cbp* and *p300* gene expression occurs. One can speculate that the decrease in Cbp/p300 content corresponds to a shift in the activity of these HATs from a general transcriptional regulator function into a more specific role in driving particular gene expression programmes. Our approach based on the down-regulation of *Cbp* and *p300*, specifically in EIS, strongly supports this hypothesis. Indeed, our data show that the partial down-regulation of both genes in EIS, instead of having a general consequence on gene expression, very specifically affects a particular category of genes. Our additional analysis showed that these genes predominantly belong to a defined category that control spermatid-specific metabolic process.

This observation therefore supports the idea that, in post-meiotic cells, Cbp and p300 become particularly involved in the control of expression of functionally linked gene circuits that are metabolic genes. As genes of category 1, those mostly affected by the down-regulation of *Cbp* and *p300*, are mainly active in spermatogonia and in EIS, our data suggest that the Cbp and p300 are involved in the regulation of genes linking these two very different cell types with regards to cell metabolism. This could be the capacity of spermatogonia and mature spermatozoa to use glucose, a feature that is not shared with other spermatogenic cell types. Spermatocytes and spermatids are located inside the blood barrier, but spermatogonia are outside the barrier. Spermatogonia, therefore, may utilize the glucose in blood as an energy source. Because a large number of germ cells are developing into spermatozoa in the seminiferous tubules, they must be supplied efficiently with other energy sources. The ATP levels in spermatozoa are, however, maintained by several glycolytic and non-glycolytic substrates as both glycolysis and mitochondrial respiration are active in mammalian spermatozoa (Boussouar & Benahmed, 2004; Alves *et al.*, 2013). It is noteworthy that, among the category 1 genes that are affected by the Cbp and p300 partial depletion, glucose and hexose metabolism pathways as well as monosaccharides have the highest scores, strongly suggesting a tight regulation of the energy metabolism of spermatozoa by Cbp and p300. Interestingly, Cbp and p300 were also shown to help maintain energy homeostasis in somatic cells settings (Bedford *et al.*, 2011). Altogether, these data not only revealed the occurrence of stage-specific gene expression regulatory circuits that control the expression of functionally related genes but also showed a role of the Cbp and p300 in this process.

ACKNOWLEDGEMENTS

Mice were bred in the High Technology Animal Facility (PHTA) of Grenoble University, with the precious help of Rachel Balouza and Marilyn Cossin. Paraffin embedding was performed on the histology platform of Grenoble Medical School. This project is supported by ANR EpiSperm2 grant and its cancer oriented developments by INCa. AG and HS are both recipients of a Marie Curie Initial Training Network funded by European Commission (FP7-PEOPLE-2011-ITN, PITN-GA-289880). FB was a recipient of "contrat jeune chercheur" INSERM, from 2006 to 2011.

REFERENCES

- Alves MG, Martins AD, Rato L, Moreira PI, Socorro S & Oliveira PF. (2013) Molecular mechanisms beyond glucose transport in diabetes-related male infertility. *Biochim Biophys Acta* 1832, 626–635.
- Bedford DC, Kasper LH, Fukuyama T & Brindle PK. (2010) Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* 5, 9–15.
- Bedford DC, Kasper LH, Wang R, Chang Y, Green DR & Brindle PK. (2011) Disrupting the CH1 domain structure in the acetyltransferases CBP and p300 results in lean mice with increased metabolic control. *Cell Metab* 14, 219–230.
- Bellvé AR. (1993) Purification, culture, and fractionation of spermatogenic cells. *Methods Enzymol* 225, 84–113.
- Boussouar F & Benahmed M. (2004) Lactate and energy metabolism in male germ cells. *Trends Endocrinol Metab* 15, 345–350.
- Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S & Khochbin S. (2010) From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 277, 599–604.
- Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, Bertrand S, Héry P, Jounier S, Depaux A, Vitte AL, Guardiola P, Pernet K, Debernardi A, Lopez F, Holota H, Imbert J, Wolgemuth DJ, Gerard M, Rousseaux S & Khochbin S. (2012) Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* 31, 3809–3820.
- Goodman RH & Smolik S. (2000) CBP/p300 in cell growth, transformation, and development. *Genes Dev* 14, 1553–1577.
- Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thevenon J, Catena R, Davidson I, Garin J, Khochbin S & Caron C. (2007) Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J Cell Biol* 176, 283–294.
- Kang-Decker N, Tong C, Boussouar F, Baker DJ, Xu W, Leontovich AA, Taylor WR, Brindle PK & van Deursen JM. (2004) Loss of CBP causes T cell lymphomagenesis in synergy with p27Kip1 insufficiency. *Cancer Cell* 5, 177–189.
- Kasper LH, Fukuyama T, Biesen MA, Boussouar F, Tong C, de Pauw A, Murray PJ, van Deursen JM & Brindle PK. (2006) Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Molecular and Cellular Biology* 26, 789–809.
- Montellier E, Rousseaux S, Zhao Y & Khochbin S. (2012) Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *BioEssays* 34, 187–193.
- Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, Shiota H, Debernardi A, Héry P, Curtet S, Jamshidikia M, Barral S, Holota H, Bergon A, Lopez F, Guardiola P, Pernet K, Imbert J, Petosa C, Tan M, Zhao Y, Gérard M & Khochbin S. (2013) Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev* 27, 1680–1692.
- Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC & Turner JM. (2008) The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat Genet* 40, 794–799.
- Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD & Lee JT. (2006) Postmeiotic sex chromatin in the male germline of mice. *Curr Biol* 16, 660–667.
- Rousseaux S, Reynoird N, Escoffier E, Thevenon J, Caron C & Khochbin S. (2008) Epigenetic reprogramming of the male genome during gametogenesis and in the zygote. *Reprod Biomed Online* 16, 492–503.
- Rousseaux S, Boussouar F, Gaucher J, Reynoird N, Montellier E, Curtet S, Vitte AL & Khochbin S. (2011) Molecular models for post-meiotic male genome reprogramming. *Syst Biol Reprod Med* 57, 50–53.
- Sassone-Corsi P. (2002) Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science* 296, 2176–2178.
- Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B & Zhao Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016–1028.
- Turner JM. (2007) Meiotic sex chromosome inactivation. *Development* 134, 1823–1831.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Cbp and p300 partial depletion does not affect mouse spermatogenesis. Haematoxylin and eosin staining of paraffin-embedded testis sections from adult 3-month-old mice. A total of six testis sections of each genotype were analysed after HE staining corresponding to approximately 100 staged tubules. All stages were found in similar proportions in both animals.

Figure S2. Mean testis weight (A), sperm counts in cauda epididymis (B) and motility parameters (C) in control (Ctrl) and in mice with *Cbp/p300* post-meiotic partial deletion (*Cbp/p300* dKO).

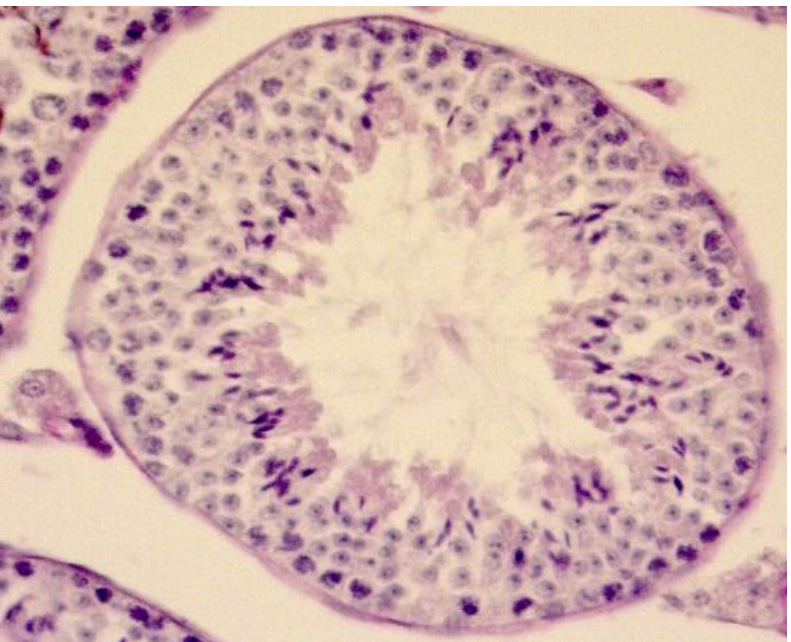
Figure S3. *Cbp* and *p300* partial depletion does not affect post-translational acetylation of histones. Seminiferous tubule sections from control (*Cbp^{fllox/fllox}; p300^{fllox/fllox}*) or mutant (*Cbp^{fllox/fllox}; p300^{fllox/fllox}; Prm1-Cre*) animals were immunostained with the indicated antibodies. Note that no obvious differences could be observed between the two genotypes.

Table S1. Motility parameters.

Movie clip S1 (Mutant.avi): Motility of epididymal sperm from a mutant mouse.

Movie clip S2 (Ctrl.avi): Motility of epididymal sperm from a control mouse.

Cbp^{flox/flox}; p300^{flox/flox}



Cbp^{Δflox}; p300^{Δflox}

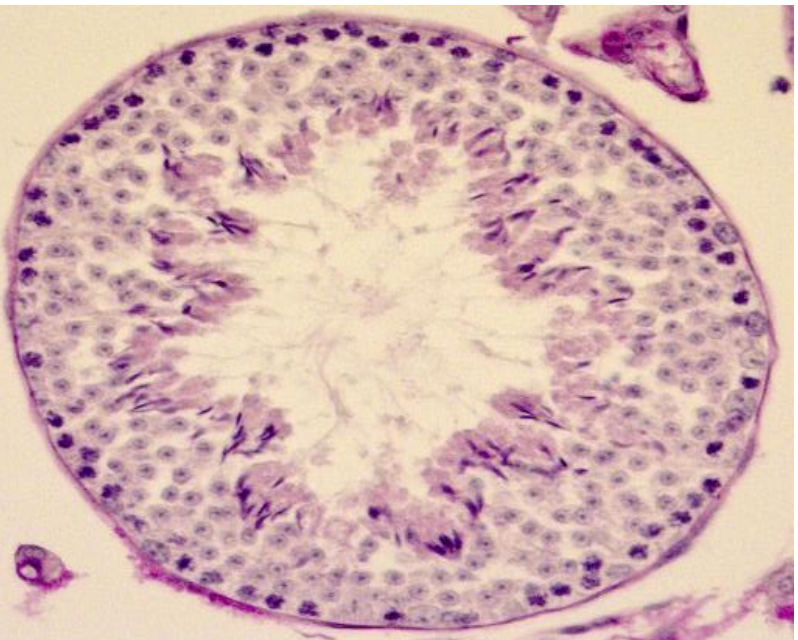


Figure S1. *Cbp* and *p300* partial depletion do not affect mouse spermatogenesis. Haematoxylin and eosin staining of paraffin-embedded testis sections from adult 3 month-old mice. A total of 6 testis sections of each genotype were analysed after HE staining corresponding to approximately 100 staged tubules. All stages were found in similar proportions in both animals.

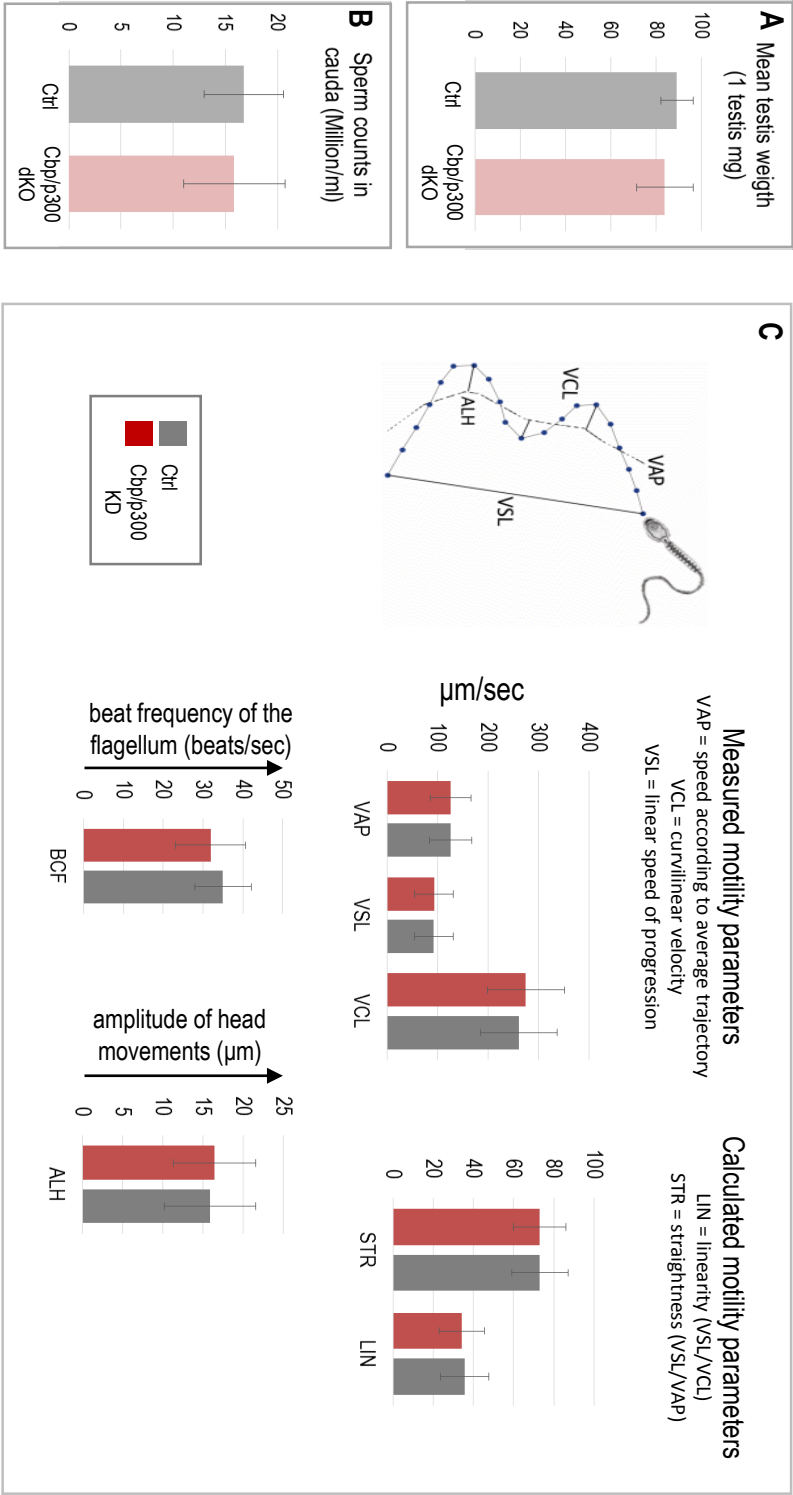


Figure S2: Mean testis weight (A), Sperm counts in cauda epididymis (B) and motility parameters (C) in control (Ctrl) and in mice with *p300/cbp* post-meiotic partial deletion (*p300/cbp* dKO)

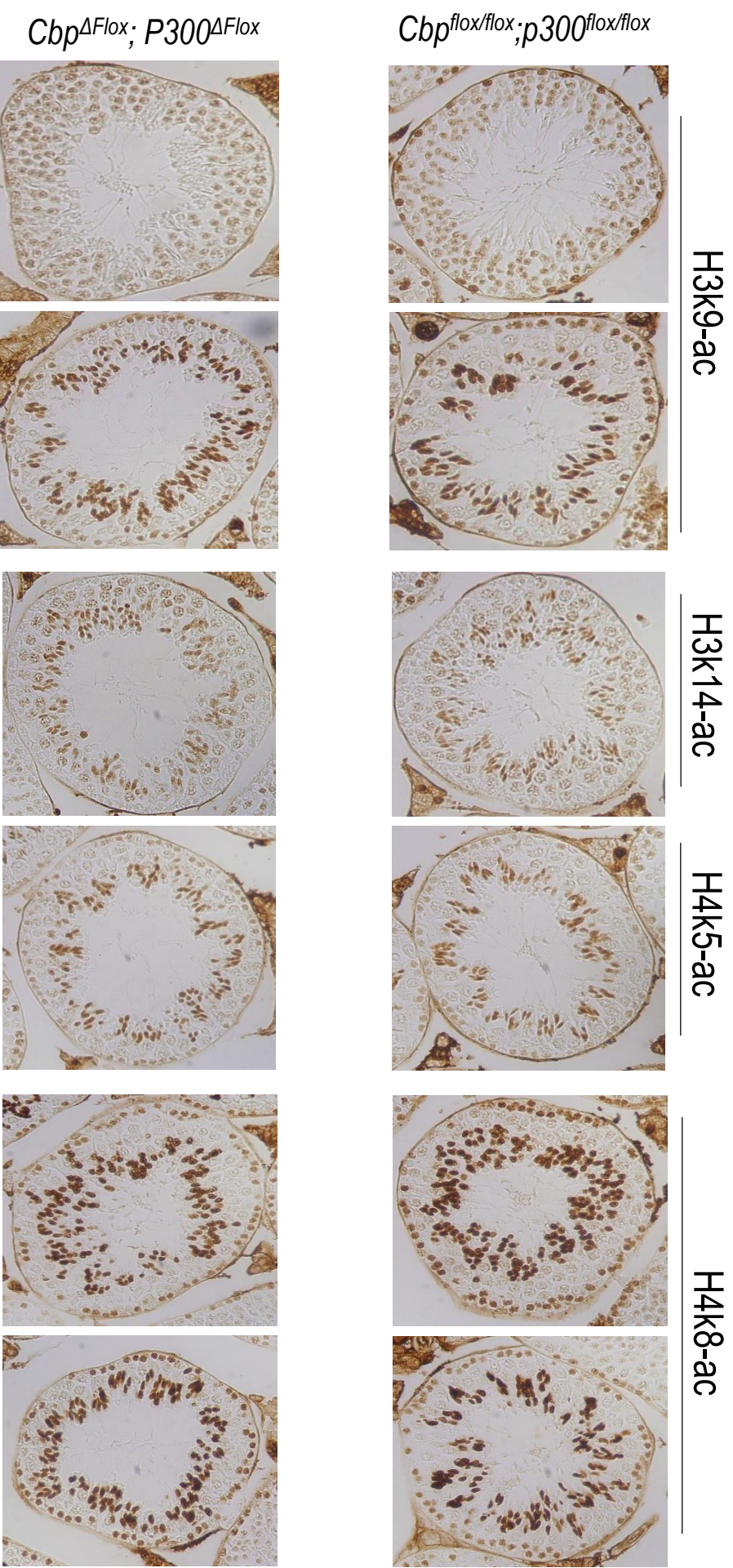


Figure S3. *Cbp* and *p300* partial depletion does not affect post-translational acetylation of histones. Seminiferous tubule sections from control (*Cbp^{flox/flox}; p300^{flox/flox}*) or mutant (*Cbp^{flox/flox}; p300^{flox/flox}; Prm1-Cre*) animals were immunostained with the indicated antibodies. Note that no obvious differences could be observed between the two genotypes.

2- Investigation of butyrylation in parallel with acetylation on lysines 5 and 8 of histone H4 during spermatogenesis.

Although our team has already shown the involvement of lysine acetylation, crotonylation and 2-hydroxyisobutyrylation in controlling the male germ cells gene expression program, more histone modifications could play important roles during mammalian spermatogenesis. The main part of my thesis focused on the investigation of the potential role of histone butyrylation in male germ cells gene expression program and in the guidance of Brdt's function.

We chose to use butyrylation as a representative of the non-acetyl histone acylations in order to deepen our understanding of their functions. Butyrylation was chosen since published data demonstrated that bromodomains start to discriminate acyl groups when their length exceeds three carbons ([Flynn et al., 2015](#)). Therefore, butyrylation has the potential to infer a different function compared to acetylation, while propionylation would have the same function as acetylation.

We focused our attention on the functional impact of butyrylation on histone H4 lysines 5 and 8, together with acetylation of the same positions, which are critical for Brdt function ([Moriniere et al., 2009](#)).

Concerning stage specific gene expression during spermatogenesis, one of the important questions is whether the different profiles of gene expression are associated with specific epigenetic changes, more specifically in acetylation and butyrylation. In this study we used several approaches including Immunofluorescence assays and ChIP-Seq with position specific antibodies against H4K5ac, H4K5bu, H4K8ac and H4K8bu to address the above-mentioned questions.

Main results obtained

2-1- Genomic distribution of histone H4 K5K8 butyrylation and acetylation

We found that histone butyrylation is an evolutionary conserved epigenetic mark. Genome-wide distribution of butyrylation in parallel with acetylation of H4 K5K8 sites was investigated using different approaches including immunofluorescence and ChIP-Seq assays.

Our ChIP-Seq study on spermatocytes and round spermatids revealed the co-occurrence of the four PTMs at gene TSS showing high expression levels. Our immunofluorescence assays showed a different labelling of H4K5 and H4K8 by acetylation and butyrylation marks in elongating spermatids. In these cells, butyrylated histones appeared to be homogenously distributed while acetylation was restricted to the sub-acrosomal region and disappeared earlier than butyrylation.

2-2- Histone butyrylation activates transcription but inhibits the binding of Brdt

p300-catalyzed lysine butyrylation on a reconstituted chromatin template *in vitro*, confirmed the positive effect of butyrylation on transcription. Since we previously showed the important role of Brdt in driving specific spermatogenic gene expression programs ([Gaucher et al., 2012](#)), these data suggested that Brdt could also bind butyrylated lysines. To assess the binding ability of Brdt to bind butyrylated histones, we compared our Brdt ChIP-seq data with our PTMs ChIP-seq data and found that the Brdt bound TSS regions were also covered by the four PTMs.

To precisely and definitely test the ability of Brdt to bind butyrylated histones, we set up a pull-down assay. In contrast to the data obtained from our ChIP-seq where Brdt was found at gene TSSs containing the four PTMs including H4K5bu, the pull-down assay demonstrated that butyrylation at H4K5 strongly interferes with Brdt binding. This apparent contradiction suggests the occurrence of a dynamic exchange of acetylation and butyrylation at active chromatin regions. In addition, our structural knowledge of Brdt-BD1 in complex with H4K5acK8ac ([Moriniere et al., 2009](#)), explains why butyrylation on lysine 5 of histone H4 cannot be tolerated.

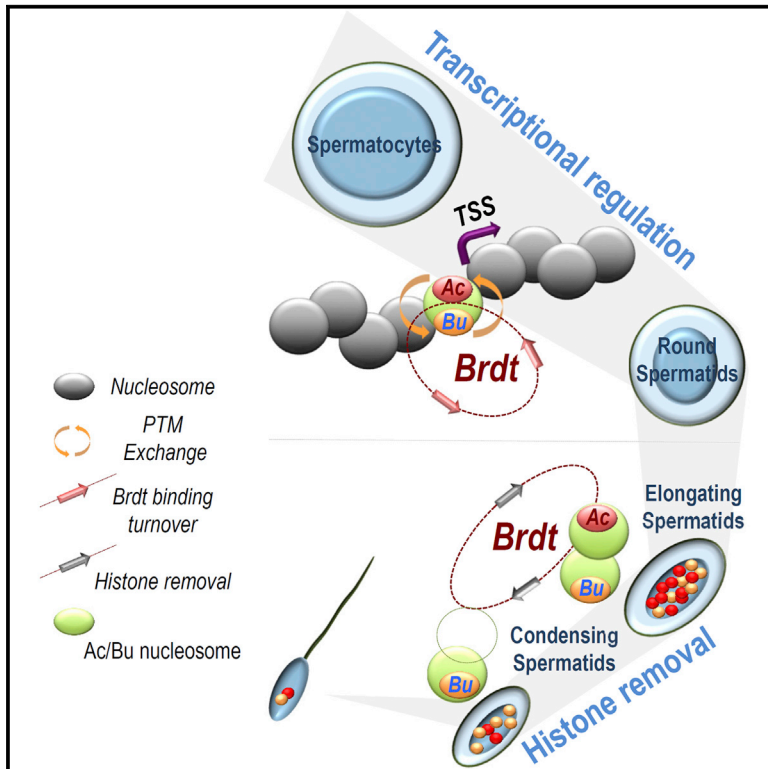
These studies opened the way for the development of important new concepts. Indeed most histone PTMs are dynamics and have relatively short half-lives. Our finding indicate that alternating PTMs on a specific position may create a dynamic state that is necessary for important processes such as chromatin remodeling and transcription.

This work also indicates that a change in the amount of acetyl-CoA or other acyl-CoAs could have important functional consequences by stabilizing a given chromatin state and by slowing down its dynamic nature.

These results have been published in Molecular Cell in 2016. For my contribution in this work, please see author contributions in the paper.

Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters

Graphical Abstract



Authors

Afsaneh Goudarzi, Di Zhang, He Huang, ..., Robert G. Roeder, Yingming Zhao, Saadi Khochbin

Correspondence

yzhao2@bsd.uchicago.edu (Y.Z.), saadi.khochbin@univ-grenoble-alpes.fr (S.K.)

In Brief

Histone butyrylation stimulates gene transcription while competing with acetylation at H4K5 to control Brdt bromodomain binding. Differential chromatin labeling with interchangeable H4 acylations is an important epigenetic regulatory mechanism controlling gene expression and chromatin reorganization.

Highlights

- Active gene TSSs are marked by competing H4 K5K8 acetylation and butyrylation
- Histone butyrylation directly stimulates transcription
- H4K5 butyrylation prevents binding of the testis specific gene expression-driver Brdt
- H4K5K8 butyrylation is associated with delayed histone removal in spermatogenic cells

Accession Numbers

GSE77277

GSE39910



Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters

Afsaneh Goudarzi,^{1,8} Di Zhang,^{2,8} He Huang,² Sophie Barral,¹ Oh Kwang Kwon,² Shankang Qi,² Zhanyun Tang,³ Thierry Buchou,¹ Anne-Laure Vitte,¹ Tieming He,⁴ Zhongyi Cheng,⁴ Emilie Montellier,¹ Jonathan Gaucher,^{1,5} Sandrine Curtet,¹ Alexandra Debernardi,¹ Guillaume Charbonnier,⁶ Denis Puthier,⁶ Carlo Petosa,⁷ Daniel Panne,⁵ Sophie Rousseaux,¹ Robert G. Roeder,³ Yingming Zhao,^{2,9,*} and Saadi Khochbin^{1,9,*}

¹CNRS UMR 5309, INSERM, U1209, Université Grenoble Alpes, Institut Albert Bonniot, 38700 Grenoble, France

²Ben May Department of Cancer Research, The University of Chicago, Chicago, IL 60637, USA

³Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10065, USA

⁴Jingjie PTM Biolab (Hangzhou) Co., Ltd., Hangzhou 310018, China

⁵EMBL Grenoble, BP 181, 71 Avenue des Martyrs, 38042 Grenoble Cedex 9, France

⁶TAGC, UMR, S 1090 INSERM Aix-Marseille Université, U928 Parc Scientifique de Luminy case 928 163, Avenue de Luminy, 13288 Marseille Cedex 9, France

⁷Université Grenoble Alpes/CEA/CNRS, Institut de Biologie Structurale, 38027 Grenoble, France

⁸Co-first author

⁹Co-senior author

*Correspondence: yzhao2@bsd.uchicago.edu (Y.Z.), saadi.khochbin@univ-grenoble-alpes.fr (S.K.)

<http://dx.doi.org/10.1016/j.molcel.2016.03.014>

SUMMARY

Recently discovered histone lysine acylation marks increase the functional diversity of nucleosomes well beyond acetylation. Here, we focus on histone butyrylation in the context of sperm cell differentiation. Specifically, we investigate the butyrylation of histone H4 lysine 5 and 8 at gene promoters where acetylation guides the binding of Brdt, a bromodomain-containing protein, thereby mediating stage-specific gene expression programs and post-meiotic chromatin reorganization. Genome-wide mapping data show that highly active Brdt-bound gene promoters systematically harbor competing histone acetylation and butyrylation marks at H4 K5 and H4 K8. Despite acting as a direct stimulator of transcription, histone butyrylation competes with acetylation, especially at H4 K5, to prevent Brdt binding. Additionally, H4 K5K8 butyrylation also marks retarded histone removal during late spermatogenesis. Hence, alternating H4 acetylation and butyrylation, while sustaining direct gene activation and dynamic bromodomain binding, could impact the final male epigenome features.

INTRODUCTION

Besides lysine acetylation, we recently identified a variety of short-chain lysine acylations in core histones, including lysine propionylation, butyrylation, 2-hydroxyisobutyrylation, crotonylation, malonylation, succinylation, and glutarylation (Chen

et al., 2007; Dai et al., 2014; Tan et al., 2011, 2014; Xie et al., 2012). Emerging data suggest that these new histone lysine acylations may have unique functions that depend not only on cell metabolism, but also on their ability to be deposited or removed by specific enzymes (Dai et al., 2014; Montellier et al., 2012; Rousseaux and Khochbin, 2015; Sabari et al., 2015; Sin et al., 2012; Tan et al., 2011). Nevertheless, the functional impact of differential histone acylation on chromatin recognition by specific factors has remained unexplored.

This study aims to understand the functional consequences of differential histone acylation. In particular, we decided to investigate histone butyrylation, because, in contrast to the acetyl (2-carbon) and propionyl (3-carbon) groups, the butyryl (4-carbon) group restricts the binding of bromodomains (Flynn et al., 2015). More specifically, we focused our attention on histone H4 at K5 and K8, whose acetylation is required to bind the first bromodomain of Brdt, a testis-specific member of the BET protein family (Morinière et al., 2009). Our previous work showed that Brdt stimulates the transcription of certain spermatogenic-specific genes by recruiting the P-TEFb complex and by directly binding to their transcriptional start sites (TSSs). Additionally, during late spermatogenesis, Brdt's first bromodomain is necessary for the replacement of histones by non-histone sperm-specific transition proteins (TPs) and protamines (Prms) (Gaucher et al., 2012). Given the critical role of H4K5 and H4K8 acetylation in Brdt-driven activities, we hypothesized that other mutually exclusive histone marks at these two residues might have key regulatory roles in sperm cell genome programming.

Here, we identify major histone lysine butyrylation sites in cells from different species, including mouse spermatogenic cells. Using spermatogenesis as an integrated biological model system, in addition to in vitro experiments and targeted proteomic approaches, we demonstrate new characteristics of active gene TSSs. Our data indicate that interchangeable acetylation

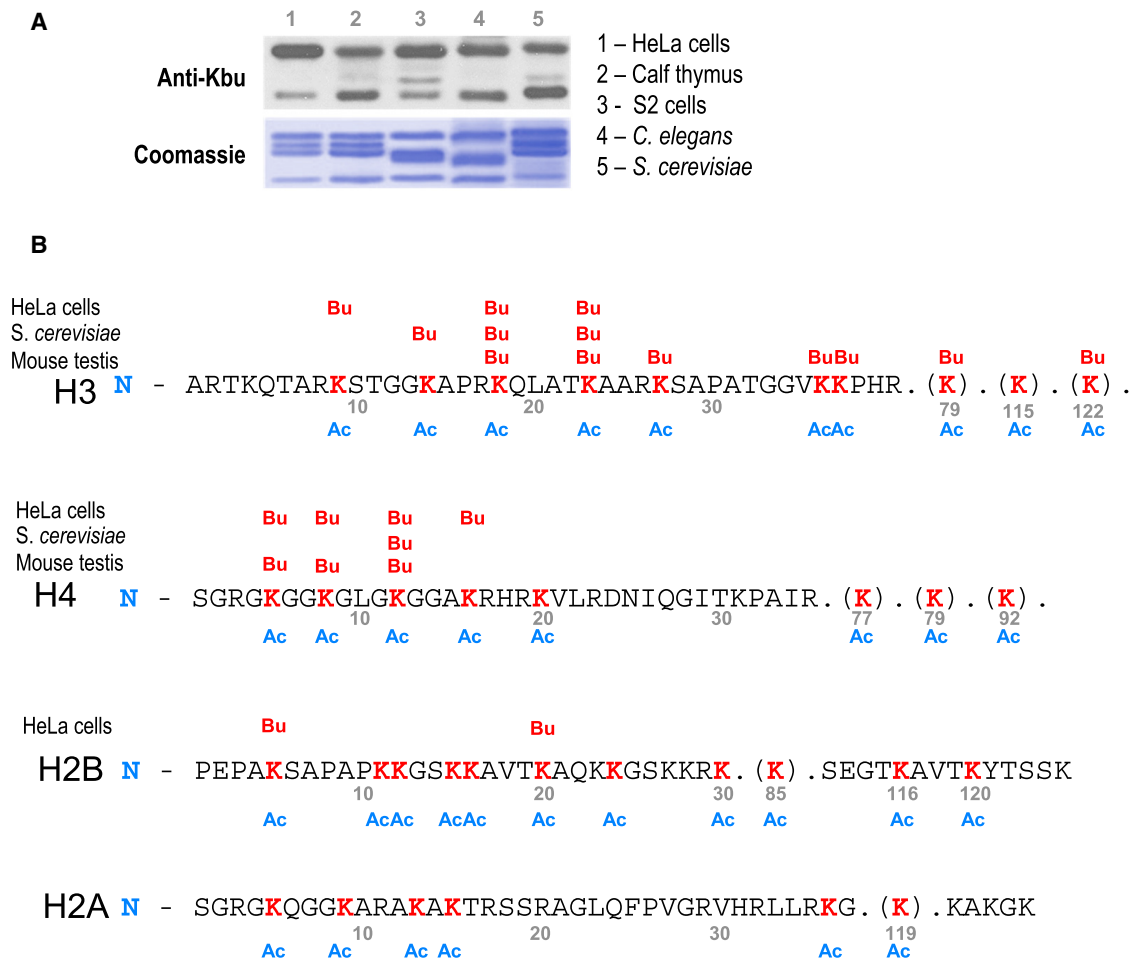


Figure 1. Histone Lysine Butyrylation Is an Evolutionarily Conserved PTM in Eukaryotic Cells

(A) The Kbu residues in core histones from the indicated species were detected by western blotting using a pan anti-Kbu antibody (upper). The corresponding Coomassie blue stained gel is shown (lower).

(B) Illustrations of histone Kac and Kbu sites in core histones identified by tandem mass spectrometry (MS/MS) (acetyl, Ac and butyryl, Bu). The annotated MS/MS spectra for histone Kbu peptides and the specific co-occurrence of K5bu and K8bu in H4 from spermatogenic cells are shown in [Data S1](#).

and butyrylation at H4K5 and H4K8 not only stimulates transcription, but could also underlie a highly dynamic interaction of histone post-translational modification (PTM)-binding factors such as Brdt. Additional data further show that stable differential use of acetylation and butyrylation could also durably affect genome organization in the maturing sperm. Altogether, these findings indicate how competition between histone acylation states could be an important epigenetic regulatory mechanism.

RESULTS

Histone Lysine Butyrylation Is an Evolutionarily Conserved PTM

To identify histone butyryllysine (Kbu) sites and study their function, we first confirmed the presence of histone Kbu by western blotting. Our data suggest that histone Kbu is an evolutionarily conserved PTM in eukaryotic cells ([Figure 1A](#)). We then used mass spectrometry to identify possible Kbu sites in core histones

from three species ([Chen et al., 2007](#); [Kim et al., 2006](#)). Kbu sites were detected in the N-terminal tails of H3 (K9, K14, K18, K23, K27, K36, K37, K79, and K122), H4 (K5, K8, K12, and K16), and H2B (K5 and K20) ([Figure 1B](#); [Data S1](#)).

Functional Significance of Histone Butyrylation

To examine the function of Kbu in chromatin biology, we used mammalian spermatogenesis as a model system. Spermatogenic cells can be roughly classified into three major types: proliferative progenitor cells (spermatogonia), meiotic cells (spermatocytes), and post-meiotic haploid cells (spermatids). During differentiation, highly specific gene expression programs are activated in both meiotic and early post-meiotic cells. Large-scale genome reorganization also takes place in spermatids, where a genome-wide replacement of histones by TPs and Prms occurs in post-meiotic cells known as elongating and condensing spermatids ([Gaucher et al., 2010](#); [Goudarzi et al., 2014](#); [Govin et al., 2004](#)).

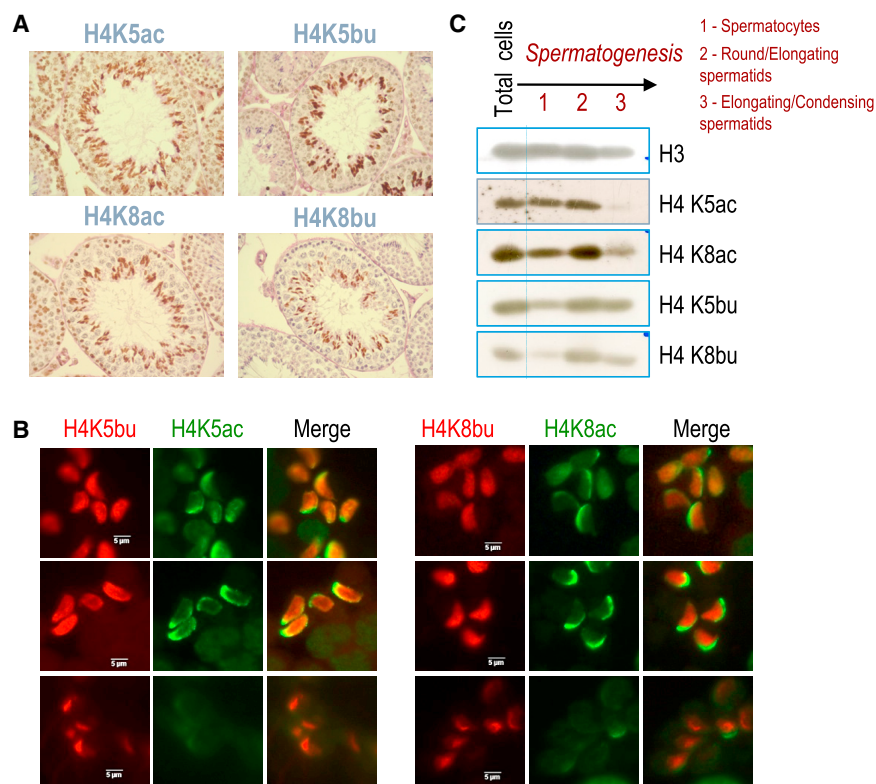


Figure 2. Occurrence of H4K5ac, K5bu, K8ac, and K8bu Marks during Mouse Spermatogenesis

(A and B) Detection by immunohistochemistry on paraffin sections of staged testis tubules (A) and by immunofluorescence in elongating/condensing spermatids (B) using the indicated antibodies are shown. The tubule sections shown in (A) represent stage IX or X of spermatogenesis. The use of mouse monoclonal anti-H4K5ac or H4K8ac antibodies together with rabbit polyclonal anti-H4K5bu or H4K8bu antibodies allowed for the co-detection of the indicated marks. (B) Spermatids are presented as a function of their differentiation along spermiogenesis (upper to lower) as judged by their morphology and degree of genome compaction. The scale bars represent 5 μ m. The images were acquired and treated under the same conditions.

(C) Total extracts from fractionated spermatogenic cells corresponding to cells enriched at the indicated stages along with extracts from the total (non-fractionated) spermatogenic cells were used to detect the indicated marks by western blotting. Compared to acetylation, longer exposure times were required for the butyrylation signal to be detected in all samples.

To confirm the existence of histone butyrylation in spermatogenic cells, we identified histone Kbu sites by mass spectrometry analysis of histones from mouse testis. We detected ten butyrylation sites including H4K5bu and H4K8bu (occurring separately or in combination), supporting the presence of these two histone marks in spermatogenic cells (Figure 1B; Data S1; see also Figure 7B). We then used highly specific anti-H4K5bu and anti-H4K8bu antibodies, along with the anti-H4K5ac and anti-H4K8ac antibodies, to investigate the stage-specific presence of these marks in spermatogenic cells. Immunohistochemistry analysis showed that H4K5 and K8 butyrylation is enhanced in elongating spermatids (Figure 2A), similar to earlier observations for histone H4K5 and H4K8 acetylation (Hazzouri et al., 2000).

To study the dynamic changes of butyrylation versus acetylation at both H4K5 and H4K8 sites in spermatogenic cells, we examined the co-existence of H4K5ac and H4K5bu as well as that of H4K8ac and H4K8bu. In elongating spermatids, H4K5ac and H4K8ac are widely distributed, but their localization becomes biased toward the sub-acrosomal regions in later stages. In contrast, H4K5bu- and H4K8bu-containing nucleosomes are homogeneously distributed in the same cells (Figure 2B). In late elongating spermatids, while acetylated histones are removed and degraded (Gaucher et al., 2012; Qian et al., 2013), butyrylated H4 species escape this wave of acetylation-dependent histone removal, to finally disappear in condensing spermatids (Figure 2B). Interestingly, immunoblotting experiments, using fractionated spermatogenic cells, confirmed that butyrylation persists longer than acetylation on H4K5 and K8 sites (Figure 2C).

These results highlight a bimodal histone removal process, whereby the removal of H4 K5/K8 butyrylated nucleosome occurs after that of H4 K5/K8 acetylated nucleosomes.

Genome-wide Distribution of H4K5 and H4K8 Acetylation and Butyrylation

Our immunofluorescence analysis showed distinctive patterns of acetylation and butyrylation at H4K5 and H4K8, with both marks showing enhanced labeling in elongating spermatids, when cellular transcription dramatically decreases (Gaucher et al., 2010). This observation motivated us to study their genomic distributions and potential functions at earlier stages of spermatogenesis, when specific spermatogenic gene expression programs are activated (Dai et al., 2014; Gaucher et al., 2012; Tan et al., 2011). Toward this goal, mouse spermatogenic cells were fractionated into spermatocytes and post-meiotic round spermatids (Dai et al., 2014; Gaucher et al., 2012; Tan et al., 2011). The two pools of cells were subjected to chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) using four anti-PTM-specific antibodies (anti-H4K5ac, -H4K8ac, -H4K5bu, and -H4K8bu). Analysis of the genomic distribution of these marks revealed that regions surrounding gene TSSs (located upstream of TSSs and at the 5'UTR) exhibit the highest coverage by the four PTMs in both spermatocytes and spermatids (Figure 3A). Further analyses showed that the four H4 PTMs are enriched at TSSs in a manner dependent on the transcriptional activity of the corresponding genes (Figure 3B).

A critical concern in the above experiments is the possibility that histone sites subject to acetylation may be butyrylated at

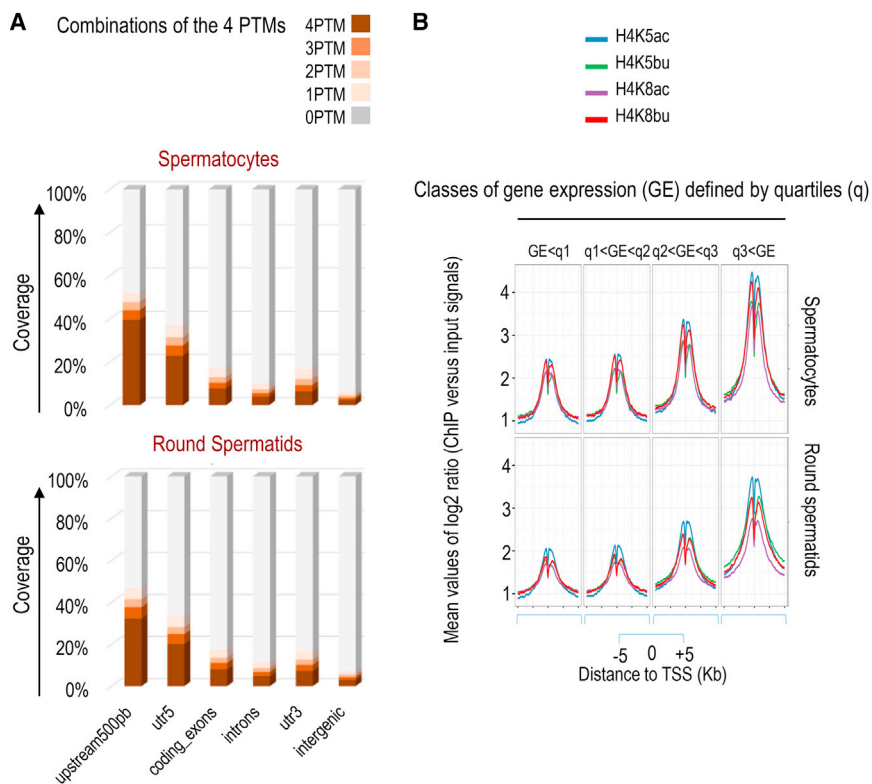


Figure 3. Genome-wide Mapping of H4K5ac, K5bu, K8ac, and K8bu Marks in Meiotic Spermatocytes and Post-meiotic Round Spermatids

(A) Bar diagrams showing the coverage of each of the genomic elements (as indicated) by different combinations of 4, 3, 2, 1, and 0 marks. 100% represents the total annotated genomic element in base pairs.

(B) Log2 ratio of ChIP-seq versus input signal corresponding to the four histone marks (with the indicated color code) is visualized over the gene TSS regions ($\pm 5,000$ bp). There are four gene categories that are ordered from left to right according to the level of their expression in the indicated cell types, spermatocytes, or round spermatids. To this aim, we used our transcriptomic analyses of these cells (Montellier et al., 2013). Antibody characterizations and ChIP-qPCR on selected regions are shown in Figure S1 and Table S1.

only low background levels, which would be misleadingly overestimated by the ChIP-seq analysis due to vastly different affinities of the antibodies used. To address this issue, we used surface plasmon resonance (SPR) to measure the affinities of antibodies for their respective targets. These measurements showed that all four antibodies have similar ranges of affinity (Figure S1A). Additionally, a ChIP-qPCR approach demonstrated that the four histone marks are significantly detected at selected genomic regions (Figures S1B and S1C). Together, these experiments confirmed that H4K5K8 butyrylation occurs at levels that largely exceed background noise.

Following these control experiments, we investigated in more detail the relationship between gene expression and the co-occurrence of the four TSS-associated histone marks. Remarkably, the most active genes were found associated with all four marks at their TSS regions (Figures 3B and 4A). In contrast, genes lacking any one of these marks at their TSS showed significantly reduced expression (Figure 4A). Other genomic elements did not show such a direct relationship between the co-existence of the four histone marks and gene expression (Figure S2).

To further analyze the relationship between the four acylation histone marks and gene activity, we took advantage of the differences in gene expression programs between spermatocytes (meiotic cells) and the transcriptionally active haploid round spermatids (generated after meiosis). Using our ChIP-seq data from these two cell types, we divided genes into four categories according to the intensity of TSS labeling by the four H4 acylation marks, namely: (1) genes bearing none of the four marks (labeled “no”) and (2–4) genes bearing all four marks, either with (2) com-

parable (“=”), (3) higher (“>”), or (4) lower (“<”) peak intensities in spermatocytes compared to round spermatids (Figure 4B). We observed that the four groups of genes belong to different gene expression programs. Genes in category (1) are largely unexpressed, while those in category (2) exhibited no change in expression level between meiotic and post-meiotic cells. In striking contrast, genes in categories (3) and (4) exhibited differential expression in the two cell types, which was positively associated with the change in intensity of TSS labeling by the four histone marks (Figure 4B).

Additional support for a positive correlation between the presence of the four PTMs in the TSSs and gene expression is the observation that all four marks are depleted on the sex chromosomes compared to autosomes, consistent with the chromosome-wide meiotic transcriptional inactivation known to characterize sex chromosomes (Figure 4C). To further investigate this observation, we specifically considered the fraction of sex chromosome-linked genes that escapes meiotic sex chromosome inactivation (Namekawa et al., 2006). While all four PTMs were identified on a majority (73%) of TSSs associated with active autosomal genes, co-occurrence of the four PTMs was observed in only a minority (23%) of TSSs associated with sex-linked genes that are reactivated in post-meiotic cells. Hence, over three-quarters of this latter category of TSSs bear between zero and three of the four histone marks. Interestingly, most of the TSSs bearing 1–3 of the PTMs were depleted of H4K8ac, but not of H4K8bu (Figure 4C). This is consistent with other studies showing depletion of acetylation on the TSSs of sex-linked post-meiotic genes and their labeling with other acyl groups such as crotonyl (Sin et al., 2012; Tan et al., 2011).

Histone Butyrylation Directly Stimulates Transcription

Our ChIP-seq data showed that, like acetylation, histone butyrylation is associated with high levels of gene expression,

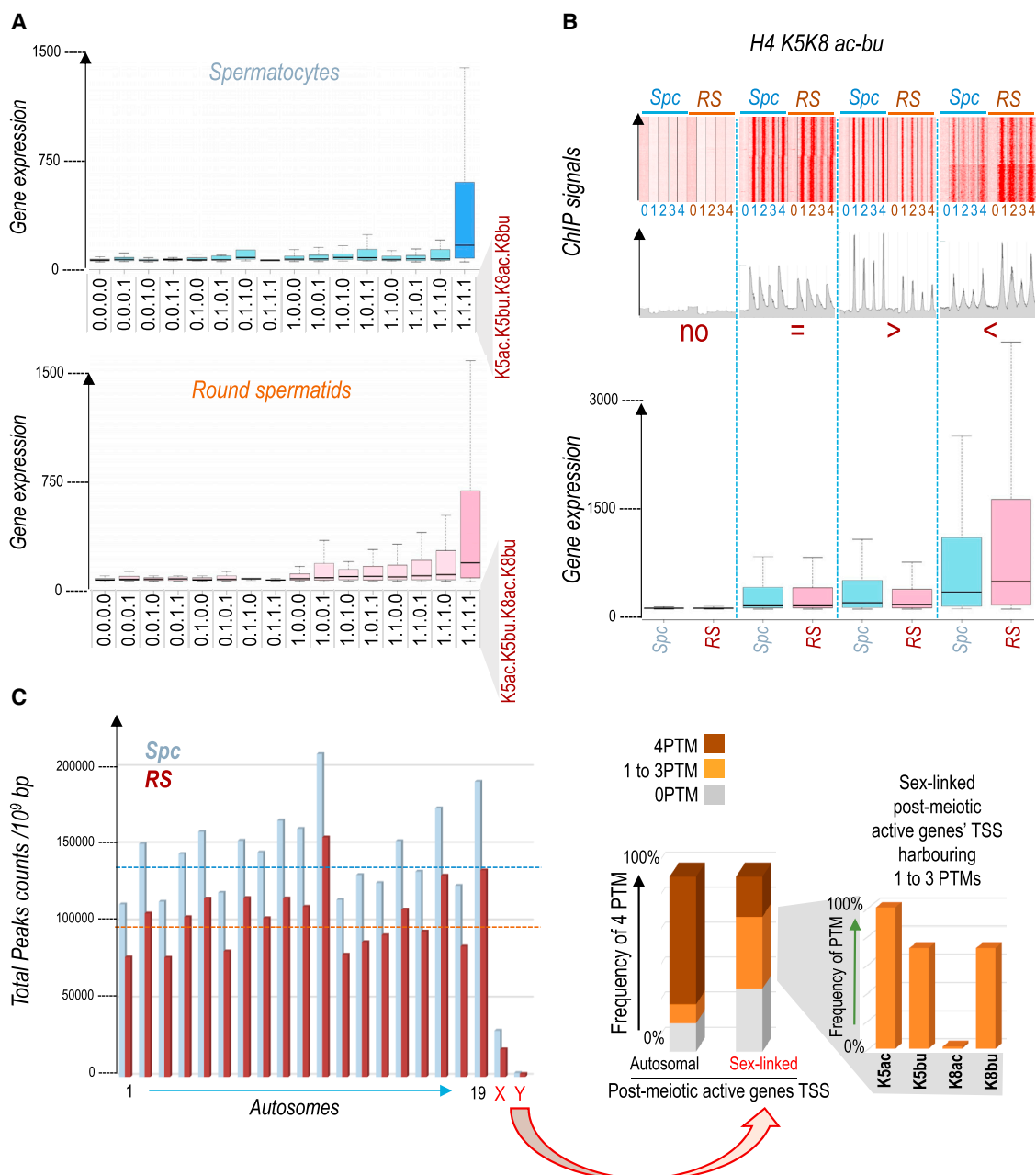


Figure 4. Tight Relationship between TSSs Labeled with H4K5ac, K5bu, K8ac, and K8bu Marks and Gene Transcriptional Activity

(A) The transcriptional activities of genes whose TSS regions bear the indicated histone H4 modifications are shown as box plots in spermatocytes (upper) and round spermatids (lower). The absence or presence of a histone mark is represented by the numbers “0” and “1”, respectively, and the four histone marks are arranged from bottom to top as follows: H4K5ac, H4K5bu, H4K8ac, and H4K8bu. The gene transcriptional activities as a function of various combinations of the four H4 PTMs on different genomic elements are shown in Figure S2.

(B) The heatmaps and profiles (upper) show the peak intensities of the ChIP signal at the TSS regions of genes associated with each of the four histone marks (“0” is input; “1” is K5ac; “2” is K5bu; “3” is K8ac; and “4” is K8bu), in spermatocytes (“Spc”), and in round spermatids (“RS”). The genes were assigned to four groups corresponding to: (1) genes with no ChIP peaks (“no”), (2) genes with moderate intensity of TSS labeling by the four marks and no significant change between spermatocytes and round spermatids (=), (3) genes with variation in TSS labeling intensities by the four marks with high meiotic peak intensity and a decreased intensity in post-meiotic cells (>), and (4) genes with an increased peak intensity in post-meiotic cells compared to spermatocytes (<). The box plots (lower) show the expression of the corresponding genes in meiotic and in round spermatids. These data were obtained using transcriptomic data of control samples from our previous work (Montellier et al., 2013; GSE46136).

(C) Total peak counts (acetyl and butyryl H4s) per million base pair for each chromosome in spermatocytes (blue bars) and round spermatids (red bars) are shown. The expected random distributions of the ChIP-seq peaks are indicated as dashed lines (upper). The right panel shows the proportion of gene TSSs harboring 0, 1, 2, 3, or 4 PTMs. (legend continued on next page)

suggesting the possibility that histone butyrylation directly stimulates gene expression. To test this hypothesis, we exploited a reconstituted activator-dependent *in vitro* transcription system. Our early studies had shown that *in vitro*, p300 and CREB binding protein (CBP) can catalyze lysine butyrylation by transferring the butyryl group from [^{14}C] butyryl-CoA to core histone proteins (Chen et al., 2007). This activity was also observed in *ex vivo* transfection experiments (Chen et al., 2007). Moreover, *in vitro* butyrylation activity was confirmed both on a reconstituted chromatin template and on histone octamers by mass spectrometry analysis. Further analysis of this *in vitro* activity revealed that p300 efficiently butyrylates the sites of interest, H4K5 and H4K8, in histone octamers as well as in chromatin (Figures 5A and 5B; Data S1).

After confirming that p300 is a histone butyryltransferase, we used a p300- and p53-dependent *in vitro* transcription system (Figure 5C) to test if histone butyrylation could stimulate transcription. We observed that p300-catalyzed histone butyrylation indeed directly stimulates transcription (Figure 5D). The mutation of lysine residues to arginine either on H3 or H4 tails eliminated acyl-CoA-stimulated transcription, indicating that acetyl/butyryl-CoA activates transcription through p300-catalyzed histone lysine acylation (Figure 5E). This experiment clearly demonstrates that, like acetylation, histone butyrylation can also directly stimulate gene transcriptional activity.

Brdt Binds to Gene TSSs Harboring H4 K5/K8 Acetylation and Butyrylation

Association of the four histone acylation marks with the TSSs of most of the highly active genes in spermatogenic cells raises the possibility that the high transcriptional activity of these genes is mediated by the binding of Brdt to the acylation marks on their TSSs. To test this hypothesis, we compared previously identified Brdt-bound TSSs (Gaucher et al., 2012) with TSSs labeled with the four histone marks. This analysis showed that most of the Brdt-bound gene TSSs also bore high levels of acetylation and butyrylation marks at H4K5 and H4K8 (Figures 6A and 6B). To study if Brdt's first bromodomain (BD1) mediates this interaction, we used spermatogenic cells from mice expressing a mutated form of Brdt lacking BD1 (Gaucher et al., 2012; Shang et al., 2007). The ChIP-seq analysis showed that, in both spermatocytes and round spermatids, the deletion of BD1 considerably weakens Brdt binding to TSSs bearing the four acylation marks (Figure 6A, Brdt Δ BD1). This result indicates that BD1 has a major role in targeting gene TSSs bearing H4K5K8 acetylation/butyrylation.

To further characterize histone acylation at gene TSSs bound by Brdt, we considered the occurrence of the four studied H4 PTMs at TSSs as a function of Brdt-binding. Remarkably, the vast majority of Brdt-associated TSSs bore all four acylation marks (Figure 6B). These observations therefore led us to question the ability of Brdt to bind H4 bearing butyrylation at either K5 or K8 or both.

H4 K5 Butyrylation Inhibits Brdt Binding

To test the ability of Brdt's bromodomains to bind a butyrylated H4 tail, we first carried out an *in vitro* binding assay using biotinylated H4 tail peptides bearing all four possible combinations of the K5ac, K5bu, K8ac, and K8bu acylation marks. We incubated the peptides with extracts from transfected Brdt-expressing Cos7 cells and performed pull-down experiments. As expected, Brdt efficiently interacts with H4 peptides that are either fully acetylated (i.e., on K5, K8, K12, and K16) or diacetylated on K5 and K8 (Figure S3A, upper). Strikingly, whereas Brdt binding was only slightly affected by the replacement of H4K8ac by H4K8bu, it was completely abolished by the substitution of H4K5ac by H4K5bu. Pull-down experiments using Brdt mutants bearing inactive bromodomains BD1 or BD2 demonstrated that Brdt binding to all the tested peptides depends on the integrity of its first bromodomain (Figure S3A, lower). These data clearly imply that the butyrylation of H4K5 inhibits the binding of Brdt to histone H4.

To validate this result, we performed the same experiment with nuclear extracts from mouse testis. Our results confirmed that butyrylation at H4K5 abolishes the binding of Brdt to H4 tails (Figure 6C, upper). As further confirmation, we repeated the pull-down assay on protein extracts from wild-type mouse testis using either fully acetylated or fully butyrylated immobilized H4 tail peptides and analyzed the bound fractions by mass spectrometry. Brdt was easily identified among the proteins affinity-isolated by the H4ac-containing peptide, but not by the H4bu-containing peptide (Figure S3B; Data S1). The use of extracts from mice testes expressing the Brdt Δ BD1 mutant confirmed that the Brdt-H4 tail interactions described above are primarily mediated by the BD1 domain (Figure 6C, lower right). This result was further corroborated by an experiment with JQ1, a BET bromodomain inhibitor, which abolished the binding of testis-derived Brdt to acetylated H4 and H4K5acK8bu peptides (Figure 6C, lower left). Taken together, these findings establish that H4K5bu, but not H4K8bu, abolishes the interaction between Brdt and the histone H4 tail.

Structural Analysis of the Effect of Butyrylation on the Brdt BD1-H4 Tail Interaction

To understand the molecular basis of the inhibitory effect of H4K5 butyrylation on Brdt binding, we carried out a structural modeling analysis. In the crystal structure of Brdt-BD1 bound to H4K5acK8ac, residue K5ac is intimately recognized by BD1, whereas K8ac makes fewer contacts (Morinière et al., 2009). Modeling shows that replacing the K8 acetyl group with a butyryl group allows the additional atoms to be accommodated without compromising any of the interactions between BD1 and the peptide (Figure 6D). In contrast, replacing the K5 acetyl group by a butyryl group results in a steric clash with residues in the domain's ZA loop, implying that some structural adjustments to the model are required to accommodate the butyryl group. Comparison with the published co-crystal structure of Brd4-BD1

1–3, or 4 of the studied H4 PTMs on autosomes and on the X chromosome in round spermatids (color coded). Of the X-linked genes that escape inactivation, the majority has either no H4K5K8 acetyl/butyryl marks or harbors one to three of these PTMs. A detailed consideration of these PTMs indicates that these TSSs are nearly always depleted in H4K8ac. The list of X-linked genes that are activated in post-meiotic cells was established based on our previous detailed post-meiotic transcriptomic analysis (Boussouar et al., 2014).

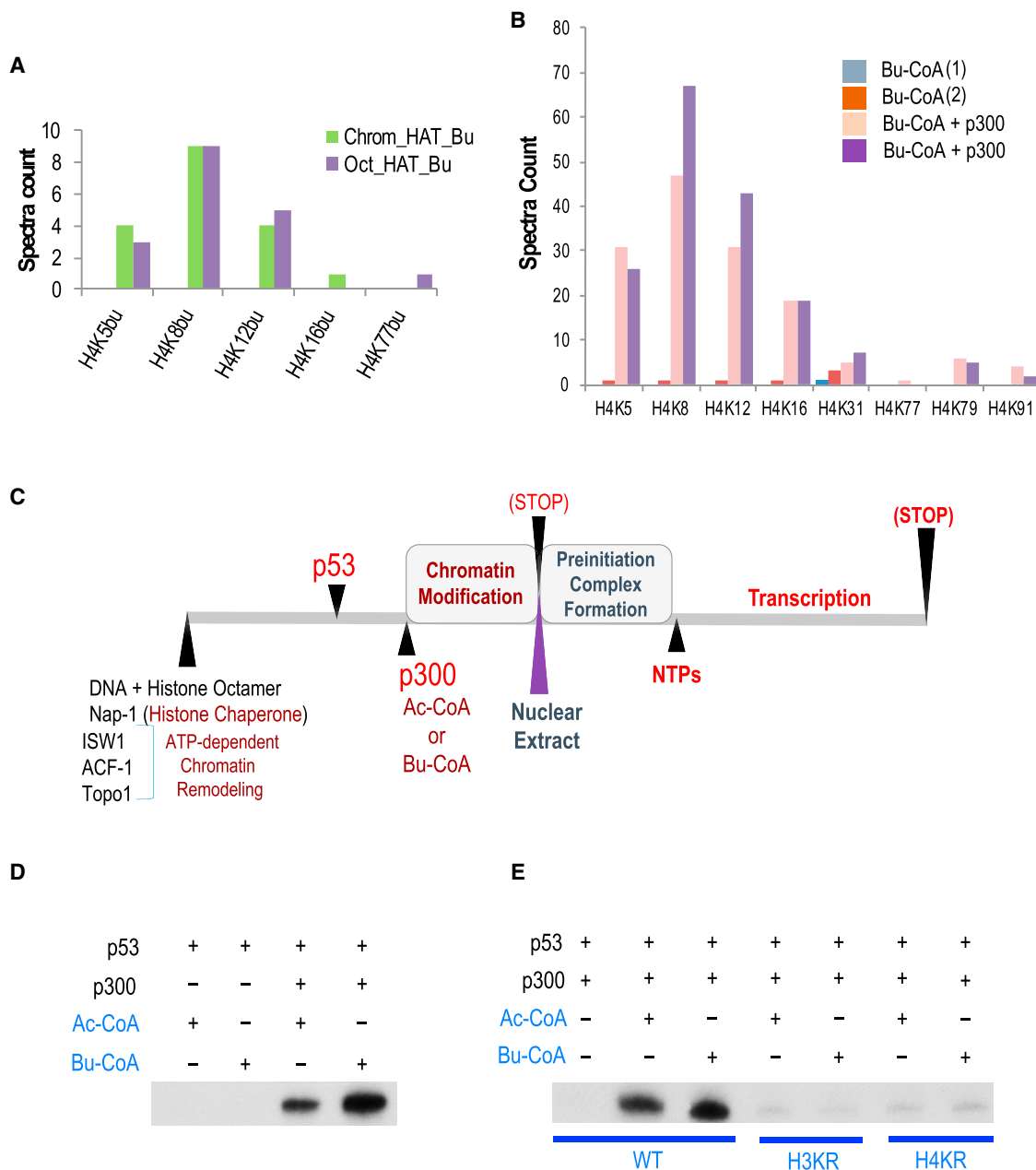


Figure 5. p300 Uses Butyryl-CoA to Butyrylate Histones in Histone Octamers and in Chromatin and to Directly Stimulate Gene Transcription

(A) In vitro reconstituted histone octamers or a chromatin template (used in the assay shown in C) were incubated with butyryl-CoA in the presence of purified p300 and butyrylated histones were detected by MS.

(B) In another assay, histone octamers were incubated with butyryl-CoA in the presence or absence of p300 to measure the degree of non-enzymatic histone butyrylation. The histograms represent the spectra counts of peptides containing the indicated butyrylated lysines in two different experiments (1 and 2: color coded). The annotated MS/MS spectra for histone Kbu peptides are presented in [Data S1](#).

(C) Schematic representation of the p53- and p300-dependent in vitro transcription assay showing the order of the added components ([An and Roeder, 2004](#)).

(D) Autoradiography of RNA products obtained by in vitro transcription under the conditions shown.

(E) The same reactions as in (D) were performed with chromatin templates generated by the use of either wild-type histone H3 and H4 or the indicated H3 and H4 K to R mutants.

bound to H3K14bu indicates the type of adjustments required to accommodate H4K5bu ([Vollmuth and Geyer, 2010](#)). Aligning the latter structure with that of the Brdt-BD1/H4K5acK8ac complex

shows that H3K14bu occupies approximately the same position as K5ac ([Figure 6E](#)). However, the bulkier butyryl group results in the displacement of a water molecule within the ligand-binding

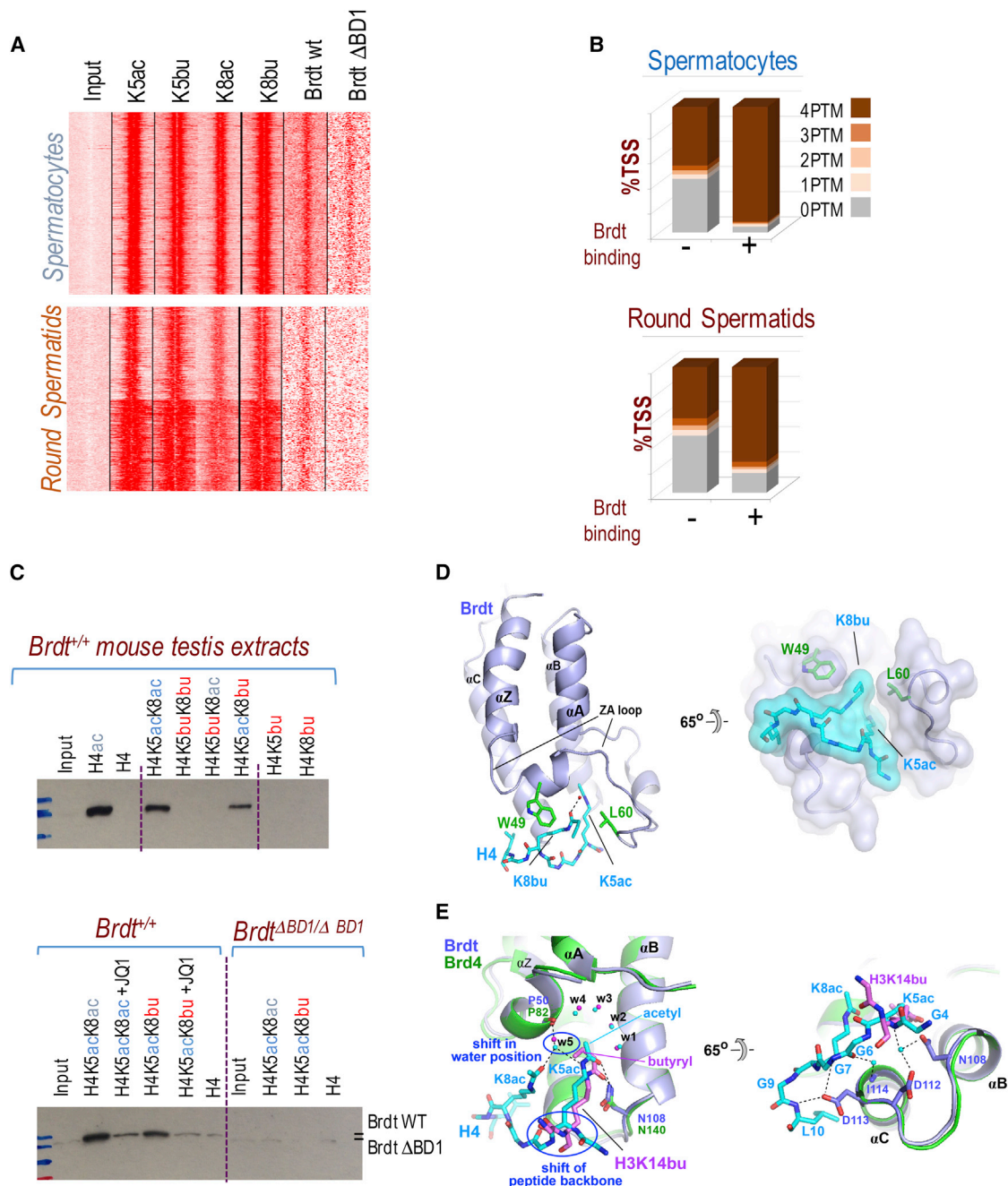


Figure 6. Brdt Is Preferentially Recruited to Gene TSS Regions Enriched in H4K5ac, K5bu, K8ac, and K8bu Marks

(A) ChIP-seq data from chromatin immunoprecipitation of Brdt either from wild-type fractionated spermatogenic cells or the corresponding fractionated cells expressing a truncated Brdt lacking its first bromodomain (Brdt Δ BD1) were obtained and compared with ChIP-seq data from the four indicated histone H4 marks from wild-type spermatogenic cells. Seqminer software illustrates gene TSSs bound by the wild-type Brdt (Brdt wt), Brdt Δ BD1, and the occurrence of the four H4 marks on the same regions. The input corresponds to the sequencing of chromatin fragments before ChIP from wild-type cells. Exactly the same profile was obtained for the input chromatin fragments from Brdt Δ BD1 cells (data not shown).

(B) Gene TSS regions were divided into two categories, either bound or unbound by Brdt, in spermatocytes (upper) or round spermatids (lower). For each category, the proportion of genes whose TSS regions are enriched with none, 1, 2, 3, or 4 of the four histone PTMs are represented with the indicated colors.

(C) Mouse total testis extracts were prepared from wild-type mice (upper and lower left) or mice expressing the truncation mutant Brdt Δ BD1 (lower right) and incubated with the indicated peptides. Brdt was then visualized after pull-down using an anti-Brdt antibody (Gaucher et al., 2012). In some experiments, BET bromodomain inhibitor JQ1 was added to the extract prior to the peptide pull-down step, as indicated. In another experiment Brdt was identified by MS/MS (Figure S3B), and the annotated MS/MS spectra for Brdt peptides are presented in Data S1. The pull-down experiments were also performed on Brdt-expressing transfected cells and the data shown in Figure S3A.

(legend continued on next page)

pocket and causes the K14bu main-chain and side-chain atoms to be shifted relative to those of K5ac. In the context of an H4K5buK8ac peptide, such shifts would be predicted to disrupt several direct and indirect hydrogen bonds between BD1 and the peptide (see figure legend for details). Thus, the modeling approach provides a plausible structural basis for the poor affinity observed for the binding of Brdt to the H4K5buK8ac peptide compared with that to the H4K5acK8ac and H4K5acK8bu peptides.

Dynamic Mixed Labeling of H4K5K8 by Acetylation and Butyrylation at Active Chromatin Sites

Our pull-down and structural modeling data strongly suggest that Brdt is inhibited from binding TSS regions where histone H4 is modified by lysine butyrylation at the K5 position. This conclusion appears contradictory to the ChIP-seq data, where H4K5bu and Brdt were both associated with the same TSS regions. A hypothesis that would reconcile these observations is that acetylation and butyrylation of H4K5 exhibit a rapid turnover. Indeed, acyltransferases associated with highly active gene TSSs might feasibly drain cellular acetyl-CoA as well as butyryl-CoA toward these sites, leading to a mixture of histone H4 acetylation and butyrylation marks. To test this hypothesis, we performed *in vitro* assays by incubating reconstituted histone octamers with purified p300 and an equimolar mixture of acetyl-CoA and butyryl-CoA. Histones were then analyzed by mass spectrometry to detect acetylated and/or butyrylated peptides. Our results show that p300 can use acyl-CoAs to catalyze acetylation and butyrylation at both H4K5 and H4K8 sites (Figure 7A), as we detected H4 peptides bearing various combinations of acetylated or butyrylated H4K5 and H4K8 (Figure 7A; Data S1). Encouraged by this *in vitro* result, we then investigated whether histone H4 isolated from spermatogenic cells also contains diverse lysine acylation marks. Mass spectrometry analysis of these samples detected H4 peptides with various combinations of acetylation and butyrylation at H4K5 and H4K8.

In addition, depending on cell type, the stoichiometry of H4K5bu and H4K8bu could be higher than some of the widely studied histone marks such as H3K4me3 (Kulej et al., 2015) (Figures 7B and S4A; Table S2), but lower than those of H4 tail acetylation, such as that of K16, which can be as high as 20% of H4 species (Kulej et al., 2015).

Acetylation is known to have a high turnover rate at gene TSSs (Crump et al., 2011). Therefore, due to continuous acylation by histone acetyltransferase (HATs) such as p300 and the

rapid turnover of these PTMs, it is feasible that H4K5 and K8 alternate between acetylated and butyrylated states. This model is supported by the detection of various combinations of H4K5K8 acetylation and butyrylation either in *in vitro* HAT assays (Figure 7A) or *in vivo* in spermatogenic cells (Figure 7B). A direct consequence of such alternating histone acetylation/butyrylation would be the dynamic binding of Brdt, which would oscillate between high- and low-affinity states depending on the acylation status of H4K5 (Figure 7C). Such dynamic histone H4 acylation could be facilitated by open nucleosomes on the corresponding gene TSSs by specific histone variants such as H2A.Lap1, a histone H2A variant capable of inducing unstable and open nucleosomes and known to associate with active gene TSSs in spermatogenic cells (Nekrasov et al., 2013; Soboleva et al., 2012). Indeed, analysis of ChIP-seq data for H2A.Lap1 revealed that the four H4 PTMs are particularly enriched on H2A.Lap1-associated TSSs (Figure S4B).

Taken together, the above results support a model whereby alternating competing histone acetylation and butyrylation underlie a dynamic interaction between the histone modifications and the cognate bromodomain.

DISCUSSION

The present study reports our findings regarding the interplay between histone acetylation and butyrylation on the histone H4 tail during sperm cell differentiation. We found that *in vitro*, p300 uses available acetyl-CoA and butyryl-CoA sources to acylate the H4 tail at all the acceptor lysines in an indiscriminate manner. A proteomic approach also revealed the co-existence of the same combinations of H4 K5K8 acetyl and butyryl marks in different spermatogenic cell types. ChIP-seq analysis on fractionated spermatogenic cells further demonstrated that H4K5/K8 acetyl/butyryl are particularly enriched at the TSSs of the most active genes. Strikingly, however, functional and structural analysis revealed that the acylation state of H4K5 is a critical determinant of Brdt binding affinity, as Brdt binds the acetylated, but not the butyrylated state. In this context, a stable maintenance of differential acetylation and butyrylation could have important functional consequences for the genome reorganization observed during spermatogenesis. We found that, in contrast to earlier stages, in elongating spermatids, H4K5 and H4K8 acetylation and butyrylation become more markedly associated with specific regions of the genome. This could have a direct consequence on the action of Brdt in these cells.

(D and E) Modeling of Brdt-BD1 bound to H4 peptides bearing mixed butyryl and acetyl groups.

(D) Hypothetical model of an H4K5acK8bu peptide bound to Brdt-BD1. The model is based on the crystal structure of the Brdt/H4K5acK8ac complex and was built by substituting the K8 acetyl group with a butyryl group. The substitution allows all contacts between BD1 and the diacetylated peptide to be preserved, while the additional atoms of the butyryl group are accommodated within the space between residues Trp49 and Leu60.

(E) Comparison of the Brdt and Brd4 BD1 bromodomains bound to diacetylated and monobutyrylated peptides, respectively. The two structures were aligned via the pocket-defining BC and ZA loops. Brd4 is shown in green, and its associated water molecules and butyrylated peptide (for which only two residues are present in the crystal structure) are shown in magenta. Brdt is shown in light blue, and its water molecules and peptide ligand are shown in cyan. For clarity, most of the ZA loop is omitted. Brdt-BD1 residues 108 and 112–114 make direct and water-mediated hydrogen bonds to the peptide backbone atoms of H4 residues K5ac, G6, G7, and L10. The shift in peptide backbone position required to accommodate a butyryl group on K5 would disrupt some or all of these interactions. In the Brdt-BD1/H4K5acK8ac complex, hydrogen bonds mediated by water molecule w5 play a key structural role in linking the two acetylated lysine residues to each other and to the backbone carbonyl of BD1 residue Pro50. In the structure of Brd4-BD1/H3K14Bu, w5 is displaced by 1.1 Å toward Pro82, while the amide nitrogen of K14Bu is displaced in the opposite direction by 1.3 Å. Thus, compared to the diacetylated peptide, the presence of the butyryl group disrupts a water-mediated hydrogen bond between BD1 and the peptide.

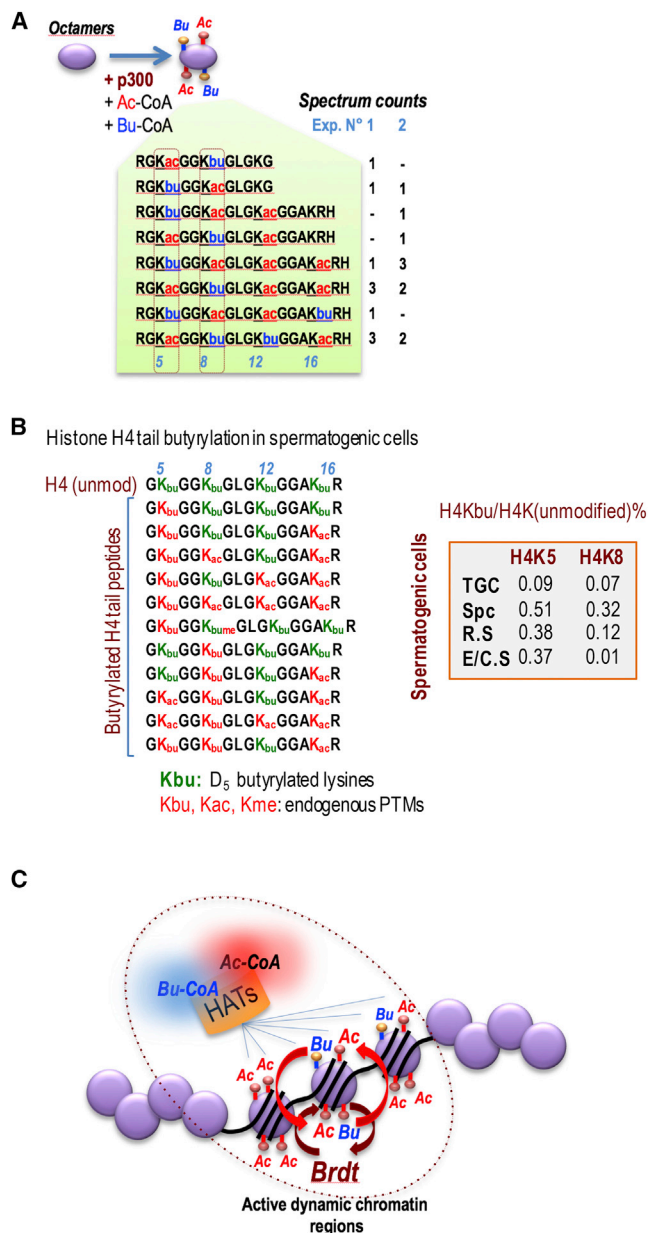


Figure 7. p300-Mediated Acetylation and Butyrylation of H4K5K8 and Functional Implications for Gene Promoter Activity

(A) In vitro reconstituted octamers were incubated with purified p300 in the presence of equal amounts of acetyl-CoA and butyryl-CoA, and the histones were subsequently analyzed by MS. The unique histone H4 peptides bearing multiple lysine modifications, which were specifically identified, are indicated along with their corresponding spectrum counts. The annotated MS/MS spectra for histone Kbu/ac peptides are presented in [Data S1](#).

(B) Stoichiometry of H4K5bu and K8bu in sperm cells. The green letters indicate the in vitro chemical butyrylation, and the red letters indicate the endogenous modifications. The % represents the respective ratios of the peptide bearing H4K5bu and H4K8bu (summed peak areas) over the corresponding unmodified H4 (total germ cells: TGC; spermatocytes: Spc; round spermatids: RS; and elongating/condensing spermatids: E/CS). The values of peak area were used for calculations ([Table S2](#); [Figure S4A](#)).

(C) Model of the functional interplay between acylation marks. We propose that at active gene TSSs, p300 (and possibly other HATs) randomly use acetyl- or

Indeed, we previously showed that in elongating spermatids, when histone hyperacetylation and a general transcriptional shut-down occur, BD1 is indispensable for the replacement of acetylated histones by TPs ([Gaucher et al., 2012](#)). Additionally, the hyperacetylated histones are known to be directly targeted for degradation by a PA200-containing specialized proteasome ([Qian et al., 2013](#)), suggesting that the Brdt-bound histone population enters this pathway. Here, we show that in elongating spermatids, butyrylated histones survive this wave of acetylation-dependent histone removal and degradation. This survival is perfectly consistent with the inability of Brdt to recognize butyrylated H4. These observations suggest that a stable differential labeling of H4 by acetylation and butyrylation may control the timing of histone removal. In this context, it is also possible to speculate that nucleosomes bearing butyrylated H4 could undergo a direct histone-to-Prm replacement due to the inability of Brdt to mediate the exchange of histones by TPs. Indeed, *TP1-TP2* double KO cells can undergo direct histone-to-Prm replacement but, in this case, Prms are unable to tightly compact the genome (for review, see [Gaucher et al., 2010](#)). The prediction is therefore that regions bearing butyrylated histones would evolve to a less compact structure in mature sperms than regions that are marked by acetylated histones before their removal. This regulatory mechanism could be an elegant way to introduce differences in genome compaction by Prms.

Differential histone tail acylation might also play important roles in the control of somatic cell gene expression, differentiation, and genome programming. Indeed, a recent study of p300-mediated histone H3K18 crotonylation revealed that an increase in cellular levels of crotonyl-CoA favors a more efficient de novo gene activation ([Sabari et al., 2015](#)). Although the precise mechanism underlying the role of H3K18cr in transcription is unknown, the differential affinity of a bromodomain protein toward crotonyl versus acetyl marks could conceivably account for the observed phenomenon.

Previous reports demonstrated a rapid turnover of histone acetylation on a sub-fraction of nucleosomes ([Clayton et al., 2006](#); [Waterborg, 2002](#)), overlapping with active gene TSSs ([Crump et al., 2011](#)). However, the functional implications of the short half-life of histone acetylation on active chromatin regions, especially on highly transcribed gene TSSs, have remained elusive. The results described in our present study are consistent with a rapid turnover of both acetyl and butyryl marks on the H4 tail. Given the differential affinity of Brdt BD1 for the acetylated and butyrylated forms of H4K5, a rapid alternation of acylation states at H4K5 would result in a highly dynamic interaction between Brdt and chromatin. Thus, it is tempting to speculate that the generally observed rapid turnover of acetyl marks may be functionally significant because it enables rapid

butyryl-CoA to modify histone tails. HDACs and histone variants, also present at dynamic chromatin regions, accelerate the turnover of these histone marks. Alternate modifications of H4K5 with acetyl and butyryl marks, while maintaining the region competent for transcription, also lead to an oscillation in Brdt binding. The co-localization of H2A.Lap1 and the four H4 PTMs at gene TSSs is shown in [Figure S4B](#). Dynamic Brdt binding could help the pre-initiation complex and RNA PolII to reassemble at TSSs and would favor successive cycles of gene transcription.

transitions between alternative states of lysine acylation. Such a mechanism, while maintaining histones permanently modified, would allow for a dynamic association with specific bromodomains, which might be important for sustaining successive cycles of transcription. Thus, either a change in the ratio of cellular acyl-CoAs (for instance caused by metabolic disorders; [Pougovkina et al., 2014](#)) or differential activities of histone deacetylase (HDACs) in removing acyl-groups ([Rousseaux and Khochbin, 2015](#)) could reprogram gene expression profiles.

In conclusion, we provide here the first demonstration that the interchangeable use of two closely related histone acylation marks at a specific site, H4K5, has important functional consequences by modulating the ability of a transcriptional regulator, Brdt, to recognize chromatin. This finding should improve our understanding of gene transcriptional regulation and its link to cell metabolism.

EXPERIMENTAL PROCEDURES

Antibodies and Other Reagents

Pan anti-Kbu antibody, anti-histone site-specific Kbu, and acetyllysine antibodies used were purchased from PTM Biolabs and anti-histone H3 and anti-histone H4 antibodies were from Abcam. Mouse monoclonal antibodies against H4K5ac and H4K8ac were a generous gift from Dr. H. Kimura. Anti-Brdt is a homemade antibody previously described ([Gaucher et al., 2012](#)). Butyryl-CoA and acetyl-CoA were purchased from Sigma-Aldrich. The modified porcine trypsin was purchased from Promega. HPLC-grade acetonitrile, water, and ethanol were purchased from EMD Chemicals. Peptides bearing one or a few acetyl and Kbu residues were custom synthesized and were verified by HPLC and mass spectrometry. JQ1 was synthesized by Charles McKenna and Elena Ferri (University of Southern California) ([Emadali et al., 2013](#)) and used as described in this reference.

The antigen recognition capacities of the antibodies used in ChIP-seq and ChIP-qPCR were determined using SPR as described in the [Supplemental Information](#).

Identification of Kbu Sites in Core Histones by Affinity Enrichment and Mass Spectrometry

Acetylation and butyrylation of histones were determined on 200 μ g of core histones extracted from the different studied cell types. The detailed procedures are described in the [Supplemental Information](#).

Quantification of H4K5K8 Butyrylation in Spermatogenic Cell Populations

Histone from spermatogenic cells underwent chemical butyrylation with deuterated (D5) butyryl anhydride and processed as described in the [Supplemental Information](#).

Analysis of Mouse Spermatogenic Cells, ChIP, Bioinformatics, and In Vitro Transcription

All the experimental procedures, including immunostaining, cell fractionation, and ChIP-seq were carried out exactly as described previously ([Dai et al., 2014](#); [Gaucher et al., 2012](#); [Tan et al., 2011](#)). ChIP-qPCR was performed following our ChIP-seq protocol, but the recovered DNA was amplified using quantitative PCR and specific primers corresponding to regions indicated in [Figure S1](#). [Table S1](#) indicates the sequence of these primers and details of the ChIP experiments are described in [Supplemental Information](#).

Brdt pull-down assays were performed as previously described ([Huang et al., 2010](#); [Pivot-Pajot et al., 2003](#)). The in vitro transcription assay was also previously described ([An et al., 2002](#); [An and Roeder, 2004](#); [Sabari et al., 2015](#); [Tang et al., 2013](#)). The bioinformatic analyses followed a pipeline similar to the one previously described ([Dai et al., 2014](#); [Gaucher et al., 2012](#); [Montellier et al., 2013](#); [Tan et al., 2011](#)) and are detailed in the [Supplemental Information](#).

p300 Expression and Purification for In Vitro Acetyl/Butyryl-Transferase Assays on Histone Octamer

In vitro acetyl/butyryl-transferase assays on histone octamer used recombinant p300 expressed and purified from baculovirus infected in Sf21 insect cells. Protein purification and the details of the HAT assays are described in the [Supplemental Information](#).

ACCESSION NUMBERS

The accession numbers for the H4 K5 and K8 acetylation and butyrylation ChIP-seq data and Brdt ChIP-seq data reported in this paper are GEO: GSE77277 and GSE39910, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, two tables, and one data file and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2016.03.014>.

AUTHOR CONTRIBUTIONS

Y.Z. and S.K. designed and coordinated the whole project. A.G. performed most of the experiments on spermatogenic cells. S.B., J.G., and E.M. were involved in the ChIP setup and S.B. helped with the final experiments. A.-L.V. designed appropriate primers and performed qPCR after ChIP from total germ cells. J.G. and D.P. performed in vitro HAT assays. Z.T. and R.G.R. performed in vitro transcription assays. S.C. helped with pull-down assays. T.B. helped with spermatogenic cell fractionation. C.P. helped with structural analyses. A.D., G.C., and D.P. helped with the bioinformatic analyses under the supervision of S.R. D.P. D.Z. performed the IP/pull-down/MS/MS experiments, PTM quantifications, antibody characterization (SPR), and coordinated the MS analyses. H.H. and O.K. performed the HPLC-MS/MS experiments and data analysis. S.Q. constructed the ChIP-seq library for sequencing. T.H. and Z.C. were involved in antibody production and characterization.

ACKNOWLEDGMENTS

The S.K. laboratory is supported by a grant from the Fondation pour la Recherche Médicale (FRM) "analyse bio-informatique pour la recherche en biologie" program as well as the ANR Episperm3 program (ANR-15-CE12-0005-02 to S. K. and D. P.) and by INCa (2013-082 program to S.K. and C. P.). D. P. laboratory is also supported by the World Wide Cancer Research foundation (16-0280). A.G. has been a recipient of a Marie Curie Initial Training Network funded by the European Commission (FP7-PEOPLE-2011-ITN, PITN-GA-289880) for three years and is now supported by "Fondation ARC." G.C. is supported by the above-mentioned FRM grant. The Y.Z. laboratory is supported by the NIH (GM105933, DK107868, and GM115961). Y.Z. is a shareholder and a member of the scientific advisory board of PTM Biolabs, Co., Ltd. T.H. and Z.C. are PTM Biolabs employees. Work in the laboratory of R.G.R. was supported by grants from the NIH and the Starr Foundation Tri-Institutional Stem Cell Initiative (2014-034). We gratefully acknowledge the generous gift of mouse monoclonal antibodies against H4K5ac and H4K8ac by Dr. H. Kimura. JQ1 was synthesized and provided to us by Charles McKenna and Elena Ferri. We also acknowledge Miho Shimada for the preparation of histone mutants.

Received: November 4, 2015

Revised: February 5, 2016

Accepted: March 10, 2016

Published: April 21, 2016

REFERENCES

An, W., and Roeder, R.G. (2004). Reconstitution and transcriptional analysis of chromatin in vitro. *Methods Enzymol.* 377, 460–474.

- An, W., Palhan, V.B., Karymov, M.A., Leuba, S.H., and Roeder, R.G. (2002). Selective requirements for histone H3 and H4 N termini in p300-dependent transcriptional activation from chromatin. *Mol. Cell* 9, 811–821.
- Boussouar, F., Goudarzi, A., Buchou, T., Shiota, H., Barral, S., Debernardi, A., Guardiola, P., Brindle, P., Martinez, G., Arnoult, C., et al. (2014). A specific CBP/p300-dependent gene expression programme drives the metabolic remodelling in late stages of spermatogenesis. *Andrology* 2, 351–359.
- Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C., Falck, J.R., Peng, J., Gu, W., and Zhao, Y. (2007). Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol. Cell. Proteomics* 6, 812–819.
- Clayton, A.L., Hazzalin, C.A., and Mahadevan, L.C. (2006). Enhanced histone acetylation and transcription: a dynamic perspective. *Mol. Cell* 23, 289–296.
- Crump, N.T., Hazzalin, C.A., Bowers, E.M., Alani, R.M., Cole, P.A., and Mahadevan, L.C. (2011). Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP. *Proc. Natl. Acad. Sci. USA* 108, 7814–7819.
- Dai, L., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., Debernardi, A., Buchou, T., Rousseaux, S., Jin, F., et al. (2014). Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat. Chem. Biol.* 10, 365–370.
- Emadali, A., Rousseaux, S., Bruder-Costa, J., Rome, C., Duley, S., Hamaidia, S., Betton, P., Debernardi, A., Leroux, D., Bernay, B., et al. (2013). Identification of a novel BET bromodomain inhibitor-sensitive, gene regulatory circuit that controls Rituximab response and tumour growth in aggressive lymphoid cancers. *EMBO Mol. Med.* 5, 1180–1195.
- Flynn, E.M., Huang, O.W., Poy, F., Oppikofer, M., Bellon, S.F., Tang, Y., and Cochran, A.G. (2015). A subset of human bromodomains recognizes butyryllysine and crotonyllysine histone peptide modifications. *Structure* 23, 1801–1814.
- Gaucher, J., Reynold, N., Montellier, E., Boussouar, F., Rousseaux, S., and Khochbin, S. (2010). From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J.* 277, 599–604.
- Gaucher, J., Boussouar, F., Montellier, E., Curtet, S., Buchou, T., Bertrand, S., Hery, P., Jounier, S., Depaux, A., Vitte, A.L., et al. (2012). Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J.* 31, 3809–3820.
- Goudarzi, A., Shiota, H., Rousseaux, S., and Khochbin, S. (2014). Genome-scale acetylation-dependent histone eviction during spermatogenesis. *J. Mol. Biol.* 426, 3342–3349.
- Govin, J., Caron, C., Lestrat, C., Rousseaux, S., and Khochbin, S. (2004). The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur. J. Biochem.* 271, 3459–3469.
- Hazzouri, M., Pivot-Pajot, C., Faure, A.K., Usson, Y., Pelletier, R., Sèle, B., Khochbin, S., and Rousseaux, S. (2000). Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur. J. Cell Biol.* 79, 950–960.
- Huang, R., Holbert, M.A., Tarrant, M.K., Curtet, S., Colquhoun, D.R., Dancy, B.M., Dancy, B.C., Hwang, Y., Tang, Y., Meeth, K., et al. (2010). Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation. *J. Am. Chem. Soc.* 132, 9986–9987.
- Kim, S.C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., et al. (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* 23, 607–618.
- Kulej, K., Avgousti, D.C., Weitzman, M.D., and Garcia, B.A. (2015). Characterization of histone post-translational modifications during virus infection using mass spectrometry-based proteomics. *Methods* 90, 8–20.
- Montellier, E., Rousseaux, S., Zhao, Y., and Khochbin, S. (2012). Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *BioEssays* 34, 187–193.
- Montellier, E., Boussouar, F., Rousseaux, S., Zhang, K., Buchou, T., Fenaille, F., Shiota, H., Debernardi, A., Héry, P., Curtet, S., et al. (2013). Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev.* 27, 1680–1692.
- Morinière, J., Rousseaux, S., Steuerwald, U., Soler-López, M., Curtet, S., Vitte, A.L., Govin, J., Gaucher, J., Sadoul, K., Hart, D.J., et al. (2009). Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 461, 664–668.
- Namekawa, S.H., Park, P.J., Zhang, L.F., Shima, J.E., McCarrey, J.R., Griswold, M.D., and Lee, J.T. (2006). Postmeiotic sex chromatin in the male germline of mice. *Curr. Biol.* 16, 660–667.
- Nekrasov, M., Soboleva, T.A., Jack, C., and Tremethick, D.J. (2013). Histone variant selectivity at the transcription start site: H2A.Z or H2A.Lap1. *Nucleus* 4, 431–438.
- Pivot-Pajot, C., Caron, C., Govin, J., Vion, A., Rousseaux, S., and Khochbin, S. (2003). Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol. Cell. Biol.* 23, 5354–5365.
- Pougovkina, O., Te Brinke, H., Wanders, R.J., Houten, S.M., and de Boer, V.C. (2014). Aberrant protein acylation is a common observation in inborn errors of acyl-CoA metabolism. *J. Inher. Metab. Dis.* 37, 709–714.
- Qian, M.X., Pang, Y., Liu, C.H., Haratake, K., Du, B.Y., Ji, D.Y., Wang, G.F., Zhu, Q.Q., Song, W., Yu, Y., et al. (2013). Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* 153, 1012–1024.
- Rousseaux, S., and Khochbin, S. (2015). Histone acylation beyond acetylation: Terra incognita in chromatin biology. *Cell J.* 17, 1–6.
- Sabari, B.R., Tang, Z., Huang, H., Yong-Gonzalez, V., Molina, H., Kong, H.E., Dai, L., Shimada, M., Cross, J.R., Zhao, Y., et al. (2015). Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol. Cell* 58, 203–215.
- Shang, E., Nickerson, H.D., Wen, D., Wang, X., and Wolgemuth, D.J. (2007). The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development* 134, 3507–3515.
- Sin, H.S., Barski, A., Zhang, F., Kartashov, A.V., Nussenzweig, A., Chen, J., Andreassen, P.R., and Namekawa, S.H. (2012). RNF8 regulates active epigenetic modifications and escape gene activation from inactive sex chromosomes in post-meiotic spermatids. *Genes Dev.* 26, 2737–2748.
- Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J. (2012). A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat. Struct. Mol. Biol.* 19, 25–30.
- Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016–1028.
- Tan, M., Peng, C., Anderson, K.A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., Zhang, Y., et al. (2014). Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. *Cell Metab.* 19, 605–617.
- Tang, Z., Chen, W.Y., Shimada, M., Nguyen, U.T., Kim, J., Sun, X.J., Sengoku, T., McGinty, R.K., Fernandez, J.P., Muir, T.W., and Roeder, R.G. (2013). SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. *Cell* 154, 297–310.
- Vollmuth, F., and Geyer, M. (2010). Interaction of propionylated and butyrylated histone H3 lysine marks with Brd4 bromodomains. *Angew. Chem. Int. Ed. Engl.* 49, 6768–6772.
- Waterborg, J.H. (2002). Dynamics of histone acetylation in vivo. A function for acetylation turnover? *Biochem. Cell Biol.* 80, 363–378.
- Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y. (2012). Lysine succinylation and lysine malonylation in histones. *Mol. Cell. Proteomics* 11, 100–107.

Supplemental Information

Dynamic Competing Histone H4 K5K8

Acetylation and Butyrylation Are Hallmarks

of Highly Active Gene Promoters

Afsaneh Goudarzi, Di Zhang, He Huang, Sophie Barral, Oh Kwang Kwon, Shankang Qi, Zhanyun Tang, Thierry Buchou, Anne-Laure Vitte, Tieming He, Zhongyi Cheng, Emilie Montellier, Jonathan Gaucher, Sandrine Curtet, Alexandra Debernardi, Guillaume Charbonnier, Denis Puthier, Carlo Petosa, Daniel Panne, Sophie Rousseaux, Robert G. Roeder, Yingming Zhao, and Saadi Khochbin

Supplemental Information

Supplemental Figures

[Figure S1](#) to [Figure S4](#) and their legend

Supplemental Tables

[Table S1](#) and [Table S2](#) and their legend

Supplemental experimental procedures

Supplemental references

Figure S1 (related to Figure 3)

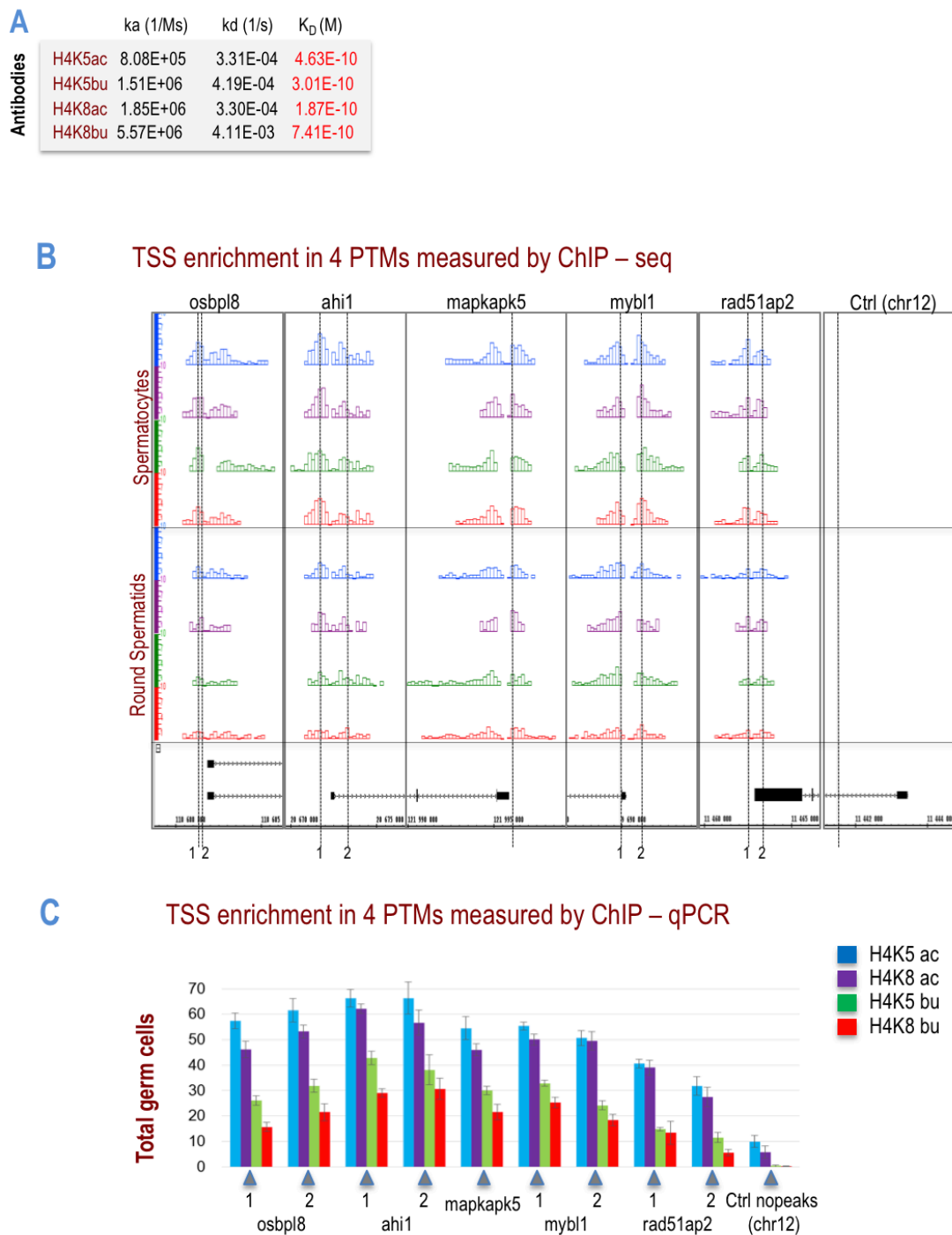


Figure S1. Characterization of antibodies used in ChIP-seq by SPR and ChIP-qPCR on several regions. (A) The indicated antibodies employed in ChIP experiments were used to measure their respective k_a , k_d and K_D . The indicated values represent the average of three measurements. (B) Genomic regions within or around TSS bearing the four studied H4 PTMs in spermatocytes and round spermatids according to ChIP-seq, were chosen for ChIP-qPCR and are shown here on the IGB genome browser as dashed lines. (C) Total germ cells were isolated and the corresponding chromatin digested with MNase and the resulting nucleosomes were immunoprecipitated with the indicated antibodies (colour coded). The histograms show the results of four ChIP-qPCR obtained from two independent ChIPs experiments. Values are represented as a proportion of the qPCR amplification of 1/400 of the input. Standard deviations of four values (2 independent experiments in duplicates) are shown.

Figure S2 (related to Figure 4A)

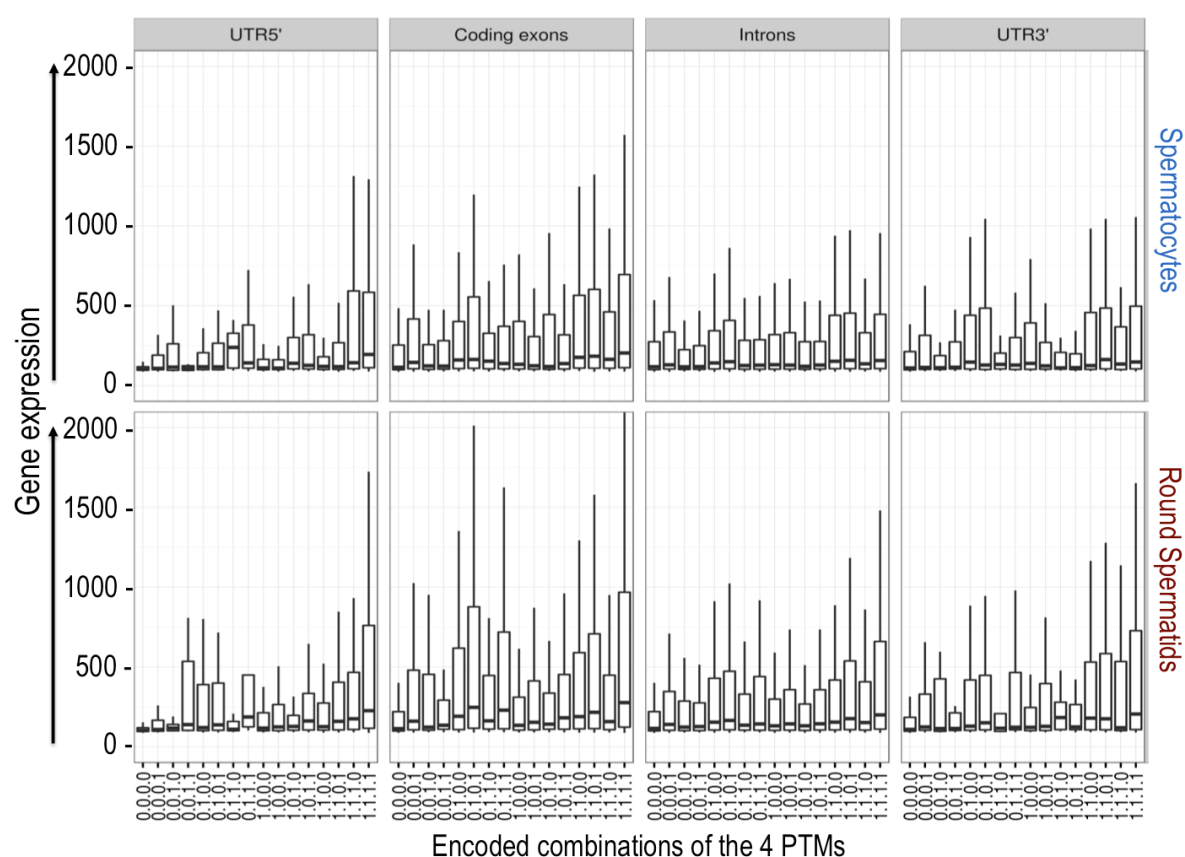


Figure S2. Relationship between genomic elements labelled with H4K5K8 acetylation / butyrylation and gene transcriptional activity

The transcriptional activities of genes whose annotated elements (5'UTR, coding exons, etc.) associated with nucleosomes bearing regions labelled with the indicated histone H4 modifications are shown as box plots in spermatocytes (upper panel) and round spermatids (lower panel). The absence or presence of a histone mark is represented by the numbers “0” and “1”, respectively, and the four histone marks are arranged from bottom to top as follows: H4K5ac, H4K5bu, H4K8ac, H4K8bu (as in Figure 4A). The transcription levels in spermatocytes and round spermatids were obtained using transcriptomic data of control samples from our previous work ((Montellier et al., 2013); GSE46136).

Figure S3 (related to Figure 6C)

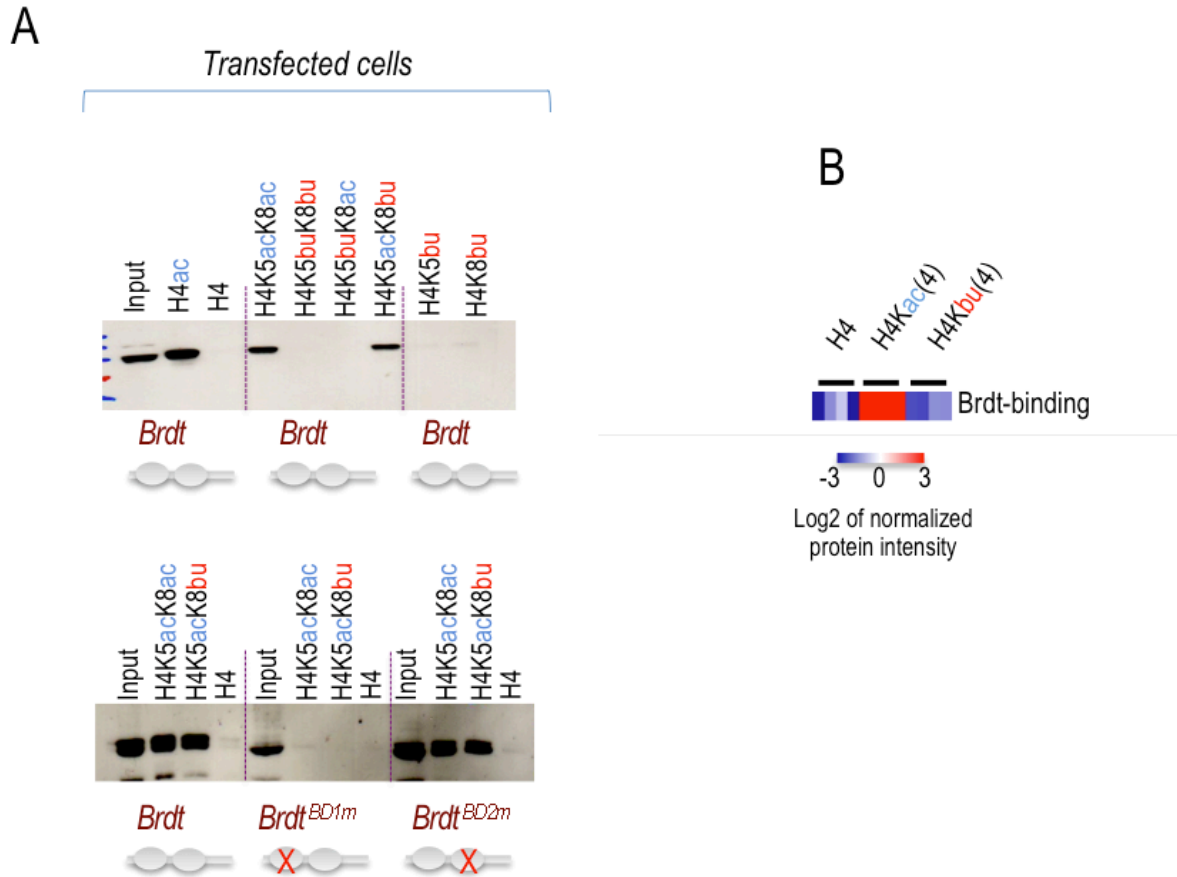


Figure S3. Butyrylation of H4K5 prevents Brdt binding to histone H4.

(A) Cos-7 cells were transfected by a GFP-Brdt construct with wild type bromodomains (deltaS construct described in (Moriniere et al., 2009)), and extracts were used to pull down Brdt with the indicated biotinylated peptides immobilized on streptavidin beads (upper panel). In another set of experiments, the same Brdt expression vector as above and expression vectors encoding Brdt with inactivating mutations in either the first ($\text{Brdt}^{\text{BD1mut}}$) or the second ($\text{Brdt}^{\text{BD2mut}}$) bromodomain were used and a pull-down was performed with the indicated peptides (lower panel). Brdt was visualized by Western blotting using an anti-GFP antibody. Input represents 5% of the extract used for each of the pull-down experiments.

(B) Wild-type mouse whole testis extracts were incubated with the indicated peptides as above but the presence of Brdt was detected by mass spectrometry. Annotated MS/MS spectra corresponding to Brdt pull-down from mouse testis extracts shown in Data File S1.

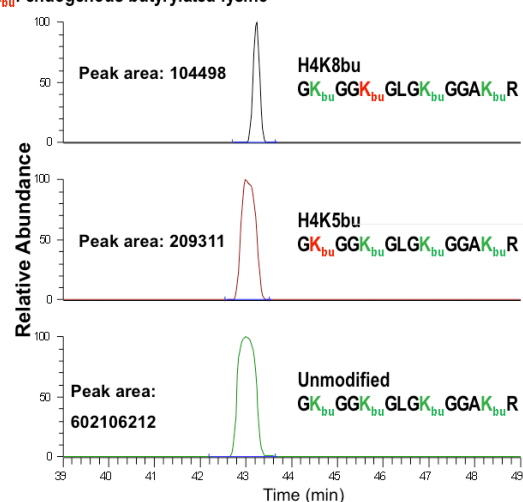
Figure S4 (related to Figure 7)

A (related to Figure 7B)

Ion chromatograms of Kbu-only peptides
in the total germ cells sample

K_{bu} : chemical labeled D₆-butyrylated lysine

K_{bu} : endogenous butyrylated lysine



$H4K5_{bu}/Unmodified = 0.0348\%$ (calculated by peak area)

$H4K8_{bu}/Unmodified = 0.0173\%$ (calculated by peak area)

B (related to Figure 7C)

Co-occurrence of H2A.Lap1 incorporation
and the four H4 PTMs at gene TSSs

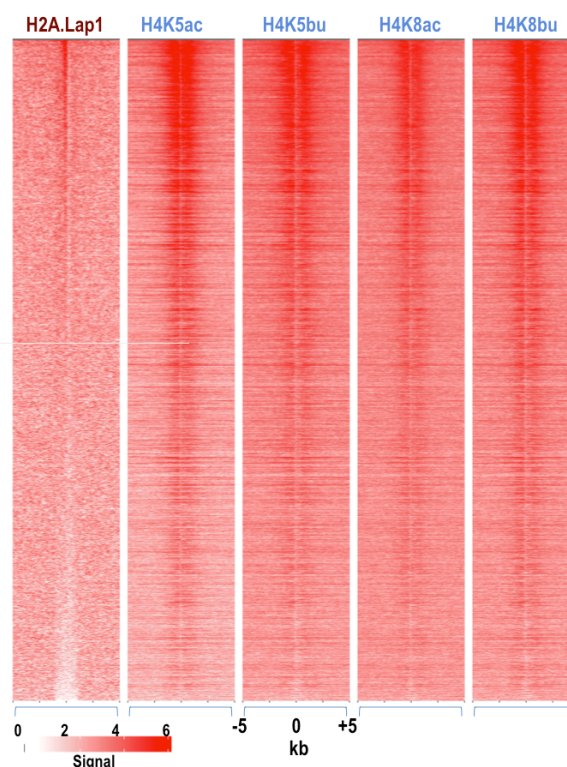


Figure S4. (A) Example of the quantification of H4K5bu and H4K8bu containing peptides in total germ cell (related to Figure 7B). Ion histograms of Kbu only H4 peptides with the indicated modifications as well as the corresponding peak area values are shown. These values were used to calculate the % of H4Kbu/unmodified H4 shown for both H4K5bu and H4K8bu. (B) TSSs with the highest level of labelling with H4K5K8 acetylation and butyrylation also show high levels of H2A.Lap1 incorporation. ChIP-seq data targeting H2A.Lap1 obtained from spermatogenic cells (Soboleva et al., 2012), were down-loaded and the presence of H2A.Lap1 in gene TSSs was compared with the presence of the four considered H4 PTMs at the same position (related to Figure 7C).

Table S1 (related to [Figure S1](#))

Table S1: Sequences of the primers used in ChIP-qPCR reactions (related to [Figure S1](#))

<i>Regions</i>	<i>Forward primer (-F)</i>	<i>Reverse primer (-R)</i>
osbp18_TSS1	CTTGCCACGTCGCTAAAACC	AGCCTAGCAGGCTTTCTTGT
osbp18_TSS2	CCTGTGAGGCCAACAGCTAA	TACCCGCAGGTCTCTGAAGT
Ahi1_TSS1	GCAGGGATGGTGACAGCTTA	CAGTCTCATGCGTGGAAGT
Ahi1_TSS2	CCTCCATGATTTTGTGTTTGTTCGT	CAATCTTCTCCCGGGTGCTC
chr12_nopeak2	GGTCACTGAACCGATGCCAA	CGTAGACATCACACAGCCTCT
mapkapk5_TSS1	TGCCGAAAGAGAAGCGGAAT	ACTTCCCAGCCAACTCCTTG
mybl1_TSS1	AATAATGATAGATGCTCTCAGGCCA	TCTGTCCTTTCTGCTGGTC
mybl1_TSS2	ACAAGGAAACTGGCCAAGACC	CACGCAAGCCTCGAGATAAC
rad51ap2_TSS1	TCAGAGGAAAGCATCGATTTTCG	AGCTGTTCTATGCGTCTCCG
rad51ap2_TSS2	GTTCTAGGGCTCCAAGAAGGT	GGCTATTGGGTGGAGCATGA

Table S2 (related to Figure 7B)

Table S2: Quantification of histone butyrylation from total or fractionated spermatogenic cells (related to Figure 7B).

Detected Peptides	MS Peak area			
	TGC	Speytes	R	E/C
GK_{bu}GGK_{bu}GLGK_{bu}GGAK_{bu}R	602106212	83898395	79250430	133044324
GK_{bu}GGK_{bu}GLGK_{bu}GGAK_{bu}R	209311	NA	NA	NA
GK_{bu}GGK_{bu}GLGK_{bu}GGAK_{ac}R	57705	NA	32936	203401
GK_{bu}GGK_{ac}GLGK_{bu}GGAK_{ac}R	90504	NA	39127	NA
GK_{bu}GGK_{bu}GLGK_{ac}GGAK_{ac}R	61294	NA	NA	NA
GK_{bu}GGK_{ac}GLGK_{ac}GGAK_{ac}R	135530	160064	229275	63231
GK_{bu}GGK_{bu}meGLGK_{bu}GGAK_{bu}R	NA	NA	NA	169186
GK_{bu}GGK_{bu}GLGK_{bu}GGAK_{bu}R	104498	NA	NA	NA
GK_{bu}GGK_{bu}GLGK_{bu}GGAK_{ac}R	312924	NA	49103	12656
GK_{ac}GGK_{bu}GLGK_{bu}GGAK_{ac}R	6079	NA	NA	NA
GK_{ac}GGK_{bu}GLGK_{ac}GGAK_{ac}R	17478	NA	47198	NA
GK_{bu}GGK_{bu}GLGK_{bu}GGAK_{ac}R	NA	269556	NA	NA

Legend:

TGC: Total germ cells; **Speytes**: spermatocytes; **R**: round spermatids; **E/C**: elongating/condensing spermatids
K_{bu}: chemically labelled D₅-butyrylated lysine;
K_{bu}me: endogenous methyllysine that is D₅-butyrylated;
K_{bu}: endogenous butyrylated lysine;
K_{ac}: endogenous acetylated lysine;
K_{me}: endogenous methylated lysine;

Legend to Data File S1

Data File S1 encompasses all the annotated MS/MS spectra

- Annotated MS/MS spectra of histone Kbu sites from different species related to [Figure 1B](#) (indicated in each spectrum).
- Annotated MS/MS spectra for histone Kbu sites from experiments shown in [Figure 5A](#) and [Figure 5B](#) (indicated in each spectrum).
- Annotated MS/MS spectra for Brdt protein after the testis-extract pull down shown in [Figure S3B](#) (indicated in each spectrum).
- Annotated MS/MS spectra for H4 with multiple acetylation/butyrylation shown in [Figure 7A](#) (indicated in each spectrum).

Extended Experimental Procedures

Kinetics of H4K5ac/bu and H4K8ac/bu antibodies binding to their corresponding antigen peptides.

The kinetic parameters for anti-H4K5ac, anti-H4K5bu, anti-H4K8ac, and anti-H4K8bu antibodies binding to their corresponding peptide antigens $\text{NH}_2\text{-SGRGK}_{\text{ac}}\text{GGK}_{\text{ac}}\text{GLGK}_{\text{ac}}\text{GGAK}_{\text{ac}}\text{RHRGGK}_{\text{biotin}}\text{-COOH}$ or $\text{NH}_2\text{-SGRGK}_{\text{bu}}\text{GGK}_{\text{bu}}\text{GLGK}_{\text{bu}}\text{GGAK}_{\text{bu}}\text{RHRGGK}_{\text{biotin}}\text{-COOH}$ were measured by Surface Plasmon Resonance (SPR) in a Bio-Rad ProteOn XPR36. The antibodies were injected at concentrations ranging from 1 to 25 nM at a constant flow rate of 30 $\mu\text{l}/\text{min}$ over immobilized acetylation or butyrylation containing peptides respectively. The association phase was monitored for 300sec and the dissociation phase was monitored for 3600sec to record the change in SPR signal. The data were fit to a “two-state” binding model. The residuals from the “two-state” binding model, indicated an excellent fit (not shown). The association rates (k_a) and dissociation rates (k_d) were measured and used to calculate the dissociation constant (K_D) of antibody-peptide complex formation as the ratio k_d/k_a .

Identification of Kbu sites in core histones by affinity enrichment and mass spectrometry

Two hundred micrograms of core histones of interest were isolated and dried in a SpeedVac. The proteins were suspended in 50 mM NH_4HCO_3 (pH 8.0). The proteins were digested with trypsin at an enzyme-to-substrate ratio of 1:50 for 16 hrs at 37°C. For the mouse testis samples related to [Figure 1B](#), chemical derivatization was performed before digestion. Briefly, the extracted histones were propionylated using propionic anhydride in 100 mM NH_4HCO_3 buffer (adjust to pH=8.0 after adding propionic anhydride) at 37 °C for 1h. The propionylation reaction was repeated, and then quenched by 100 mM ethonalamine at room temperature for 1h. No chemical derivatization was performed to the other samples. For quantifications of H4 butyrylation in mouse spermatogenic cells another protocol was followed, which is presented in an independent section (below). The tryptic digest was incubated with 5 μg of anti-Kbu antibody in CLN buffer (20 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.5% NP-40, pH 8.0) at 4°C for 6 hrs with gentle shaking. Protein A-agarose was then added to immunoprecipitate the antibody-antigen complexes. The bound peptides were washed twice with the CLN buffer, and twice with the CLN buffer without NP-40, and then eluted three times with 50 μl of 1% TFA in water. The eluates were combined and dried in a SpeedVac. The resulting peptides were cleaned in C18 Zip Tips (Millipore Co., Temecula, CA) according to the manufacturer’s instruction, prior to nano-HPLC/MS analysis.

The HPLC/MS/MS analysis was carried out by nano-HPLC/LTQ mass spectrometry as described previously ([Chen et al., 2007](#)). Briefly, the enriched peptides bearing Kbu was dissolved in 10 μl of HPLC buffer A (0.1% (v/v) formic acid in water), and 2 μl were injected into an Agilent nano-HPLC system (Agilent, Palo Alto, CA). Peptides were separated on a custom-made capillary HPLC column (50-mm length x 75- μm inner diameter) containing Jupiter C12 resin (4- μm particle size, 90-Å pore diameter, Phenomenex, St. Torrance, CA) and electrosprayed directly into the mass spectrometer using a nanospray source. The LTQ mass spectrometer was operated in a data-dependent mode, cycling between acquiring one MS spectrum and MS/MS spectra of the 10 strongest ions in that MS spectrum.

MS/MS of synthetic peptides was carried out using the same LTQ settings as those used for the analysis of *in vivo* peptides.

Quantification of H4K5K8 butyrylation in spermatogenic cell populations

To quantify the histone H4K5K8 butyrylation, an *in vitro* chemical labelling method was used followed by targeted MS detection. Briefly, the acid extracted histones from total or fractionated mouse germ cells were chemically butyrylated with deuterated (D5) butyryl anhydride (catalog number D-7421, C/D/N Isotopes Inc., Quebec, Canada), using a procedure previously reported by Becker laboratory (Feller et al., 2015). Under these conditions, all unmodified and mono-methylated lysine residues become butyrylated and can be distinguished from endogenous butyrylation in mass spectrometric analysis because of the mass difference between endogenous butyryl group and D5-labeled butyryl group. Since the chemical labelling disables the tryptic digestion at lysine residues, we only focussed on the H4K5K8-containing peptide sequence window ‘GKGGKGLGKGGAKR’.

To calculate the stoichiometry of H4K5bu and H4K8bu sites, a method developed by Garcia laboratory (Lin and Garcia, 2012) was used. Briefly, the peak areas of unmodified and all differentially modified forms of a histone peptide of interest are summed and each individual form of a histone peptide i. e., H4K5bu, is expressed as a fraction of the summed of unmodified peak areas. Considering the high structural similarity between lysine butyrylation and acetylation, diversely acylated peptides were assumed to have similar ionization efficiencies.

ChIP-seq and ChIP-qPCR, Chromatin preparation and immunoprecipitation

Male germ cells were fractionated into spermatocytes and round spermatids enriched cell suspensions from OF1 wild type mice (ChIP-seq) or 129svj/C56BL6 mixed background mice (ChIP-qPCR), and these male germ cell fractions were then used to prepare mononucleosomes as previously described (Dai et al., 2014). Each ChIP experiment was performed on 100 µg of mononucleosomes and the DNA corresponding to the immunoprecipitated nucleosomes, as well as before immunoprecipitation (input) were purified (as previously described in (Dai et al., 2014)).

For ChIP-qPCR, immunoprecipitated DNA fragments were amplified along with serial dilutions of input DNA using SYBR Green mix: “Brilliant III Ultra fast SYBR Green QPCR master Mix” (AGILENT) on a : CFX96 Real-Time System, C1000 Touch Thermal Cycler (BIO RAD). The genome positions of the primers used are indicated in Figure S1 and their sequences are reported in Table S1.

For the ChIP-seq, libraries were prepared following the Ovation Ultralow DR Multiplex System protocols (NuGEN Technologies INC, San Carlos, CA), with a size selection (200-700bp) step using SPRIselect beads (Beckman Coulter, Brea, CA). Single end sequencing was performed using Illumina HiSeq 2500. Base calls were performed using CASAVA version 1.8.2 (by sequencing facility). Total read counts of ChIP-seq experiments ranged from 24.2 to 38.7 Million reads.

Bioinformatic analyses

ChIP-seq reads were aligned to the mm9 genome assembly using bowtie version 2.2.5 with default arguments. Read counts were computed using Deeptools bamCoverage version 1.5.11 (binSize=50; fragmentLength=200; missingDataAsZero=yes). Seqminer (<http://sourceforge.net/projects/seqminer/>; (Ye et al., 2011)) was used to visualize aligned raw reads as heatmaps on TSS regions. Peaks were called using SICER (<http://home.gwu.edu/~wpeng/Software.htm>; (Zang et al., 2009)) with the following parameters: W200, G600, FDR 0.01. The IGB genome browser (<http://bioviz.org/igb/index.html>; (Nicol et al., 2009)) was used for visualizing peaks on particular gene regions. Raw data (FastQ files) and processed data (.bed and .wig outputs of Sicer peak caller) are available on GEO (<http://www.ncbi.nlm.nih.gov/geo/>, dataset no GSE77277).

p300 expression and purification for in vitro acetyl- or butyryl-transferase assays on histone octamer

Flag tagged human p300 (324-2094) was cloned into pFASTBAC1 vector (Invitrogen). We used EMBAC-Y baculovirus for the expression of flag p300. The recombinant baculoviruses were propagated in Sf21 insect cells with Sf-900 III SFM medium (Invitrogen). For the production of p300, Hi5 insect cells in Express Five SFM medium were infected with recombinant baculovirus at a multiplicity of infection of 2, maintained in shake flasks at 27 °C and harvested by centrifugation 72 h post infection. Recombinant baculovirus-infected cell pellet was resuspended in TNZ buffer (20 mM Tris pH 7.5, 300 mM NaCl, 5µM ZnCl₂, Complete Protease Inhibitors EDTA-free (Roche Applied Science)) and lysed by sonication. The lysate was then clarified by centrifugation at 100,000 g for 1h. The clarified insect cell lysate was incubated with pre-equilibrated Anti-FLAG M2-agarose affinity gel (Sigma) for 1 h and then applied to a 25-ml Econo column (Bio-Rad). The resin was washed with 200 ml of TNZ buffer, and then the bound protein was eluted using 0.1 mg/ml triple FLAG peptide (sequence, MDYKDHDADYKDHDIDYKDDDDK) dissolved in TNZ buffer. The eluate was concentrated using an Amicon-ultra 30 unit (Amicon, Inc.) and further purified on a Hi-Load 16/60 Superdex 200 gel filtration column (GE Healthcare) equilibrated in HNZ buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 5µM ZnCl₂, 0.5mM DTT). The final protein was concentrated to 5 mg/ml and aliquots were flash-frozen in liquid nitrogen and stored at -80 °C.

Untagged drosophila histones H2A, H2B, H3, H4 were cloned into pET21 vector (Novagen). We used IPTG-induced BL21 bacterial cells for expression of histones. Recombinant histones were purified and refolded according to standard procedures.

Standard reactions, with purified proteins, were performed in reaction buffer (25 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 10 % Glycerol, 1x Complete EDTA-free protease inhibitor (Roche)) with 100 ng/mL TSA, 2 µg of purified octamer and 50 µM Acetyl-CoA or 50 µM Butyryl-CoA or 25 µM Acetyl-CoA and 25 µM Butyryl-CoA (Sigma, A2181 and B1508). Reactions were incubated in presence or not of purified p300 during 1 h at 30 °C. Reactions were stopped by addition of Laemmli buffer and samples were used for immunoblotting or mass spectrometry.

Supplemental references

- Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C., Falck, J.R., Peng, J., Gu, W., and Zhao, Y. (2007). Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* 6, 812-819.
- Dai, L., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., Debernardi, A., Buchou, T., Rousseaux, S., Jin, F., *et al.* (2014). Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat Chem Biol* 10, 365-370.
- Feller, C., Forne, I., Imhof, A., and Becker, P.B. (2015). Global and specific responses of the histone acetylome to systematic perturbation. *Mol Cell* 57, 559-571.
- Lin, S., and Garcia, B.A. (2012). Examining histone posttranslational modification patterns by high-resolution mass spectrometry. *Methods Enzymol* 512, 3-28.
- Montellier, E., Boussouar, F., Rousseaux, S., Zhang, K., Buchou, T., Fenaille, F., Shiota, H., Debernardi, A., Hery, P., Curtet, S., *et al.* (2013). Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev* 27, 1680-1692.
- Moriniere, J., Rousseaux, S., Steuerwald, U., Soler-Lopez, M., Curtet, S., Vitte, A.L., Govin, J., Gaucher, J., Sadoul, K., Hart, D.J., *et al.* (2009). Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 461, 664-668.
- Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A., and Loraine, A.E. (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* 25, 2730-2731.
- Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J. (2012). A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat Struct Mol Biol* 19, 25-30.
- Ye, T., Krebs, A.R., Choukrallah, M.A., Keime, C., Plewniak, F., Davidson, I., and Tora, L. (2011). seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* 39, e35.
- Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* 25, 1952-1958.

3- p300-catalyzed histone acylations

The recent investigations reported that p300 in addition to acetylation, can also catalyse crotonylation, butyrylation, and propionylation ([Chen et al., 2007](#); [Cheng et al., 2009](#); [Sabari et al., 2015](#)). In a study initiated in the laboratory of our collaborator in EMBL, the crystal structure of p300-acyl-CoAs complexes bound to different acyl-CoAs such as butyryl-CoA, propionyl-CoA, and crotonyl-CoA was determined. These structures gave us important hints on the basis of p300-dependent histone acylations and helped us to better understand how a change in cell metabolism could drive a specific histone acylation.

Main results obtained

3-1- p300 discriminates between different acyl-CoAs

The work from Daniel Panne's group based on the analysis of crystal structure of p300 in complex with several acyl-CoAs, reveals how p300 accommodates different acyl-CoAs. This study reports p300 activity becomes weaker by increasing the length of acyl-chains. The restricted activity of histone acylation by p300 allowed us to show the determinants that are necessary to enhance the use of acyl groups by p300.

This work has important applications in the understanding of the regulation of histone acylations. It shows that p300 prefers Acetyl-CoA over other CoA derivatives with longer fatty acid chains. Therefore to efficiently acylate histones, p300 would need a factor that induces some structural transitions. This could be achieved by its partners or by yet undefined mechanisms.

The results of this study are published in Nature Chemical Biology in 2016. For my contribution in this work, please see author contributions in the paper.

Structure of p300 in complex with Acyl-CoA variants

Zuzanna Kaczmarek^{1,2}, Esther Ortega^{1,2}, Afsaneh Goudarzi³, He Huang⁴, Sunjoo Kim⁴, José A.

Márquez¹, Yingming Zhao⁴, Saadi Khojbin³, Daniel Panne^{*1}

¹European Molecular Biology Laboratory, 38042 Grenoble, France.

³CNRS UMR 5309, INSERM, U1209, Université Grenoble Alpes, Institut Albert Bonniot, 38700 Grenoble, France

⁴Ben May Department of Cancer Research, The University of Chicago, Chicago, IL 60637, USA

²Co-first author

*Corresponding author: email: panne@embl.fr, tel: +33476207925

Abstract

Histone acetylation plays an important role in transcriptional activation. Histones are also modified by chemically diverse acylations that are frequently deposited by p300, a transcriptional co-activator that uses a number of different acyl-CoA cofactors. Here we report that while p300 is a robust acetylase, its activity gets weaker with increasing acyl-CoA chain length. Crystal structures of p300 in complex with propionyl-, crotonyl-, or butyryl-CoA show that the aliphatic portions of these co-factors are bound in the lysine substrate-binding tunnel in a conformation that is incompatible with substrate transfer. Lysine substrate binding is predicted to remodel the acyl-CoA ligands into a conformation compatible with acyl-chain transfer. Remodeling requires that the aliphatic portion of acyl-CoA is accommodated in a hydrophobic pocket in the enzymes active site. The size of the pocket and its aliphatic nature restricts against long chain and charged acyl-CoA variants, presumably explaining the co-factor preference for p300.

Introduction

Spatial and temporal regulation of chromatin is an important aspect of all DNA templated processes in eukaryotic cells including gene transcription, DNA replication and repair. Chromatin function is regulated at several levels including incorporation of histone variants, remodeling by ATP-dependent remodeling enzymes and posttranslational modification (PTM) of histones. Enzymes that catalyze histone PTMs frequently use co-factors that are derived from primary metabolism, providing a link between the metabolic state of the cell and gene regulation¹. For example, chromatin acetylation, a modification usually associated with active gene transcription, is catalyzed by acetyltransferases (HATs), enzymes that transfer an acetyl group from the metabolite acetyl-CoA to the ϵ -amino group of lysine². Lysine acetylation provides binding sites for effector proteins containing bromodomains, a domain that is prevalent in nuclear proteins that interact with chromatin³. Histone deacetylases (HDACs) remove such chromatin modifications and the combined action of writing, reading and erasure of histone acetylation, and histone modifications in general, provides a mechanism for differential regulation of chromatin function and gene expression⁴.

In addition to acetylation, histone lysine residues can also be modified by chemically diverse acyl-modifications, including formylation, propionylation, butyrylation, 2-hydroxyisobutyrylation, β -hydroxybutyrylation, crotonylation, malonylation, succinylation, glutarylation, 4-pentynoyl-, 3-phosphoglyceroylation, palmitoylation and myristoylation⁵⁻¹⁵. These modifications are derived from their respective charged acyl-CoAs, small molecule intermediates derived from pathways of cellular energy metabolism. As the concentrations of various acyl-CoAs fluctuate as a function of the metabolic state of the cell, metabolism potentially directly influences cellular signaling and transcriptional activation by modulating differential protein and histone acylation^{5,6,9,16,17}.

The identification of various histone lysine acylations raises the question of which enzymes mediate their deposition, how they are removed, and if there are binding domains that recognize such modifications. A recent investigation of the capacity of human bromodomains to bind acyllysine modifications showed that while most bromodomains analyzed are able to bind to acetyl- and propionyllysine, most bromodomains do not bind the longer butyryl- or crotonyllysine modification¹⁸. Thus, the longer acylations can potentially inhibit binding of bromodomain proteins as shown recently for histone H4K5 butyrylation which prevents binding of Brdt, with the result that histone removal is delayed in spermatogenic cells¹⁹. Recent data also support the notion that there are specialized 'reader' domains such as the YEATS domain that preferentially recognize the crotonyllysine modification^{20,21}.

Removal of long-chain lysine acylations is mediated by several HDACs: SIRT5 preferentially removes acidic acyl modifications such as malonyl^{12,22}, succinyl^{12,22,23} or glutaryl¹⁰. Short chain fatty acyl modifications such as propionyl, butyryl or crotonyl are removed by several sirtuins²⁴⁻²⁸ while long-chain fatty acyl modifications can be removed by SIRT6^{25,29}.

How enzymes catalyze diverse lysine acylations is less well understood. Based on the sequence divergence, HATs have been grouped into at least five different families³⁰. Three HAT family members have been reported to catalyze short-chain lysine acylation *in vitro*: p300/CBP (KAT3B and KAT3A nomenclature according to ref.³¹) catalyze propionylation, crotonylation and butyrylation on chromatin, p53 and p300 itself^{5,24,32}. p300-mediated crotonylation also occurs in cells and crotonylation levels are directly regulated by the cellular concentration of crotonyl-CoA³². Like acetylation, crotonylation and butyrylation by p300 also stimulate gene transcriptional activity in cell-free assays^{19,32}. Some data also suggest that p300 can catalyze acidic acyllysine modifications including glutarylation and succinylation^{10,33}. Other HATs such as PCAF and hGCN5 catalyze propionylation of a histone H3 peptide, and yeast Esa1, a MYST

family member, catalyzes propionylation of a histone H4 peptide³⁴. The activity of PCAF gets weaker with increasing acyl chain length³⁵. Recently it was shown that the deletion of *hat1*, *gcn5* and *rtt109* in yeast led to a considerable reduction of H3K9 crotonylation *in vivo*, with the most dramatic effects upon *gcn5* deletion suggesting that this modification is incorporated by yGCN5²⁰. Structural and biochemical data show however, that hGCN5 cannot hydrolyze longer acyl-CoA variants³⁶. While the active site of hGCN5 can bind propionyl-CoA and butyryl-CoA, the latter adopts a conformation that does not allow catalysis to occur indicating that the longer acyl-chain variants act as competitive inhibitors for hGCN5³⁶. Other HATs such as hMOF and hTIP60 also appear not to use the longer crotonyl-CoA as substrate³². Thus while most HATs are more restrictive with regards to acyl chain length, p300 can utilize a chemically diverse set of acyl-CoA cofactors.

To start to address how p300 can accommodate chemically and sterically diverse acyl-CoA cofactors in the active site in a conformation that is compatible with the binding of protein substrates and transfer of the acyl moiety, we have systematically analyzed p300-mediated histone H3-H4 acylation. In agreement with previously published data, we found that short chain acyl-CoAs variants can be used as p300 substrates but that the acyltransferase activity gets weaker with increasing acyl chain length. X-ray crystal structures of the p300 core in complex with propionyl-, crotonyl-, and butyryl-CoA revealed that the active site of p300 accommodates longer acyl chains without major structural rearrangements. The aliphatic portions of the acyl-CoA variants are positioned in the lysine substrate-binding tunnel and would sterically clash with the incoming substrate. We propose that substrate engagement results in remodeling of the short chain acyl-CoA ligands into a conformation compatible with acyl-chain transfer. We identify a hydrophobic pocket in the enzyme's active site that is predicted to accommodate the aliphatic portion of the remodeled acyl-CoA. The size of the pocket and its aliphatic nature apparently restricts against long chain and charged acyl-CoA variants, presumably explaining the co-factor

preference for p300. The lack of such a pocket in other HATs such as hGCN5, TIP60 and MOF explains the more restricted acyl-CoA cofactor requirement for these acetyltransferases.

Results

p300 can use a diverse array of acyl-CoAs as substrates

p300 activation is accompanied by trans-autoacetylation in a lysine-rich autoinhibitory loop spanning residues 1520–1581 (AIL) that is embedded in the HAT domain³⁷. To assess which acyl-CoA variants are compatible with p300 autoacylation, we used a trans-autoacetylation assay in which one catalytically deficient p300 variant (p300_core) is acylated by a wild-type version of the enzyme (p300s) and the reaction is monitored by liquid chromatography-mass spectrometry (LC-MS) (Fig. 1a–c). The p300_core variant used in these experiments contained an inactivating mutation of the catalytic Tyr residue (Y1467F), which reduces but does not completely abolish p300 activity³⁸. Accordingly, insect cell purified p300_core showed a series of peaks (+42Da) corresponding to autoacetylation of the enzyme (Fig. 1a). Upon incubation with p300s, we observed a series of +42Da adducts corresponding to acetylated p300, presumably located in the lysine-rich AIL. Similarly, we observed a series of adducts upon addition of butyryl-CoA or propionyl-CoA indicating that these acyl-CoA variants can be used by p300 in a trans-autoacylation reaction (Fig. 1b–c). Due to the high reactivity of these acyl-thiol compounds it is important to distinguish between enzyme-catalyzed and nonenzymatic acylation. As we only observed an increase in p300_core acylation in the presence of p300s but not when incubated only with acyl-CoA, we conclude that p300 can use these acyl-CoAs as acylation donors (Fig. 1a–c).

Next, we investigated whether p300 can also incorporate these acylations into protein substrates. We incubated p300s with histones H3 and H4 that were recombinantly produced in *E. coli* and

refolded to form H3-H4 tetramers and analyzed the reaction products by LC-MS. These *Xenopus* histones were already partially acetylated prior to incubation with p300s (Fig. 1d). In the presence of acetyl-CoA we observed significantly increased acetylation (Fig. 1d). In the presence of butyryl-, or propionyl-CoA, we observed additional mass peaks that corresponded to the respective acyl-CoA variant used (Fig. 1e–f). However, we observed fewer modified peaks as compared to the acetyl-CoA treated sample, indicating that the activity of p300 gets weaker with increasing acyl chain length. We also analyzed histone acylations using western blot. We found that p300 also supported crotonylation and β -hydroxybutyrylation, a modification shown recently to occur in chromatin ¹⁴, but incubation with 2-hydroxyisobutyryl-CoA resulted in nonenzymatic acylation, presumably due to high reactivity of this acyl-CoA variant (Supplementary Results, Supplementary Fig. S1a,b).

To further analyze the observed decrease in catalytic efficiency we compared the catalytic rates of hydrolysis of the acyl-CoA variants by p300 in the presence of a histone H3 peptide. We found that while p300 efficiently hydrolyzed acetyl- and propionyl-CoA, its butyrylation and crotonylation activity were less efficient (Fig. 1g). p300 propionylated the H3 peptide approximately 3-fold more slowly while it butyrylated about 40-fold and crotonylated nearly 62-fold more slowly as compared to its acetyltransferase activity (Fig. 1g). Thus, we observed a direct correlation between the length of the acyl-CoA chain and the efficiency of hydrolysis with the longer variants being less efficiently hydrolyzed by p300.

Structure determination of acyl-CoA bound p300

To elucidate the structural basis for the acyl-CoA chain preference of p300, we determined the structure of p300 bound to propionyl-, crotonyl- and butyryl-CoA. Previous structural studies of the catalytic ‘core’ required introduction of the inactivating Y1467F mutation and deletion of the flexible AIL that harbors multiple lysine acetylation sites (p300_core Δ , Supplementary Fig. S1c)

³⁹. Expression of p300_coreΔ in insect cells results in residual autoacetylation, and successful crystallization previously required further deacetylation with SIRT2. To facilitate structural studies we have established a co-expression system for p300_coreΔ and SIRT2. Co-expression in insect cells allowed us to express high yields of p300_coreΔ (Supplementary Fig. S2a) and size-exclusion chromatography showed that such preparations are homogeneous (Supplementary Fig. S2b). Mass spectrometry analysis showed that p300_coreΔ, purified in the absence of SIRT2, is autoacetylated as revealed by a series of peaks shifted by ~ +42 Da relative to the unacetylated p300_coreΔ (Supplementary Fig. S2 c). Upon co-expression of SIRT2, mostly the deacetylated form was obtained (Supplementary Fig. S2 c). Thus, p300 preferentially uses endogenous acetyl-CoA but not other acyl-CoA variants, at least under the fermentation conditions used upon overexpression in insect cells. This modified purification procedure allowed us to crystallize the p300_coreΔ bound to various ligands without the additional deacetylation step that was previously required ³⁹. All crystal structures were determined by molecular replacement using the p300 core structure (PDB code 4BHW) with the Lys-CoA ligand omitted from the search model. Data collection and refinement statistics for all structures are listed in Supplementary Table 1.

Structure of p300 bound to various acyl-CoA ligands

The overall structure of p300 in complex with the different CoA ligands is essentially identical to that seen in the p300/lysyl-CoA complex ³⁹. The structure encompasses the Bromodomain, the CH2 region containing a PHD domain that is interrupted by a RING domain, and the HAT domain (Fig. 2a). The Bromodomain, PHD, RING and HAT domains adopt an assembled configuration in which the RING domain is juxtaposed to the HAT substrate-binding site. The CoA binding pocket in the butyryl-CoA liganded structure has the same architecture as in the lysyl-CoA structure (Fig. 2b). The substrate binding loop between $\alpha 4$ and $\beta 5$ encapsulates the CoA ligand: Arg1462 and Lys1456 sandwich the adenosine ring, Arg1410 makes several critical

hydrogen bonds to phosphates and the pantetheine arm of CoA makes extensive interactions with the substrate binding loop (Fig. 2b).

The structures obtained with the lysyl-CoA bisubstrate inhibitor mimic a postreaction CoA-lysine intermediate while the structures of the isolated HAT domain in complex with acetyl-CoA and CoA represent pre- and post-reaction conformations, respectively³⁸⁻⁴⁰. We were able to crystallize p300_coreΔ in the absence of extraneously added acyl-CoA ligand. Simulated annealing omit maps contained clear density for a CoA ligand (Supplementary Fig. S3c). Mass spectrometry analysis confirmed the identity of the ligand as hydrolyzed CoA (Supplementary Fig S3a,b). Thus p300_coreΔ co-purified from insect cells with endogenous CoA but not with acetyl-CoA or other acyl-CoA variants. To understand how other acyl-CoA variants are accommodated, we crystallized p300 with propionyl-, butyryl- and crotonyl-CoA. Omit maps calculated in the absence of acyl-CoA clearly confirm the position of the ligands reported here (Fig. 2c–e) and of lysyl-CoA and acetyl-CoA reported previously (Fig. 2f–g)³⁸⁻⁴⁰. Previous structures containing the lysyl-CoA bisubstrate inhibitor show that the substrate lysine invades the active site through a hydrophobic tunnel comprised of Tyr1397, Trp1436, Tyr1446 and Cys1438 that interact with the aliphatic portion of the lysine side chain (Fig. 3a). The propionyl-CoA co-crystal structure showed that the aliphatic propionyl moiety is positioned in the lysine substrate-binding tunnel (Fig. 3b). As in the lysyl-CoA inhibitor, the carbonyl oxygen engages in a hydrogen bond with the backbone amide of Leu1398 and the sulfur atom is directed towards Tyr1467, the proposed general acid to protonate the sulfur after the acyl transfer reaction³⁸. In the butyryl-CoA and crotonyl-CoA structures, the aliphatic acyl chains are similarly positioned in the lysine substrate-binding tunnel (Fig. 3c, d). Due to the presence of a more rigid trans-unsaturated Cα-Cβ bond, the crotonyl moiety adopts a more elongated conformation as compared to the butyryl chain (Fig. 3d). Superposition with lysyl-CoA shows that the configuration of the

propionyl-, butyryl- or crotonyl- moieties is mutually exclusive with lysine substrate binding (Fig. 3e).

p300 uses a direct-transfer mechanism in which acyl-CoA binds first to the HAT domain followed by substrate lysine⁴¹. As the acyl-moieties were not in a conformation compatible with substrate binding, we hypothesized that lysine substrate engagement must displace the acyl-group from the substrate-binding tunnel to get access to the carbonyl carbon of acyl-CoA. Such a reaction mechanism would require that the displaced acyl-chain be accommodated in the enzyme active site. Inspection of the CoA binding pocket shows that the ligand snakes through a narrow tunnel that opens to the substrate-binding site where the lysine moiety of lysyl-CoA binds. This tunnel contains a cavity bounded by the hydrophobic residues Ile1395, Leu1398, Leu1418 and Ile1435. In the acetyl-CoA bound structure, the methyl moiety of the acetyl group is oriented towards this hydrophobic pocket (Fig. 3f)⁴⁰. Reorientation from the inhibitory position in the substrate binding tunnel into the cavity would require a $\sim 180^\circ$ rotation of the carbonyl such that the oxygen hydrogen bonds to a conserved water molecule that is coordinated by the backbone amide of Trp1436 and carbonyl of Ile1395. In this configuration the aliphatic portion of the acyl-CoA ligand would be positioned in the hydrophobic cavity and the carbonyl in an orientation suitable for nucleophilic attack of the deprotonated primary amine of substrate lysine on the scissile bond. A model of this configuration, which likely represents the conformation just before the reaction occurs, is shown in Fig. 3g.

Conservation of the active site and mutagenesis

HATs have been grouped into five different families, based on the sequence similarity within the HAT domain³⁰. Each family contains a conserved core region including a 3-stranded β -sheet with a parallel α -helix spanning one side of the sheet (Fig. 4a). This core region makes conserved interactions with the acyl-CoA cofactor. Surrounding the core are α/β segments that are

structurally divergent between different family members. A molecular surface rendering of the acyl-CoA binding pocket of p300 shows a cavity that is in close proximity to the methyl group of acetyl-CoA (Fig. 4a). Analysis of this cavity revealed that it extends into the interior of the protein (Fig. 4b). hGCN5, which has a more restricted acyl-CoA substrate specificity³⁶, does not contain such a cavity (Fig. 4 c,d). We therefore predicted that p300 accommodates extended acyl-CoA ligands due to the presence of this aliphatic pocket in its active site. Structural comparison reveals that a pocket at an equivalent position is used in enzymes that catalyze transfer of longer acyl chains including N-Myristoyl transferase (NMT) and *N*-acyl-L-homoserine lactone (acyl-HSL) synthases. NMT, a member of the GCN5 family, catalyzes attachment of myristic acid (*n*-Tetradecanoic acid) onto the N terminus of a number of proteins. The myristoyl group of CoA is accommodated in an extended aliphatic back pocket that is located at an equivalent position of the p300 pocket (Fig. 4e, f). Acyl-HSL synthases are structurally related to acetyltransferases and are involved in quorum sensing in gram-negative bacteria^{42,43}. Acyl-HSL synthases produce *N*-acyl-L-homoserine lactones from the substrates *S*-adenosyl-L-methionine (SAM) and acylated acyl carrier protein (acyl-ACP)⁴³. The acyl-chain, the length of which can vary from C4–C18, is accommodated in an extended aliphatic pocket, the size of which appears to dictate substrate specificity⁴³. Structural superposition shows that the extended pocket of Acyl-HSL synthases is at an equivalent position as that predicted to accommodate the acyl-moiety in p300 (Supplementary Fig. S4a,b). The acyl-chain binding pocket of NMT, TofI and EsaI are significantly larger and more extended as compared to p300, presumably explaining why p300 uses short-chain but not longer acyl-CoA variants. The small size of the pocket could explain why hydrolysis of crotonyl- and butyryl-CoA is less efficient as compared to propionyl- and acetyl-CoA. The lack of such a pocket in other HATs such as hGCN5, TIP60 and MOF likely explains the more restricted acyl-CoA cofactor requirement for these acyltransferases.

The hydrophobic back pocket of p300 is bounded by partially solvent accessible residues

including Met1376, Ile1395, Leu1398, Leu1418 and Ile1435. Ile1395 and Ile1435 are located at the entrance of the pocket and are nearly invariant among HAT enzymes. To assess the relevance of these residues for acyl-chain transfer, we mutated Ile1395 and Ile1435 to Methionine. N-terminally HA-tagged p300 variants were transiently expressed in COS cells and the impact on activity measured using an IP-HAT assay³⁹. The I1395M, I1435M or the double I1395M/I1435M mutants revealed similar acetylation activity of histones H3-H4 tetramers as compared to wild-type p300 showing that these mutations do not affect the overall structure and acetylation activity (Fig. 5a). We observed however, that butyrylation activity was increased for either the I1395M or the double mutant I1395M/I1435M but not for I1435M (Fig. 5a). Introduction of a more bulky residue at this position, I1395F, resulted in a p300 variant that retained acetylation activity but showed diminished butyrylation (Fig. 5b). For crotonylation activity, we also observed an increase with the I1395M mutant while I1395F resulted in decreased signal (Fig. 5b). In contrast, for propionylation we did not observe strong effects by mutations in the cavity (Fig. 5b). To more quantitatively characterize the impact on histone acylation, we analyzed the acetylation and butyrylation reaction products using a spectral counting method by MS (Fig. 5c, d and Supplementary note)⁴⁴. MS analysis of the acetylation reaction showed that the I1395M mutant showed decreased acetylation at all positions except H4K12 (Fig. 5 c). In contrast, we observed a modest increase for histone H3 and H4 butyrylation in the samples treated with p300 I1395M (Fig. 5 c,d). While at several lysine sites only a modest increase was observed, their cumulative effect could increase the binding affinity of pan-specific anti-butyryl-lysine antibodies and possibly explain the increased signal in the western blot. Together these results indicate that Ile1395 plays a potential role as a ‘gatekeeper’ residue that controls access of longer acyl-CoA variants into the hydrophobic back pocket of p300. A more flexible Met amino acid residue at this position results in enhanced butyrylation activity while a more bulky and rigid Phe residue interferes with butyrylation activity, presumably by restricting access of the longer acyl-CoA variants to this back pocket.

Proposed reaction mechanism

Extensive structural, biochemical, mutational, and enzymatic analyses have provided insights into the catalytic mechanism of p300^{38,41,45}. p300 utilizes a ‘Theorell-Chance’ mechanism, a special type of sequential mechanism in which the ternary complex between enzyme, acetyl-CoA and substrate is short-lived. The sequential reaction mechanism is ordered with acetyl-CoA binding first to p300 before histone substrates bind. Two highly conserved residues, Tyr1467 and Trp1436, play a key role: Tyr1467 acts as a general acid and Trp1436 helps orient the target lysine residue of the protein substrate into the active site³⁸. While there is no clear general base for catalysis, recent simulations suggest that Tyr1394 and Asp1507 are involved in a proton transfer reaction which allows deprotonation of the target substrate lysine⁴⁶. Acetyl-CoA is bound in a conformation that allows direct nucleophilic attack on the carbonyl carbon. However for the longer acyl-CoA variants, their aliphatic portions are accommodated in the substrate-binding tunnel in a conformation that is incompatible with acyl-chain transfer (Supplementary Fig. 5a). We propose that substrate engagement results in displacement of the acyl chain into the hydrophobic cavity made up of Met 1376, Ile1395, Leu1398, Leu1418 and Ile1435 and positioning of the lysine amide near the carbonyl carbon atom (Supplementary Fig. 5b). The deprotonated lysine -NH₂ is predicted to attack the carbonyl carbon of acyl-CoA, while Y1467 acts as a general acid to protonate the thiolate anion of CoA (Supplementary Fig. 5c). In the final reaction, the tetrahedral intermediate is resolved and the acyl-lysine-containing product leaves, followed by CoASH (Supplementary Fig. 5d). Presumably due to the requirement of substrate-assisted rearrangement of the acyl-CoA chain and the restricted size of the back pocket, the rate of transfer catalyzed by p300 is more efficient with short acyl-CoA chains.

Discussion

Products of cellular metabolism are frequently essential co-factors for chromatin-modifying enzymes and emerging evidence suggests that metabolism can directly influence chromatin modification and gene expression. Recently a variety of chemically different lysine acylations, derived from their respective acyl-CoA metabolites, have been identified in chromatin⁴⁷. Due to the presence of a reactive thioester group, acyl-CoA metabolites are intrinsically reactive and therefore one major question is if the observed lysine acylations arise nonenzymatically or if there are dedicated enzymes that catalyze their incorporation into chromatin. We have sought to address this issue by studying p300, a well-known histone acetyltransferase, which has been shown to utilize a number of chemically diverse acyl-CoA co-factors to modify histone lysine residues. In addition to lysine acetylation, p300 catalyzes lysine propionylation, crotonylation, butyrylation and, as we show here, β -hydroxybutyrylation^{5,24,32}. Like acetylation, crotonylation, butyrylation and β -hydroxybutyrylation are enriched on active gene promoters thus suggesting that differential histone acylation functionally connects acyl-CoA metabolism and transcriptional regulation^{14,19,32}.

We have determined crystal structures of p300 bound to propionyl-CoA, crotonyl-CoA and butyryl-CoA and show that the extended aliphatic chains are accommodated in the substrate binding tunnel in a conformation that is incompatible with lysine substrate binding. While it has been clearly established that p300 can catalyze lysine crotonylation and butyrylation, we now show that p300 distinguishes between the different acyl-CoA molecules and that the enzymes' reactivity gets weaker with increasing acyl-chain length. We propose that the decrease in catalytic efficiency is due to a substrate-assisted rearrangement of the acyl-CoA chain and due to the restricted size of an aliphatic back pocket, which we predict to transiently accommodate the rearranged acyl chain. In agreement with this model, replacement of I1395 in the back pocket with a more flexible Met residue lead to a gain-of-function phenotype while introduction of a

more bulky and rigid Phe residue resulted in a loss-of-function phenotype for butyryl-CoA. Structural analysis of hGCN5 shows that, like p300, the butyryl-CoA C3 and C4 carbons occupy the lysine substrate-binding tunnel³⁶. However, hGCN5 does not have an extended aliphatic back pocket to accommodate the acyl-chain and thus for hGCN5 butyryl-CoA is a competitive inhibitor³⁶. Other HATs such as, hMOF and hTIP60 also lack a more extended back pocket, which presumably explains why these enzymes are more restrictive with regards to acyl-chain length. Notably, mutations designed to increase the size of the back pocket in hGcn5 and hMOF also allow these enzymes to accommodate larger size acyl groups⁴⁸.

As shown previously by Sabari et al.³², the concentration of cellular crotonyl-CoA influences transcriptional regulation by modulating differential p300-mediated acylation³². In such cell culture experiments, acetyl-CoA can be ~1,000 fold more abundant than crotonyl-CoA, and it remains unclear if the physiological concentrations of different acyl-CoA variants are as dynamic³². So far few studies have attempted to quantify the relative abundance of different acyl-CoA variants *in vivo*. Measurement of the different acyl-CoAs in the livers of fasting rats showed a relative molar ratio of 40:23:10:1 of acetyl-, propionyl-, butyryl- and crotonyl-CoA⁴⁹. Other measurements have shown a molar ratio of 9:1 and 20:1 of acetyl-, and propionyl-CoA in the livers of normal and fastened rats, respectively⁵⁰. Together these data show consistently higher acetyl-CoA levels under different metabolic regimes. Combined with the lower rate of hydrolysis of the longer acyl-CoA variants by p300, this likely explains why acetylation in chromatin is more abundant as compared to other acylations. However it is possible that the longer histone acylations are enriched on sites of high histone acylation turnover, such as gene transcriptional start sites (TSSs) and thus compete with histone acetylation^{19,32}. We also consider the possibility that other classes of enzymes with more efficient longer-chain acyl transferase activity are involved in these TSS histone acylations.

Accession codes

Atomic coordinates and structure factors of the reported crystal structure have been deposited in the Protein Data Bank under the accession codes: 5LKU (endogenous CoA); 5LKT (butyryl-CoA); 5LKZ (crotonyl-CoA); 5LKX (propionyl-CoA).

Acknowledgments

We thank the staff of ESRF and EMBL-Grenoble for assistance and support in using beamlines ID29, 23-1 and 23-2 and the HTX facility. We thank Luca Signor (IBS, Grenoble) and Prasad Phapale (EMBL) for mass spectrometry analysis. Correspondence and requests for materials should be addressed to D.P. This work used the platforms of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). Z.K. was supported by a fellowship from the EMBL Interdisciplinary Postdoc Programme under Marie Skłodowska-Curie actions Cofund (grant agreement number 291772). This work was supported by the ANR Grant ‘Episperm3’ (ANR-15-CE12- 0005-02) to S. K. and D. P and the Worldwide Cancer Research foundation (grant #16-0280) to D.P. Work in S. K. laboratory is supported by INCa, Fondation pour la Recherche Médicale and Fondation ARC. Work in the Y.Z. laboratory is supported by NIH grants GM105933, DK107868 and GM115961.

Author Contributions

Z. K. performed the structure determination of p300 under supervision of J.A.M. and D.P. E.O. and Z.K. performed biochemical assays. E.O. prepared the constructs and performed protein expression and purification. A.G. performed the IP-HAT experiments under the supervision of S.Kh. H.H. and S.Kj. performed mass spectroscopy experiments under the supervision of Y. Z. D.P. designed and coordinated the project, advised and assisted on all aspects of the project and wrote the manuscript. All authors discussed the results and commented on the manuscript.

392

393 **Competing Financial Interests Statement**

394 No competing financial interests are declared.

395

396 **References**

- 397 1. Janke, R., Dodson, A.E. & Rine, J. Metabolism and epigenetics. *Annu Rev Cell Dev Biol*
398 **31**, 473-96 (2015).
- 399 2. Roth, S.Y., Denu, J.M. & Allis, C.D. Histone acetyltransferases. *Annu Rev Biochem* **70**,
400 81-120 (2001).
- 401 3. Dhalluin, C. et al. Structure and ligand of a histone acetyltransferase bromodomain.
402 *Nature* **399**, 491-6 (1999).
- 403 4. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature* **403**,
404 41-5 (2000).
- 405 5. Chen, Y. et al. Lysine propionylation and butyrylation are novel post-translational
406 modifications in histones. *Mol Cell Proteomics* **6**, 812-9 (2007).
- 407 6. Tan, M. et al. Identification of 67 histone marks and histone lysine crotonylation as a new
408 type of histone modification. *Cell* **146**, 1016-28 (2011).
- 409 7. Yang, Y.Y., Grammel, M. & Hang, H.C. Identification of lysine acetyltransferase p300
410 substrates using 4-pentynoyl-coenzyme A and bioorthogonal proteomics. *Bioorg Med*
411 *Chem Lett* **21**, 4976-9 (2011).
- 412 8. Xie, Z. et al. Lysine succinylation and lysine malonylation in histones. *Mol Cell*
413 *Proteomics* **11**, 100-7 (2012).
- 414 9. Dai, L. et al. Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark.
415 *Nat Chem Biol* **10**, 365-70 (2014).
- 416 10. Tan, M. et al. Lysine glutarylation is a protein posttranslational modification regulated by
417 SIRT5. *Cell Metab* **19**, 605-17 (2014).
- 418 11. Zhang, Z. et al. Identification of lysine succinylation as a new post-translational
419 modification. *Nat Chem Biol* **7**, 58-63 (2011).
- 420 12. Peng, C. et al. The first identification of lysine malonylation substrates and its regulatory
421 enzyme. *Mol Cell Proteomics* **10**, M111 012658 (2011).
- 422 13. Moellering, R.E. & Cravatt, B.F. Functional lysine modification by an intrinsically
423 reactive primary glycolytic metabolite. *Science* **341**, 549-53 (2013).

- 424 14. Xie, Z. et al. Metabolic Regulation of Gene Expression by Histone Lysine beta-
425 Hydroxybutyrylation. *Mol Cell* **62**, 194-206 (2016).
- 426 15. Huang, H., Lin, S., Garcia, B.A. & Zhao, Y. Quantitative proteomic analysis of histone
427 modifications. *Chem Rev* **115**, 2376-418 (2015).
- 428 16. Cai, L., Sutter, B.M., Li, B. & Tu, B.P. Acetyl-CoA induces cell growth and proliferation
429 by promoting the acetylation of histones at growth genes. *Mol Cell* **42**, 426-37 (2011).
- 430 17. Lin, H., Su, X. & He, B. Protein lysine acylation and cysteine succination by
431 intermediates of energy metabolism. *ACS Chem Biol* **7**, 947-60 (2012).
- 432 18. Flynn, E.M. et al. A Subset of Human Bromodomains Recognizes Butyryllysine and
433 Crotonyllysine Histone Peptide Modifications. *Structure* **23**, 1801-14 (2015).
- 434 19. Goudarzi, A. et al. Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation
435 Are Hallmarks of Highly Active Gene Promoters. *Mol Cell* **62**, 169-80 (2016).
- 436 20. Andrews, F.H. et al. The Taf14 YEATS domain is a reader of histone crotonylation. *Nat*
437 *Chem Biol* (2016).
- 438 21. Li, Y. et al. Molecular Coupling of Histone Crotonylation and Active Transcription by
439 AF9 YEATS Domain. *Mol Cell* **62**, 181-93 (2016).
- 440 22. Du, J. et al. Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase.
441 *Science* **334**, 806-9 (2011).
- 442 23. Park, J. et al. SIRT5-mediated lysine desuccinylation impacts diverse metabolic
443 pathways. *Mol Cell* **50**, 919-30 (2013).
- 444 24. Cheng, Z. et al. Molecular characterization of propionyllysines in non-histone proteins.
445 *Mol Cell Proteomics* **8**, 45-52 (2009).
- 446 25. Feldman, J.L., Baeza, J. & Denu, J.M. Activation of the protein deacetylase SIRT6 by
447 long-chain fatty acids and widespread deacylation by mammalian sirtuins. *J Biol Chem*
448 **288**, 31350-6 (2013).
- 449 26. Bao, X. et al. Identification of 'erasers' for lysine crotonylated histone marks using a
450 chemical proteomics approach. *Elife* **3**(2014).
- 451 27. Garrity, J., Gardner, J.G., Hawse, W., Wolberger, C. & Escalante-Semerena, J.C. N-
452 lysine propionylation controls the activity of propionyl-CoA synthetase. *J Biol Chem*
453 **282**, 30239-45 (2007).
- 454 28. Smith, B.C. & Denu, J.M. Acetyl-lysine analog peptides as mechanistic probes of protein
455 deacetylases. *J Biol Chem* **282**, 37256-65 (2007).
- 456 29. Jiang, H. et al. SIRT6 regulates TNF-alpha secretion through hydrolysis of long-chain
457 fatty acyl lysine. *Nature* **496**, 110-3 (2013).
- 458 30. Yuan, H. & Marmorstein, R. Histone acetyltransferases: Rising ancient counterparts to
459 protein kinases. *Biopolymers* **99**, 98-111 (2013).

460 31. Allis, C.D. et al. New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-6
461 (2007).

462 32. Sabari, B.R. et al. Intracellular Crotonyl-CoA Stimulates Transcription through p300-
463 Catalyzed Histone Crotonylation. *Mol Cell* (2015).

464 33. Hu, A., Britton, L.M. & Garcia, B.A. Investigating the specificity of histone
465 acetyltransferase activity for producing rare modifications on histones using mass
466 spectrometry. *The 62nd Annual American Society for Mass Spectrometry Conference on*
467 *Mass Spectrometry and Allied Topics, Baltimore* (2014).

468 34. Berndsen, C.E., Albaugh, B.N., Tan, S. & Denu, J.M. Catalytic mechanism of a MYST
469 family histone acetyltransferase. *Biochemistry* **46**, 623-9 (2007).

470 35. Leemhuis, H., Packman, L.C., Nightingale, K.P. & Hollfelder, F. The human histone
471 acetyltransferase P/CAF is a promiscuous histone propionyltransferase. *Chembiochem* **9**,
472 499-503 (2008).

473 36. Ringel, A.E. & Wolberger, C. Structural basis for acyl group discrimination by human
474 Gcn5L2. *bioRxiv* (2016).

475 37. Thompson, P.R. et al. Regulation of the p300 HAT domain via a novel activation loop.
476 *Nat Struct Mol Biol* **11**, 308-15 (2004).

477 38. Liu, X. et al. The structural basis of protein acetylation by the p300/CBP transcriptional
478 coactivator. *Nature* **451**, 846-50 (2008).

479 39. Delvecchio, M., Gaucher, J., Aguilar-Guerrero, C., Ortega, E. & Panne, D. Structure of
480 the p300 catalytic core and implications for chromatin targeting and HAT regulation.
481 *Nature Structural and Molecular Biology* **21**(2013).

482 40. Maksimoska, J., Segura-Pena, D., Cole, P.A. & Marmorstein, R. Structure of the p300
483 histone acetyltransferase bound to acetyl-coenzyme A and its analogues. *Biochemistry*
484 **53**, 3415-22 (2014).

485 41. Thompson, P.R., Kurooka, H., Nakatani, Y. & Cole, P.A. Transcriptional coactivator
486 protein p300. Kinetic characterization of its histone acetyltransferase activity. *J Biol*
487 *Chem* **276**, 33721-9 (2001).

488 42. Gould, T.A., Schweizer, H.P. & Churchill, M.E. Structure of the *Pseudomonas*
489 *aeruginosa* acyl-homoserinelactone synthase LasI. *Mol Microbiol* **53**, 1135-46 (2004).

490 43. Watson, W.T., Minogue, T.D., Val, D.L., von Bodman, S.B. & Churchill, M.E. Structural
491 basis and specificity of acyl-homoserine lactone signal production in bacterial quorum
492 sensing. *Mol Cell* **9**, 685-94 (2002).

493 44. Carvalho, P.C., Hewel, J., Barbosa, V.C. & Yates, J.R., 3rd. Identifying differences in
494 protein expression levels by spectral counting and feature selection. *Genet Mol Res* **7**,
495 342-56 (2008).

- 496 45. Karukurichi, K.R. & Cole, P.A. Probing the reaction coordinate of the p300/CBP histone
497 acetyltransferase with bisubstrate analogs. *Bioorg Chem* **39**, 42-7 (2011).
- 498 46. Zhang, X. et al. Catalytic mechanism of histone acetyltransferase p300: from the proton
499 transfer to acetylation reaction. *J Phys Chem B* **118**, 2009-19 (2014).
- 500 47. Rousseaux, S. & Khochbin, S. Histone Acylation beyond Acetylation: Terra Incognita in
501 Chromatin Biology. *Cell J* **17**, 1-6 (2015).
- 502 48. Yang, C. et al. Labeling lysine acetyltransferase substrates with engineered enzymes and
503 functionalized cofactor surrogates. *J Am Chem Soc* **135**, 7791-4 (2013).
- 504 49. King, M.T. & Reiss, P.D. Separation and measurement of short-chain coenzyme-A
505 compounds in rat liver by reversed-phase high-performance liquid chromatography. *Anal*
506 *Biochem* **146**, 173-9 (1985).
- 507 50. Hosokawa, Y., Shimomura, Y., Harris, R.A. & Ozawa, T. Determination of short-chain
508 acyl-coenzyme A esters by high-performance liquid chromatography. *Anal Biochem* **153**,
509 45-9 (1986).

510

511

Figure 1. Acyltransferase activity of p300. (a–f) Mass spectrometry analysis using electrospray ionization (ESI) of the p300_core or histone H3 before (black) and after acylation. (a) p300 acetylation (red), (b) p300 butyrylation (blue), (c) p300 propionylation (green), (d) H3 acetylation (red), (e) H3 butyrylation, (f) H3 propionylation. The molecular mass and the number of acylation events are indicated. (g) Acyltransferase activity of p300s. The rates of catalysis by p300s were measured using the indicated acyl-CoA cofactors and N-terminal histone H3 peptide. The concentrations of acyl-CoA and substrate peptide H3 (20-mer) were 250 μ M and 100 μ M, respectively. p300s was at a concentration of 0.2 μ M. Experiments were done in triplicate and the error bars represent the standard deviation.

Figure 2. Structure of p300 in complex with acyl-CoA variants. (a) Ribbon representations of the p300 core structure. The Bromodomain (Bd), RING- and PHD- domains are shown in yellow, green and red, respectively. The N- and C- subdomains of the HAT domain are shown in blue and grey, respectively. The position of the deleted autoinhibitory loop (AIL) is indicated with an arrow. Butyryl-CoA is shown in stick representation. (b) The inset displays some of the critical components of the HAT active site: The substrate binding loop; the proton donor amino acid residue Y1467 (mutated to F in our structure); Amino acid residue W1436 helps to orient the substrate lysine while residues K1456 and R1462 are making contacts to the Adenosine ring of CoA. Amino acid residues in the aliphatic cavity are indicated: I1435, L1398, I1395. (c) propionyl-CoA. (d) butyryl-CoA. (e) crotonyl-CoA. (f) lysyl-CoA (PDB code 4BHW). (g) acetyl-CoA (PDB code 4PZS). acyl-CoAs are shown in stick representation. Difference density omit maps contoured at 1.5σ are shown.

Figure 3. Comparison of the structure of p300 in complex with different acyl-CoA ligands. (a) lysyl-CoA (PDB code 4BHW). (b) propionyl-CoA. (c) butyryl-CoA. (d) crotonyl-CoA. (e) Superposition of lysyl-CoA and butyryl-CoA structures. (f) acetyl-CoA (PDB code 4PZS). (g) Model of a substrate lysine displacing the butyryl chain into the aliphatic pocket.

Figure 4. Comparison of p300 with GNAT acyltransferases. The structurally conserved core regions of the HAT domains are colored in blue and the flanking N- and C- terminal regions are colored in grey. The inset displays the van der Waals surface, colored in grey, around the acetyl group of acetyl-CoA (a) p300: The non-conserved loop L1 is in red and the cofactor is shown as yellow sticks. (b) The acetyl-CoA binding cavity of p300 is shown as a transparent grey surface.

(c) hGCN5 (PDB code 1Z4R). (d) The acetyl-CoA binding cavity of hGCN5 is shown as a transparent grey surface. (e) NMT: N-myristoyl transferase (PDB code 1IID). (f) The myristate binding pocket is shown as a transparent surface.

Figure 5. Analysis of the impact of p300 mutagenesis on the histone acylation profile. *In vitro* reconstituted histone H3-H4 tetramers were incubated with the indicated acyl-CoA variant in the presence or absence of the indicated full-length p300 variant. (a) Top, loading control of hemagglutinin (HA)-tagged p300, determined by immunoblotting with anti-HA antibody. Middle, levels of acylation determined by immunoblotting with pan anti-acetylsine or pan anti-butyrylsine antibodies. Bottom, loading control of histone H3, determined by immunoblotting with anti-H3. Reactions in the left panel were incubated with acetyl-CoA and reactions in the right panel with butyryl-CoA. Ctrl, control in the absence of p300; WT, wild type. Uncropped images are available in Supplementary Fig. 6 (b) As in panel (a) but for additional p300 mutants and including crotonylation and propionylation. Uncropped images are available in Supplementary Fig. 7 (c) The histograms represent the spectral counts of peptides derived from histone H4 from acetylation reactions with p300 wt (red) or mutant I1395M (green). Background acetylation is shown in dark blue. Spectral counts of peptides derived from butyrylation are p300 wt (light blue) or mutant I1395M (orange). Background butyrylation is shown in magenta. The lack of signal indicates that no non-enzymatic butyrylation was detected. (d) Acetylated and butyrylated peptides derived from histone H3. Color code is as in panel (c). Each sample, was analyzed 3 times and the error bars represent the standard deviation of the individual measurements.

Online Methods

Constructs

The p300_core sequence encoding residues 1048–1664 was cloned into the RsrII and XhoI sites of pFASTBAC DUET (Invitrogen) vector as a GST-fusion protein with a cleavable TEV sequence. The sequence contained a Y1467F mutation and a flexible SGGSG amino acid linker replacing the region 1520–1581. A cDNA encoding residues 38–389 of human SIRT2 was cloned into the NotI and NcoI sites of the pFASTBAC DUET_p300 vector to obtain the co-expression vector. FLAG-tagged p300s encoding residues 324–2094 was cloned into the RsrII and HindIII sites of pFASTBAC1. Point mutations were introduced by QuikChange mutagenesis (Agilent) and all constructs confirmed by DNA sequencing. p300s was found to contain a T1935A mutation which was already present in the PCR template.

Expression and Purification

Baculoviruses were maintained in Sf21 insect cells with Sf-900 III SFM medium (Invitrogen). Deacetylated p300 ‘core’ was produced in Hi5 (Invitrogen) insect cells in Express Five SFM medium as described previously³⁹. Flag tagged human p300 (324–2094) was purified as shown previously¹⁹.

Histone Acylation assays

Recombinant histones were purified and refolded according to standard procedures⁵¹. Standard reactions, with 1 µg purified p300, were performed in reaction buffer (25 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM DTT, 0.1 mM EDTA, 10 % Glycerol, 1x Complete EDTA-free protease inhibitor (Roche)) with 100 ng/mL TSA, 2 µg of purified H3-H4 and 50 µM acyl-CoA. Reactions were incubated in presence or not of purified p300 during 1 h at 30 °C. Background nonenzymatic histone acylation was determined under identical conditions but p300 was omitted

from the reaction. Reactions were stopped by addition of SDS-PAGE loading buffer followed by immunoblotting. Alternatively, reaction kinetics were measured by using the DTNB assay. Previous data indicate a K_m of $\sim 1.2 \mu\text{M}$ for acetyl-CoA and $\sim 1\text{--}10 \mu\text{M}$ for histone peptides⁴¹. We chose saturating conditions for both acyl-CoA ($250 \mu\text{M}$) and H3 (20-mer) histone substrate peptide ($100 \mu\text{M}$) assuming that the K_M for acyl-CoA and histone substrate binding are not significantly altered in reactions containing longer acyl-CoA variants. $0.2 \mu\text{M}$ of p300s was incubated in $180 \mu\text{l}$ of 100 mM sodium phosphate pH 6.8, 300 mM NaCl, $5 \mu\text{M}$ zinc chloride with acyl-CoA variants and H3 histone substrate peptide. In agreement with previously published data, HAT activity was in the linear range for 30 minutes at 30°C and the acetyl-CoA concentration was saturating. The reaction was stopped by addition of $30 \mu\text{l}$ quench buffer (6 M guanidinium hydrochloride, 100 mM Sodium Phosphate dibasic, pH 6.8). To measure CoASH production, $30 \mu\text{l}$ of 6 mM DTNB in 100 mM Sodium Phosphate, pH 6.8 and 10 mM EDTA were added and the absorbance at 412 nm recorded. Assays were repeated three times. Background absorbance was subtracted from the absorbance determined for individual reactions. CoASH concentrations were obtained from a standard curve done using serial dilutions of CoASH (Sigma). The different acyl-CoAs used were: acetyl-CoA (Sigma A2056), propionyl-CoA (Sigma P5397), crotonyl-CoA (Sigma 28007), butyryl-CoA (Sigma 20785).

Mammalian cell culture

COS7 cells were grown at 37°C under $5\% \text{ CO}_2$ atmosphere in DMEM medium supplemented with 10% FBS, 2% Glutamine and 1% Penicillin/Streptomycin (Invitrogen). Cells were grown to $\sim 60\text{--}70\%$ confluence on the day of transfection. Transfections were done with Lipofectamin 2000 (Invitrogen), using $1 \mu\text{l}$ of Lipofectamine for $1 \mu\text{g}$ vector. Cells were treated with 100 ng/ml TSA (Trichostatin A) from SIGMA (T8552) for $3\text{--}4 \text{ hrs}$ before harvest.

Cell extracts, Immunoprecipitation and Western blots

After 24 hrs of transfection, COS7 cells were harvested and extracts prepared in LSDB 500 buffer (50 mM HEPES pH 8.0, 3 mM MgCl₂, 500 mM KCl, 20% Glycerol, 0.1% NP-40, 1 mM DTT, 100 ng/ml TSA, Complete EDTA-free Protease inhibitor, Roche). After centrifugation, the cell extracts were diluted to 250 mM KCl. For immunoprecipitation, cell extracts were incubated with anti-HA Rat monoclonal antibody (Roche, clone 3F10) for 2 h at 4 °C with rotation and then incubated with magnetic protein G beads (Invitrogen) overnight at 4 °C. After a thorough wash with LSDB 250 mM buffer (50 mM HEPES pH 8, 3 mM MgCl₂, 250 mM KCl, 20% Glycerol, 0.1% NP-40, 1mM DTT, 100 ng/ml TSA, Complete EDTA-free Protease inhibitor Roche), beads were used for HAT assays. For immunoblotting, proteins were run on a 4–12 % Bis-Tris SDS-PAGE (NuPAGE Precast gel, Invitrogen) followed by transfer onto a nitrocellulose membrane (Hybond C+, GE). Membranes were blocked with 5% skim milk in PBST (PBS, 0.1% Tween-20) or 3% BSA in TBST buffer (TBS: Tris base, NaCl pH 7.6, 0.1% Tween-20) then probed with an anti-KAc (lysine acetylation) rabbit polyclonal antibody (Cell Signaling, #9441), anti-HA mouse monoclonal antibody (Covance, MMS101R), anti-pan K-bu (lysine butyrylation, PTM Biolabs), anti-pan K-cr (lysine crotonylation, PTM Biolabs) or anti-pan K-pr (lysine propionylation, PTM Biolabs), respectively. All antibody dilutions were done as recommended by the manufacturers. Membranes were washed in PBST or TBST buffer before and after incubation with anti-rabbit or anti mouse HRP-conjugated secondary antibody (GE Healthcare, NA934 or NA931), and protein bands were visualized on film after the ECL reaction (GE Healthcare).

Kac/Kbu quantification

Histones were cut from the gel and washed with 50% ethanol (overnight) and water (20 min, twice). After washing, the gel bands were cut into small pieces and dehydrated by acetonitrile for 5 min. Then the gel pieces were washed twice with 100 mM NH₄HCO₃ and dehydrated by acetonitrile for 5 min followed by overnight trypsin (Promega) digestion at 37°C in 100 mM

NH₄HCO₃ buffer. Tryptic peptides were extracted with 50% acetonitrile/5% TFA, 75% acetonitrile/0.1% TFA, and acetonitrile. The solution was vacuum concentrated and the peptides were re-suspended in 0.1% TFA followed by desalting using C18-Ziptips. Tryptic peptides were loaded onto a custom-made capillary column (10 cm length, 75 µm internal diameter) packed with Jupiter C12 resin (4 µm particle size, 90 Å pore size, Phenomenex Inc.) on a nanoLC-1D plus HPLC system (Eksigent Technologies LLC, Dublin, CA). Peptides were eluted with a gradient of 5% to 90% HPLC buffer B (0.1% formic acid in 90% acetonitrile, v/v) in buffer A (0.1% formic acid in water, v/v) over 76 min at a flow rate of 200 nL/min. The eluted peptides were ionized and introduced into a LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific) using a nanospray source. Full MS scans were acquired over the range m/z 300–1800 with a mass resolution of 60000 at m/z 400. MS/MS was operated with a top-20 data-dependent method. For all experiments, the dynamic exclusion time was set to 5s. Database searching was performed with Mascot 2.1.0 against the Uniprot *Xenopus* proteins database. The search parameters included: protein N-terminal acetylation, methionine oxidation, lysine acetylation, lysine butyrylation, lysine mono-/di-/tri-methylation, and arginine mono-/di-methylation as variable modifications; two max missed cleavage sites; precursor ion mass tolerance 20 ppm; MS/MS tolerance 0.5 Da. All the identified peptides were manually verified.

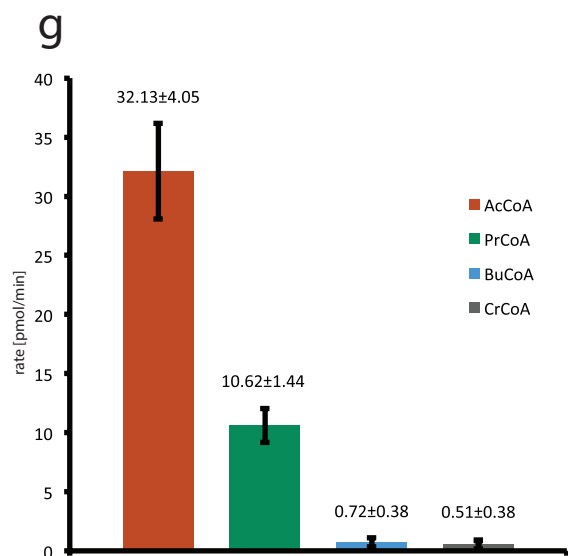
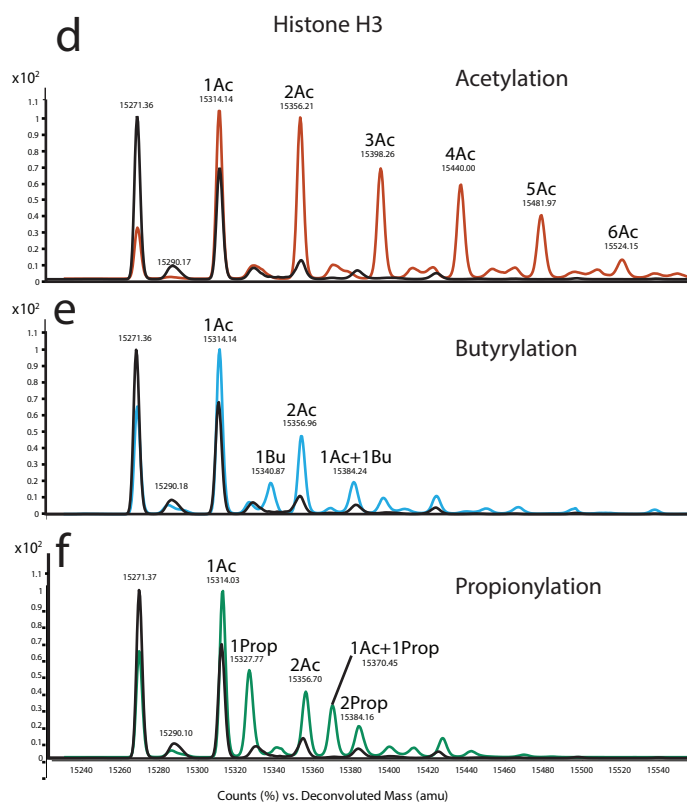
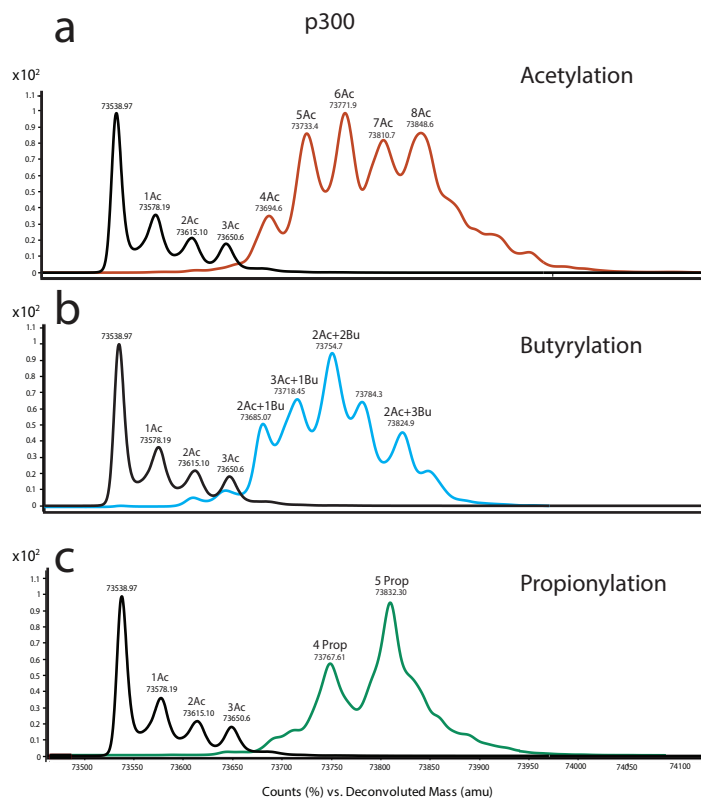
Crystallization and structure determination

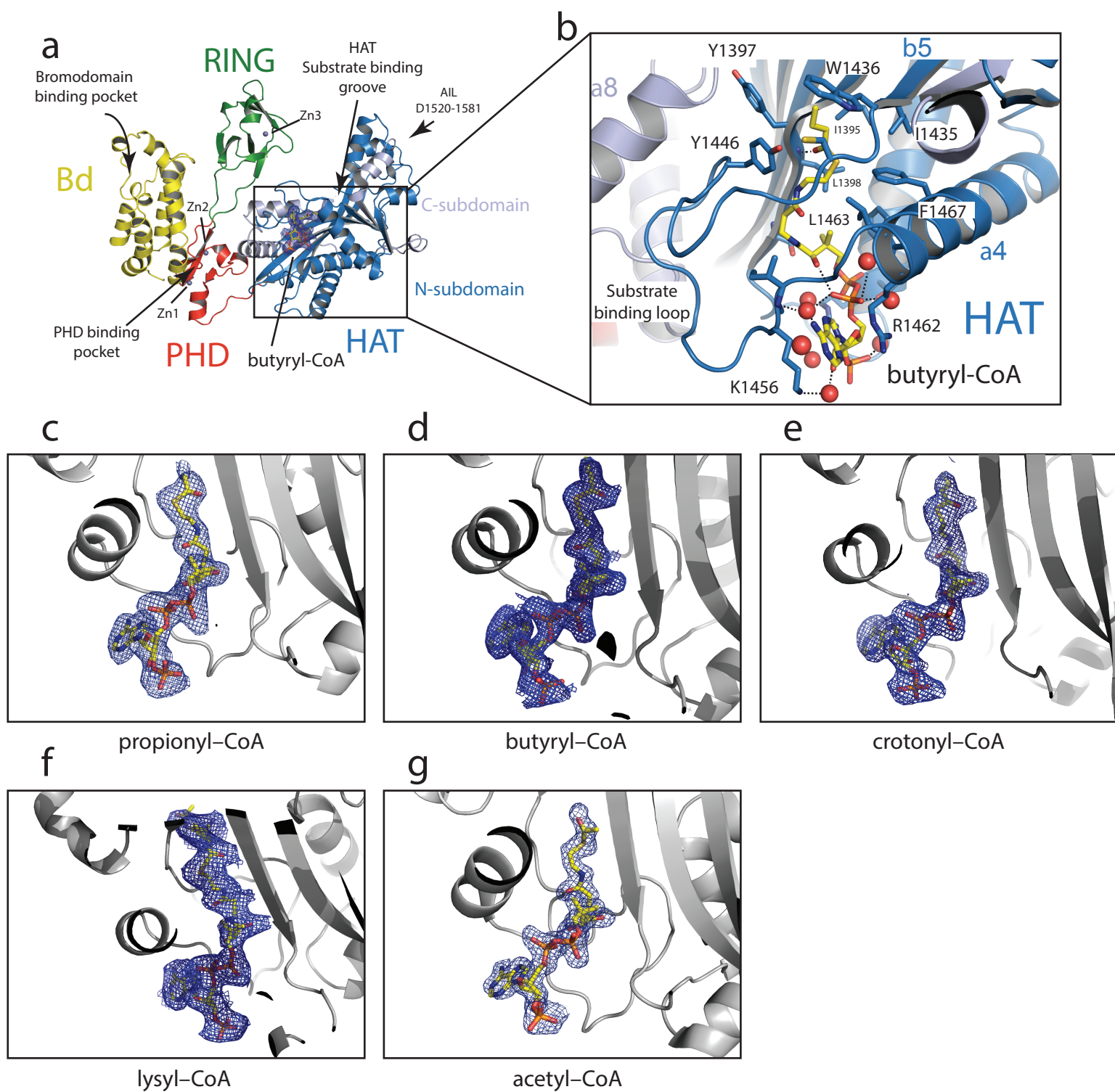
For crystallization, the deacetylated p300 ‘core’ at 4.5 mg ml⁻¹ was mixed with a 3-fold molar excess of the acyl-CoA as described previously³⁹. Crystals in the C222₁ space group were grown by sitting-drop vapour diffusion at 4 °C after mixing equal volumes of protein and crystallization solution containing 100 mM Tris, pH 7.5–8.0, 12–17.5% polyethylene glycol MME (monoethyl ether) 2000. Crystals that contain propionyl-CoA were obtained in the presence of 2% DMSO. Crystals were cryoprotected in 20–25% glycerol and frozen in liquid nitrogen. We collected native diffraction data at the ESRF on beamlines ID29 and ID23-2 under a nitrogen gas stream at

100 K, at a wavelength of 0.9763 and 0.8726 Å, respectively. All data were processed with XDS (Supplementary Table 1) and the structures determined by molecular replacement with the model for the p300 core (PDB entry 4BHW). The final models contain residues 1046–1664 with a deletion of residues 1517–1581. Analysis of the refined structure by MolProbity showed that there are no residues in disallowed regions of the Ramachandran plot (Supplementary Table 1).

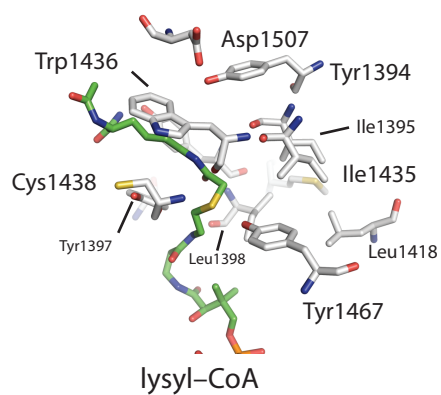
Online method references

51. Luger, K., Rechsteiner, T.J. & Richmond, T.J. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* **304**, 3-19 (1999).

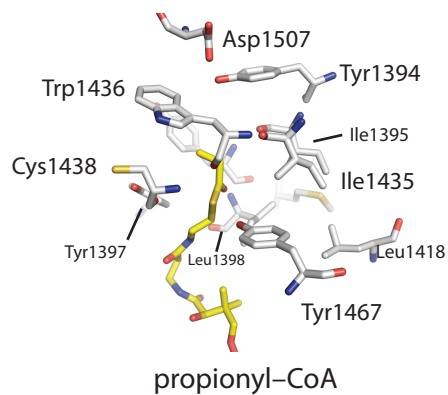




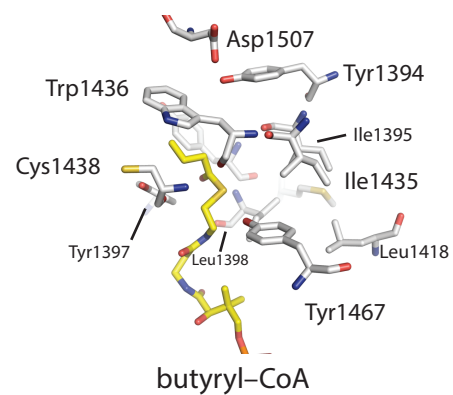
a



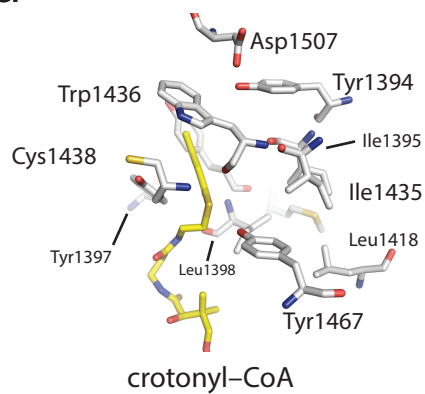
b



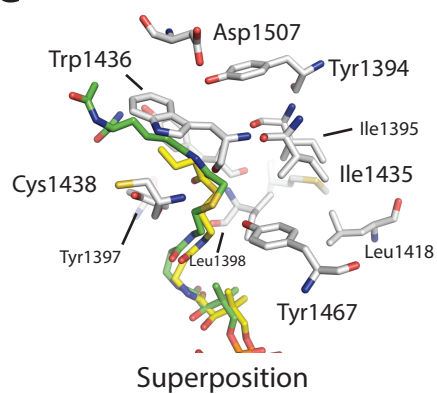
c



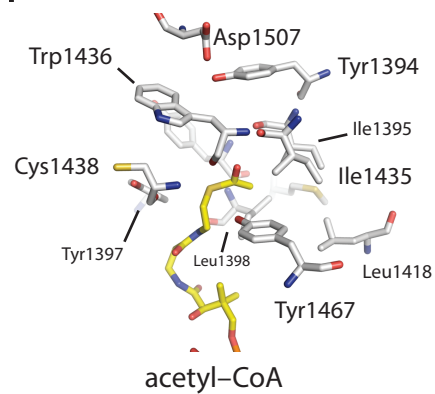
d



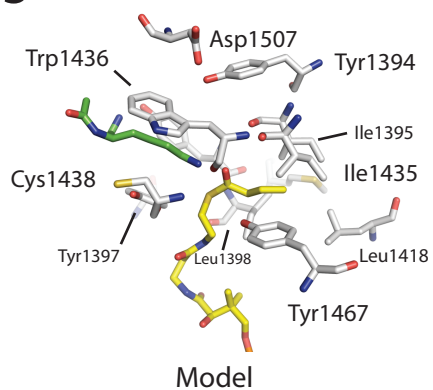
e

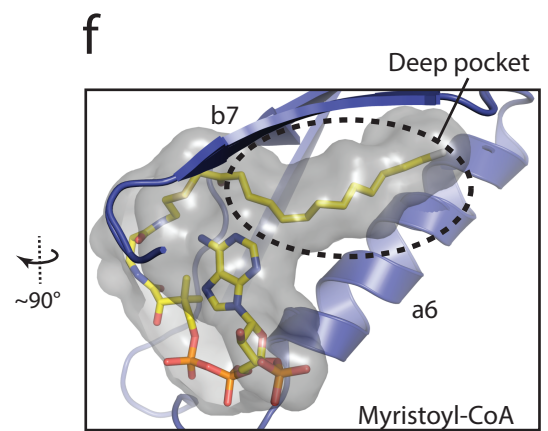
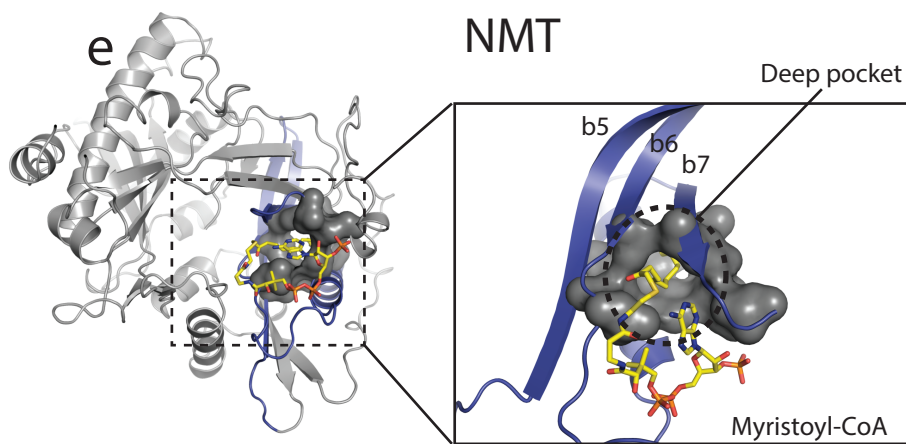
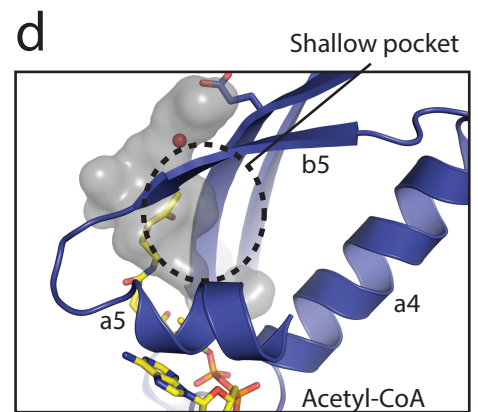
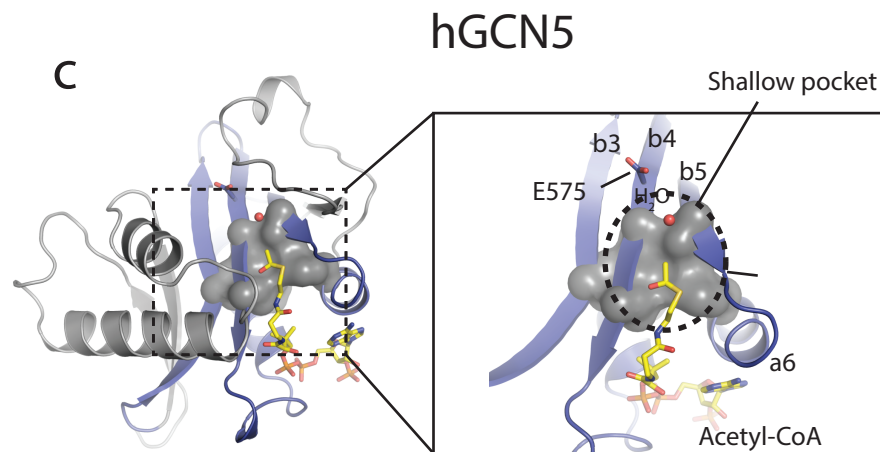
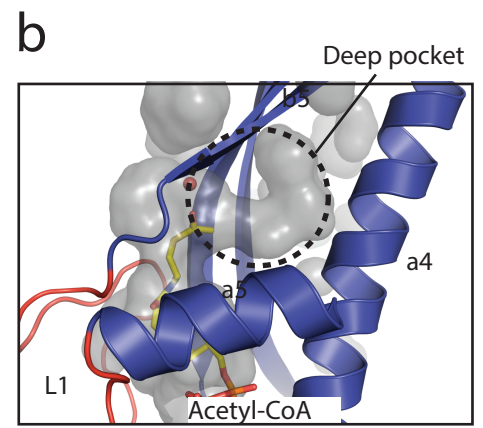
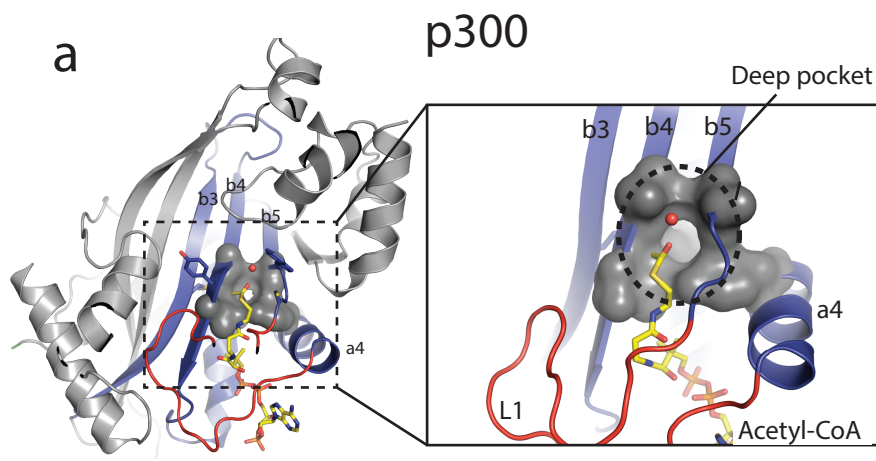


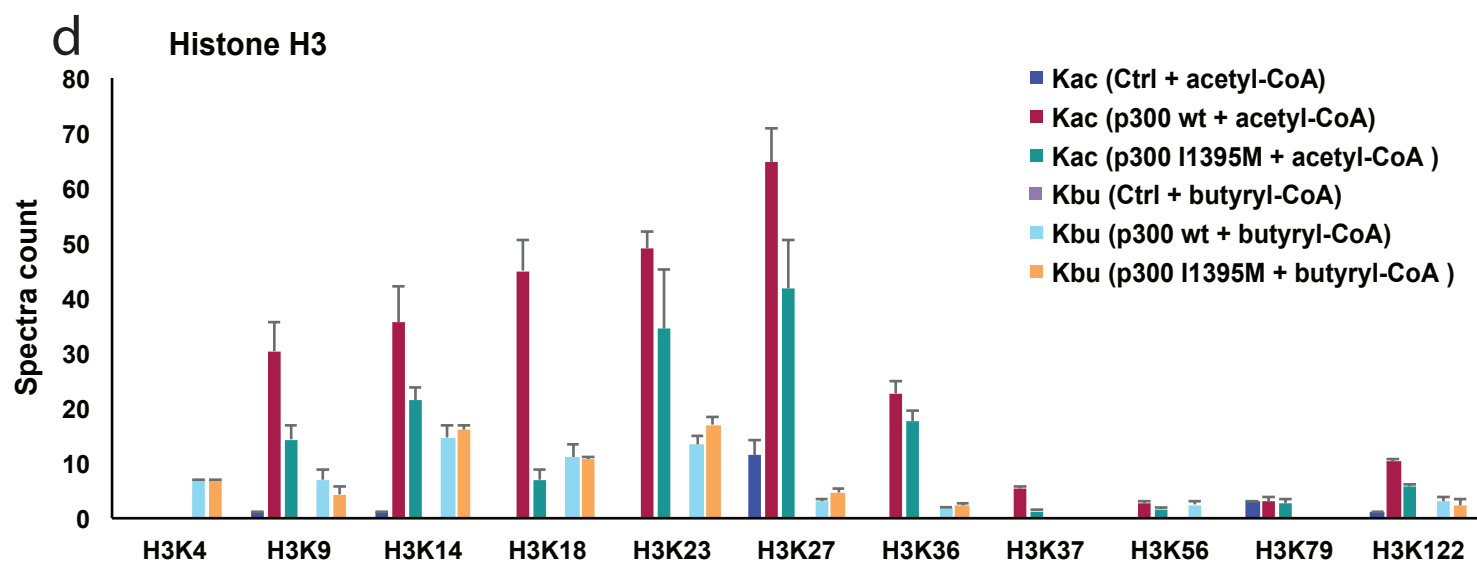
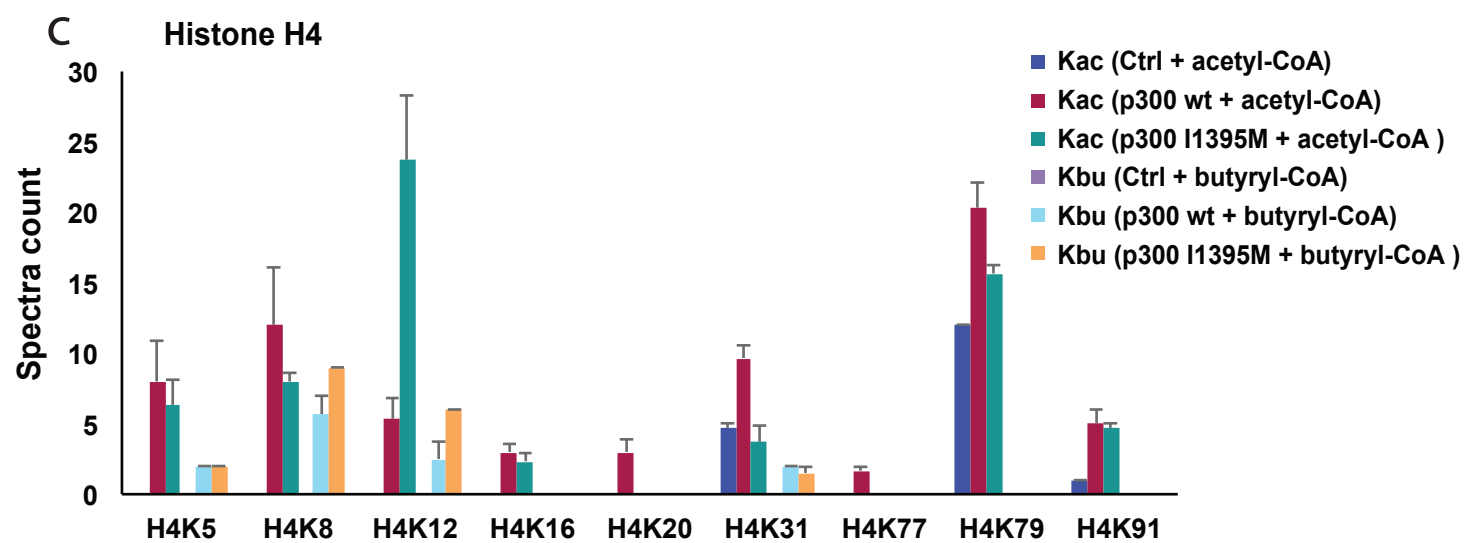
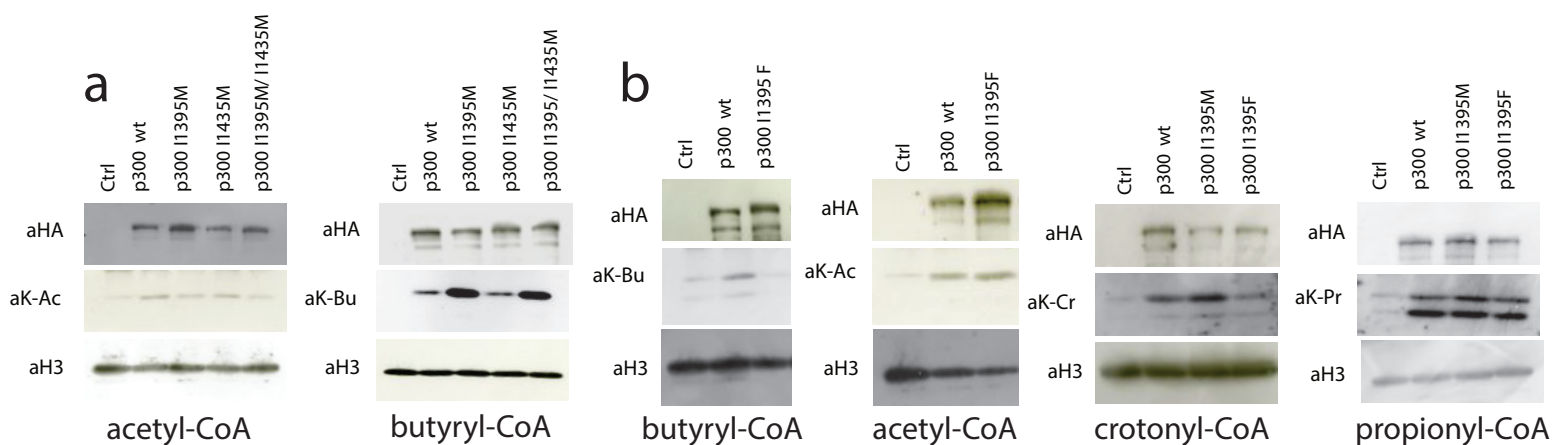
f



g







1
2
3
4
5
6
7
8
9
10
11
12
13
14

Supplementary Information

Structure of p300 in complex with Acyl-CoA variants

Zuzanna Kaczmarska^{1,2}, Esther Ortega^{1,2}, Afsaneh Goudarzi³, He Huang⁴, Sunjoo Kim⁴, José A. Márquez¹, Yingming Zhao⁴, Saadi Khochbin³, Daniel Panne^{*1}

¹European Molecular Biology Laboratory, 38042 Grenoble, France.
³CNRS UMR 5309, INSERM, U1209, Université Grenoble Alpes, Institut Albert Bonniot, 38700 Grenoble, France
⁴Ben May Department of Cancer Research, The University of Chicago, Chicago, IL 60637, USA

15 **Supplementary Results**
16

17 **Supplementary Tables**18 **Supplementary Table 1:** Data collection, phasing and refinement statistics

	endogenous CoA	butyryl-CoA	crotonyl-CoA	propionyl- CoA
Data collection				
Space group	C222 ₁	C222 ₁	C222 ₁	C222 ₁
Cell dimensions <i>a, b, c</i> (Å)	93.96, 156.40, 110.68	92.48, 154.69, 109.23	93.60, 155.91, 110.09	93.81, 155.26, 109.79
Wavelength	0.8726	0.9763	0.8726	0.8726
Resolution range (Å)	50.00-3.50 (3.59-3.50)	50.00-2.04 (2.09-2.04)	30.00-2.50 (2.56-2.50)	30.00-2.52 (2.54-2.52)
<i>R</i> _{meas}	38.3 (112.8) *	9.8 (99.9) *	16.5 (104.7) *	15.0 (110.5) *
<i>I</i> / σ <i>I</i>	5.65 (2.29) *	12.65 (2.00) *	9.93 (1.81) *	10.49 (1.48) *
CC (1/2)	97.8 (78.8) *	99.8 (65.3) *	99.5 (65.8) *	99.5 (62.7) *
Completeness (%)	99.7 (100) *	99.5 (99.1) *	99.8 (100) *	99.6 (98.3) *
Redundancy	7.04 (7.34) *	5.5 (5.5) *	6.2 (6.5) *	5.2 (5.3) *
Refinement				
Resolution range (Å)	50.00-3.50	50.00-2.04	30.00-2.50	30.00-2.52
No. reflections	10032	47507	26845	25997
<i>R</i> _{work} / <i>R</i> _{free}	0.200 / 0.242	0.177 / 0.204	0.198 / 0.234	0.207/0.239
No. atoms	4580	4961	4714	4749
Protein	4513	4497	4478	4506
<i>B</i> -factors				
Protein	66.21	45.04	49.46	57.19
Zinc ions	92.18	51.71	53.23	64.13
CoA variant	42.90	31.97	36.38	38.31
Water	24.74	45.44	37.64	40.51
R.m.s deviations				
Bond lengths (Å)	0.009	0.009	0.009	0.011
Bond angles (°)	1.390	1.304	1.404	1.321

*Values in parentheses are for highest-resolution shell.

19

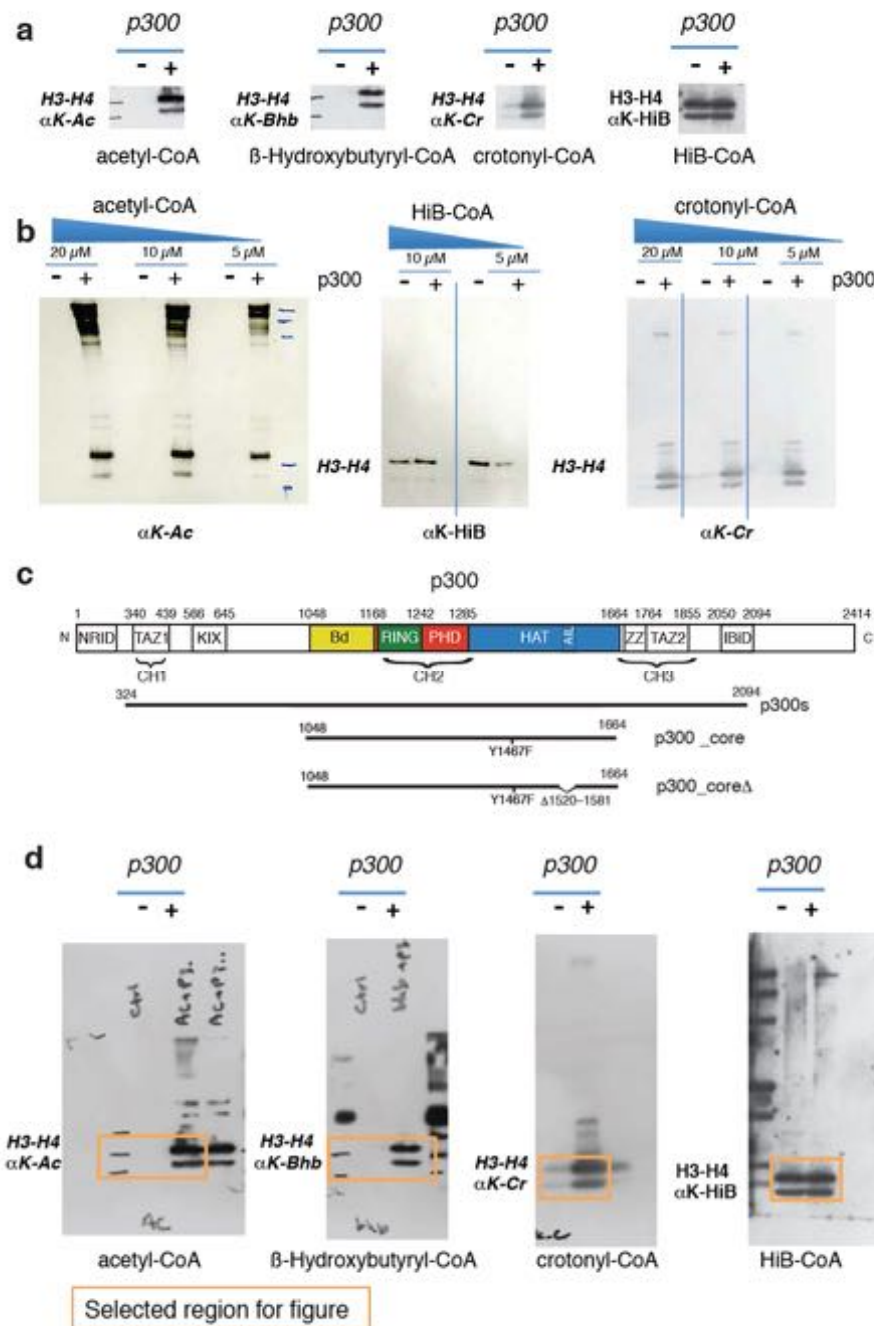
20

21

22

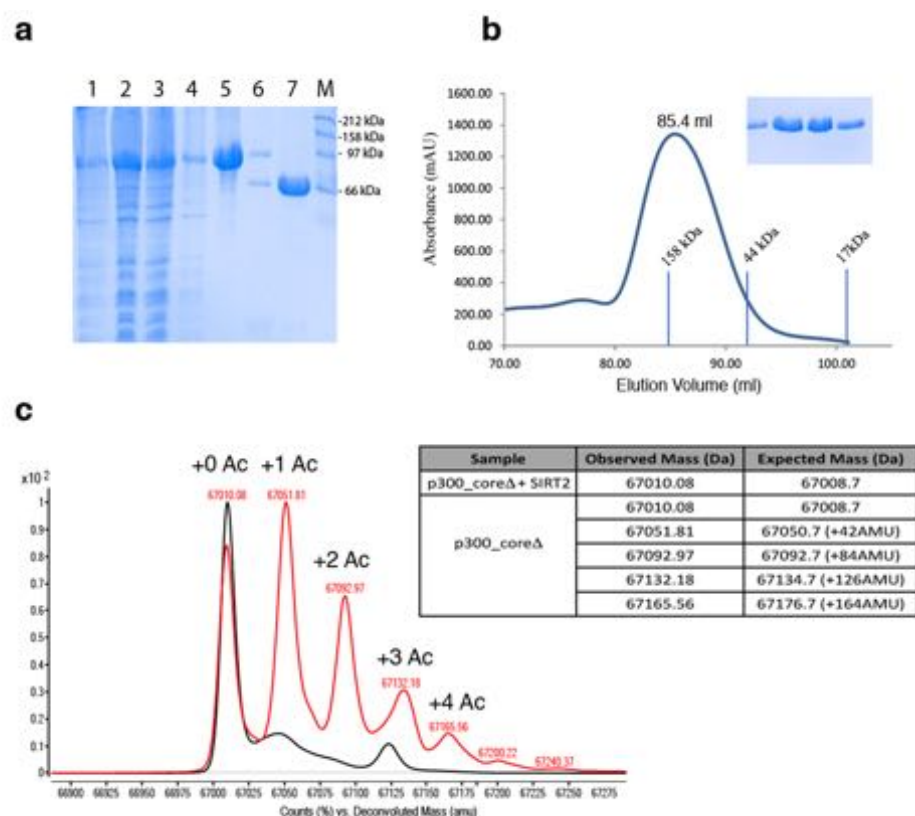
23

24 **Supplementary Figures**
25



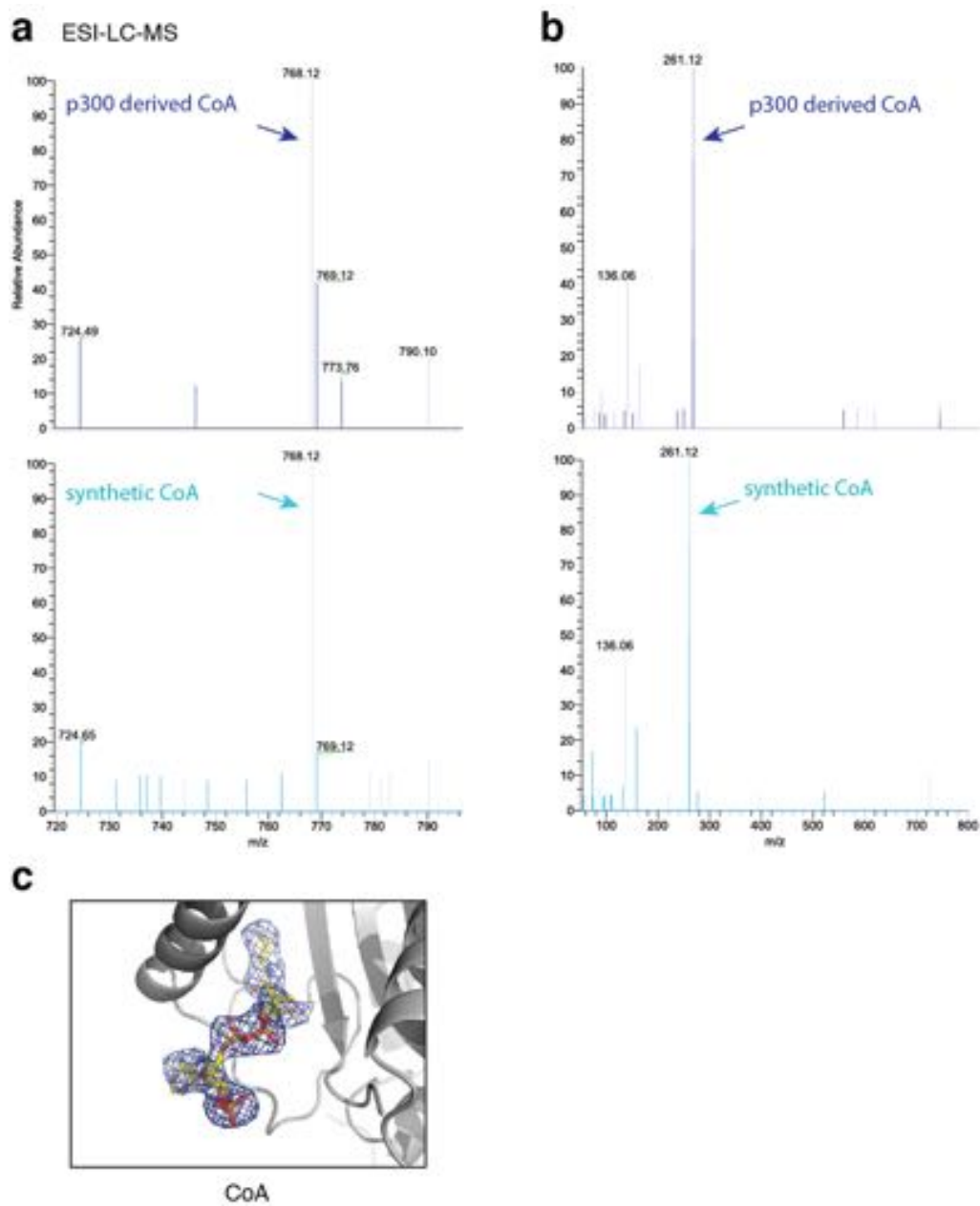
26
27 **Supplementary Figure 1. Acyltransferase activity of p300.** The samples were analyzed by
28 SDS-PAGE and western blotting with the indicated antibodies. (a) HAT assay on histone
29 tetramers H3-H4. H3-H4 was incubated for 30 minutes in the absence (-) or presence (+) of
30 p300 with 50 μ M acyl-CoA. β -hydroxybutyrylation and crotonylation required p300,
31 incubation with Hib-CoA resulted in nonenzymatic acylation. (b) Incubation at lower acyl-
32 CoA concentrations confirms that Hib-CoA is reactive in the absence of p300. The reactions

33 were repeated at least twice with consistent results. (c) Domain organization of p300 and
 34 constructs used. NRID (nuclear receptor interaction domain), TAZ1 (transcriptional adaptor
 35 zinc finger domain 1), KIX (kinase inducible domain of CREB interacting domain),
 36 Bromodomain (Bromodomain), RING (Really Interesting New Gene), PHD (Plant
 37 Homeodomain), HAT (Histone Acetyltransferase domain), ZZ (ZZ-type zinc finger domain),
 38 TAZ2 (transcriptional adaptor zinc finger domain 2), and NCBD (nuclear coactivator binding
 39 domain or IRF3-binding domain (IBID)). The TAZ1, the RING-PHD segment and the ZZ-
 40 TAZ2 domains are also known as the CH1 (cysteine/histidine-rich region 1), CH2 and CH3
 41 regions, respectively. (d) Uncropped images of the Western blots shown in (a)
 42
 43

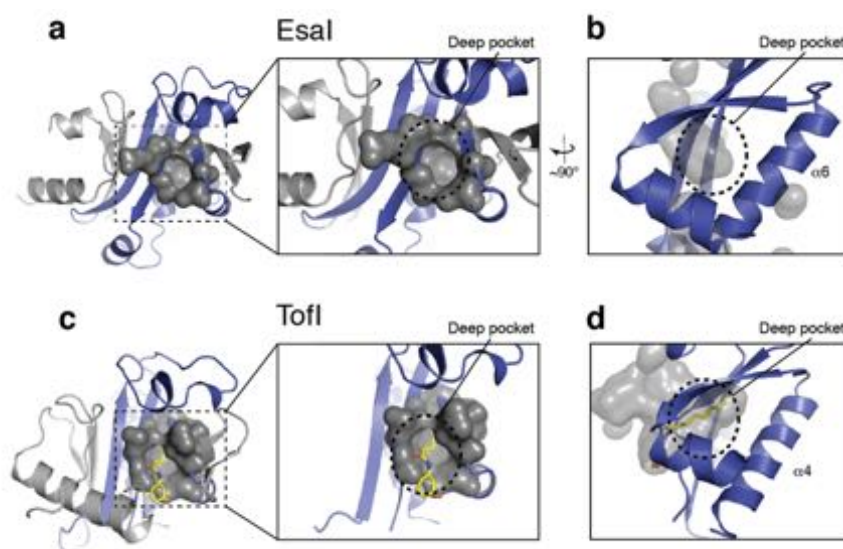


44
 45 **Supplementary Figure 2. Expression and Purification of p300 'core'.** (a) SDS-PAGE
 46 analysis of samples collected during p300 expression and purification. 1 crude lysate, 2
 47 soluble lysate, 3 flow-through from glutathione resin, 4 wash, 5 boiled glutathione resin, 6
 48 glutathione resin after rTEV cleavage, 7 Load on HiLoad Superdex 200 16/60 column. (b)
 49 p300 eluted as a single peak at 85.4 ml from the gel filtration column. Inset, SDS-PAGE
 50 analysis of the central fractions of the peak. Elution volume corresponded to that of a p300
 51 monomer. (c) ESI- Mass spectrometry analysis showed a series of peaks corresponding to

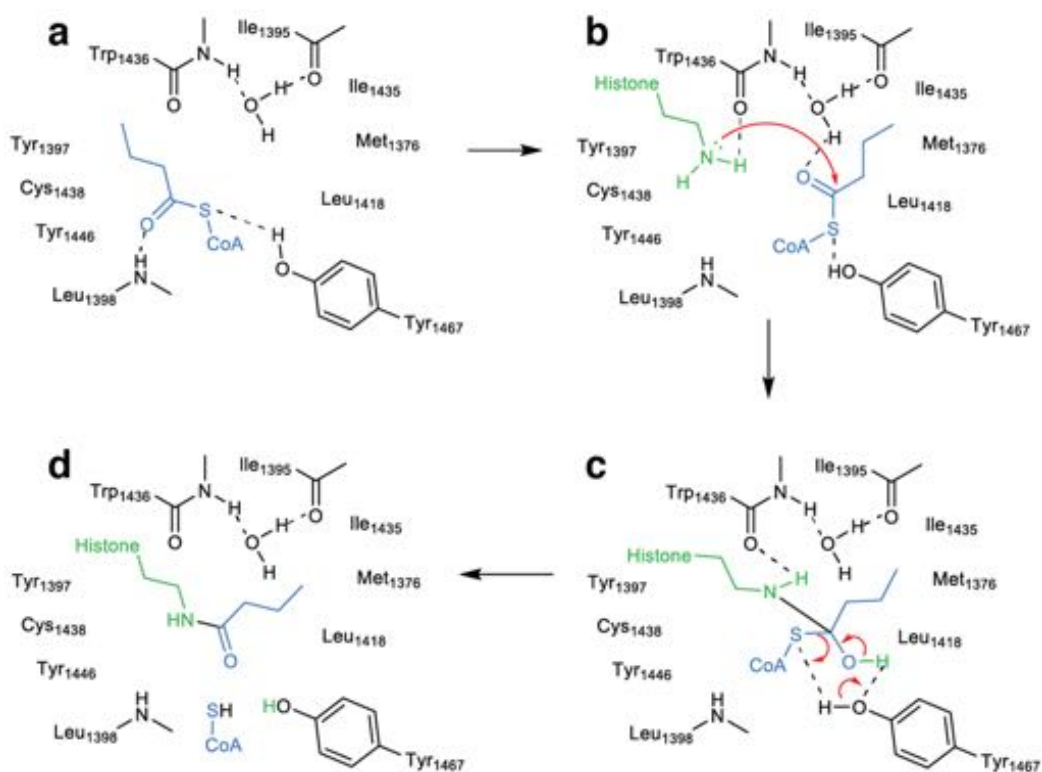
52 increasing acetylation states (+42 AMU) in the absence of SIRT2 (red line). Upon co-
 53 expression of SIRT2, mostly deacetylated p300 was obtained (black line).
 54



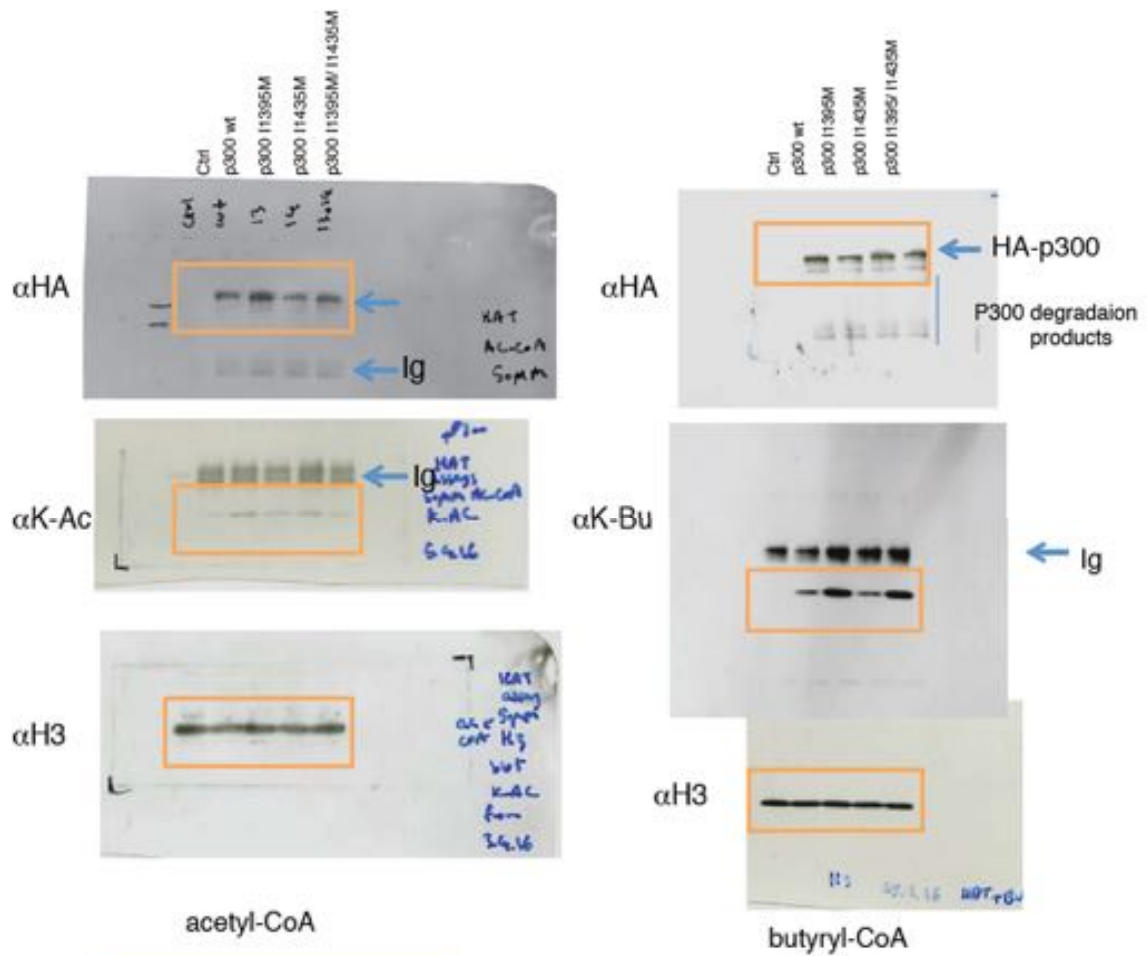
55
 56 **Supplementary Figure 3. p300 ‘core’ co-purifies with Coenzyme A.** (a) MS spectra of the
 57 CoA ligand derived from purified p300 (dark blue) and synthetic CoA (light blue) are shown.
 58 Arrows highlight the ion product with m/z 766.1, indicative of CoenzymeA. (b) Tandem MS
 59 fragmentation spectra of the p300-derived CoA ligand and synthetic CoA. (c) The crystal
 60 structure of p300_coreΔ (grey) in complex with CoenzymeA (yellow) co-purified from insect
 61 cells. A difference density omit map contoured at 1.5σ is shown in blue.



Supplementary Figure 4. Structurally related Acyl-HSL synthases have a deep acyl-chain binding pocket. (a) Esal structure (PDB code 1K4J). The structurally conserved core regions that are similar to acetyltransferases are colored in blue and the flanking N- and C-terminal regions are colored in grey. The inset displays the van der Waals surface, colored in grey, around the acyl-chain binding site. (b) The acyl-chain binding cavity of Esal is shown as a transparent grey surface. (c) TofI (PDB code 3P2H). The J8-C8 inhibitor is shown as yellow sticks. The inset displays the van der Waals surface, colored as above. (d) The acyl-chain binding cavity of TofI is shown as a transparent grey surface.

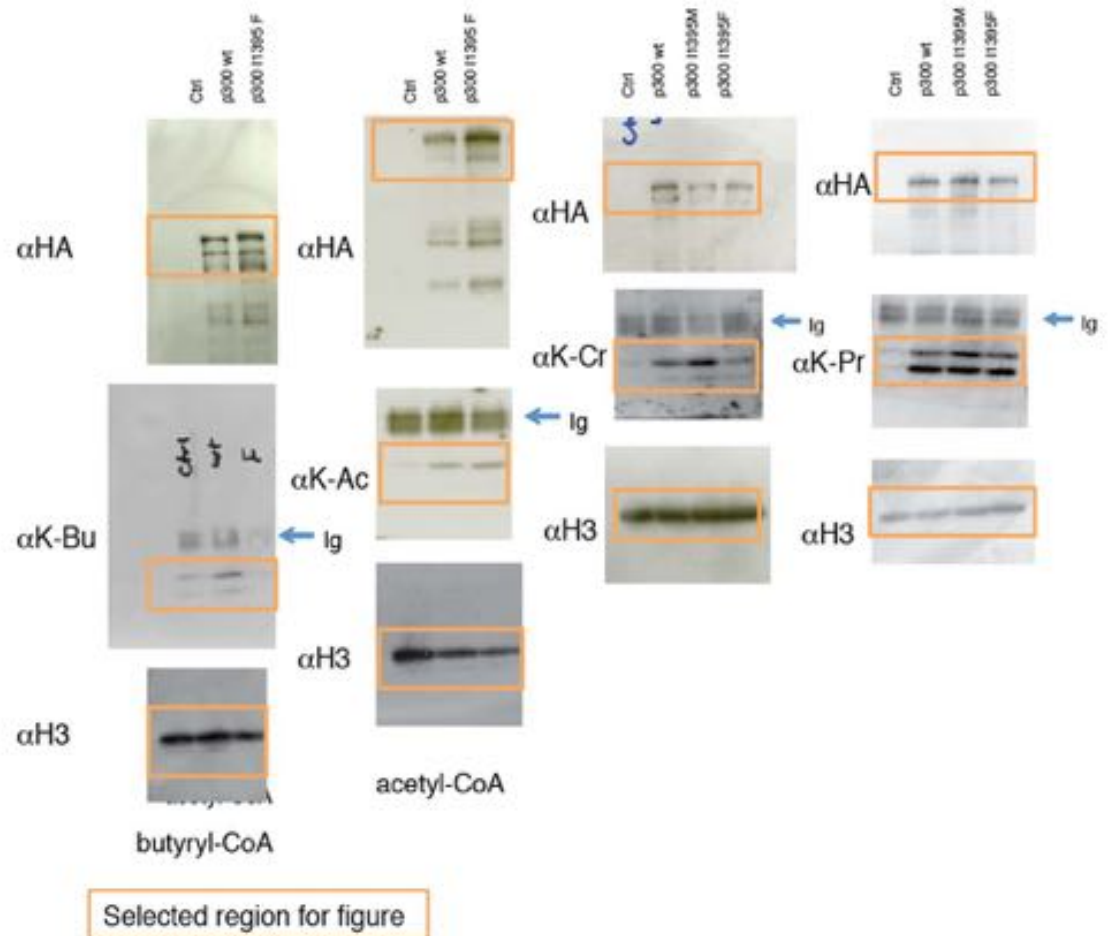


Supplementary Figure 5. Proposed mechanism of p300 acyl chain transfer. (a) The p300 active site residues are drawn in black and the acyl-CoA substrate in blue. Important residues are indicated. (b) Lysine substrate (green) engages the substrate binding tunnel and W1436 promotes deprotonation. The aliphatic portion of acyl-CoA rotates into the hydrophobic cavity allowing nucleophilic attack of the lysine attacks onto the carbonyl of acyl-CoA. (c) Y1467 acts as a general acid to protonate the leaving group. (d) The Acyl-lysine-containing product leaves, followed by CoASH.



Selected region for figure

Supplementary Figure 6. Uncropped images of Western blots shown in Figure 5a.



Supplementary Figure 7. Uncropped images of Western blots shown in Figure 5b.

Discussion and Perspectives

The recent discovery of lysine acylations opens many questions on the importance and function of these new modifications in the biology of chromatin. The present work focusing on butyrylation is the result of a collaboration between my host lab and Professor Yingming Zhao ([Tan et al., 2011](#)).

In this study, we have sought to address several points including the genome-wide distribution, function, and effect of histone butyrylation on chromatin-related activities. Our study focused mainly on H4K5K8 because of Brdt's first bromodomain's ability to bind H4K5K8 in its acetylated form ([Moriniere et al., 2009](#)) and because Brdt is an important driver of spermatogenesis.

Our data indicate that butyrylation, as a new histone modification, presents functional similarities as well as differences with acetylation at the same positions.

Mouse spermatogenic cells were used to unravel the role and function of histone butyrylation. Spermatogenesis provides an excellent model to explore the potential function of new histone PTMs and to relate it to gene activity. In fact, each cell population expresses a specific gene expression program in a stage-specific manner. In addition, sex chromosomes become transcriptionally silent in meiotic cell and only a fraction of these genes become reactivated in post-meiotic cells ([Montellier et al., 2012](#)).

- Genome-wide mapping of butyrylation and acetylation on H4K5 and H4K8 designates butyrylation as an active histone mark

The use of this spermatogenic cell system and specific antibodies against acetylation and butyrylation on lysines 5 and 8 of histone H4 allowed us to map the genomic distribution of these marks and evaluate the impact of histone butyrylation on gene expression by performing ChIP-seq on spermatocytes and round spermatids. We found that the four considered histone marks are enriched at the TSS region of genes. This observation encouraged us to investigate the expression of these genes. Interestingly, genes marked with the four PTMs were found highly active, and the loss of only one of these marks at gene TSSs was associated with a significant reduction of gene expression levels. In addition, the absence of these four marks

during MSCI constituted another argument in favor of a role for butyrylation in stimulating gene expression.

Recently my host laboratory also reported increased histone crotonylation, Kcr and histone 2-hydroxyisobutyrylation (Khib) at the TSS of X-linked genes that escape the above mentioned sex chromosome inactivation in post-meiotic cells compared to spermatocytes (Tan et al., 2011; Dai et al., 2014). Based on these observations, as well as on our observation that acetylation and butyrylation marks are depleted on these same regions, we can propose differential functions for various histone acylations in the context of a repressive environment, such as that of sex chromosomes. We propose that histone acetylation and butyrylation cannot sustain gene expression in a repressive environment, while Kcr and Khib do.

This differential action of histone acylation could be explained by the inability of HDACs to equally remove individual acyl marks. While most of these newly discovered acyl lysine modifications could be erased with the activity of NAD⁺-dependent Sirtuins, the removal of some of them also depends on the activity of Zn²⁺-dependent HDACs. For instance, SIRT5 removes succinyl, malonyl, and glutaryl. SIRT1 and SIRT2 erase crotonyl, propionyl, and butyryl. Sirt3 has been reported as the main *in vivo* decrotonylase (Bao et al., 2014).

All these considered acylations seem to be mediated by at least some the known HATs and to directly stimulate transcription. This is shown by the ability of the HAT p300 to ensure butyrylation (as previously reported for crotonylation). In an *in vitro* assay, the incubation of p300 with reconstituted histone octamer and chromatin in the presence of equal amounts of acetyl-CoA and butyryl-CoA followed by MS analysis, demonstrated the ability of this HAT to modify all the lysines present on the histone H4 tail randomly, and more specifically on our sites of interest (Lysines 5 and 8 of histone H4). We used a p300- and p53-dependent *in vitro* transcription system to show that histone butyrylation regulates gene expression positively (Goudarzi et al., 2016). This result is in line with our ChIP-seq analysis on spermatocytes and round spermatids, which also showed the co-occurrence and enrichment of H4K5K8 acetyl and butyryl marks at gene TSSs suggesting that histone butyrylation has a role in gene activation *in vivo*.

In summary, while histone butyrylation, like acetylation, crotonylation and hydroxybutyrylation, is clearly an activator of transcription it also shows marked differences

compared to these acylations. Butyrylation is a four C acyl group, which, like Kcr, repulses most of the known bromodomains (Flynn et al., 2015) but, unlike Kcr, it is depleted from the sex chromosome in post-meiotic cells (Goudarzi et al., 2016). Therefore each type of acyl group could have a specific function and hence their "choice" by the upstream enzymatic machinery could have important functional implications.

- Regulation of gene expression and p300 activity: toward differential histone lysine acylations via metabolism

The cellular (cytoplasmic and nuclear) reservoir of acetyl-CoA used to modify histone is mainly regulated through the function of ATP-citrate lyase (ACL) (Wellen et al., 2009). This enzyme generates acetyl-CoA by transforming citrate to oxaloacetate and acetyl-CoA. Additionally, Acyl-CoA Synthetase (ACSS2) ensures the production of acetyl-CoA through an alternative pathway by adding acetate to free CoA-SH. Additionally, ACSS2 has been shown to generate acetyl as well as non-acetyl acyl-CoAs by converting short chain fatty acids (SCFAs) such as acetate, crotonate, and β -hydroxybutyrate respectively to acetyl-CoA, crotonyl-CoA, and β -hydroxybutyryl-CoA (Sabari et al., 2015; Xie et al., 2016).

Manipulation of the cellular concentrations of acyl-CoAs by targeting corresponding enzymes with siRNAs and by treating cells with SCFAs could provide hints concerning the possible links between metabolism, chromatin and gene expression. Recent studies have shown that the knockdowns of ACL and of ACSS2 increase and decrease the levels of H3K18cr, respectively. In addition, the treatment of cells with crotonate enhanced H3K18cr level in a dose-dependent manner (Sabari et al., 2015; Xie et al., 2016). Another work from Zhao group demonstrates that elevated levels of β -hydroxybutyrate (bhb) in the liver and kidney of starved mice enhances histone lysine β -hydroxybutyrylation (Kbhb) and induces gene expression (Xie et al., 2016).

In a recent study on patients with metabolic disorders caused by mutations in the enzymes involved in the metabolic pathways such as malonyl-CoA decarboxylase (MCD), propionyl-CoA carboxylase (PCC), and mice deficient of short-chain acyl-CoA dehydrogenase (SCAD), Pougovkina et al reported increased levels of lysine malonylation, propionylation, and

butyrylation due to the accumulation of the corresponding acyl-CoAs (Pougovkina et al., 2014). To date, there is no evidence that lysine acylations could be catalyzed with enzymes that show restricted and specific activities. However, p300 has been reported to have acetyl- and acyltransferase activity to transfer acetyl-, propionyl-, crotonyl-, succinyl-, and glutaryl groups (Chen et al., 2007; Cheng et al., 2009; Sabari et al., 2015; Tan et al., 2014). If the same HAT catalyses different lysine acylations, the question is how does p300 choose among all these acyl group donors capable of competing with acetyl-group?

From the discussion above we know that, depending on the context, i. e., repressive environment, the functional output of different acylations could vary.

A study from Allis and colleagues showed that an increased concentration of crotonyl-CoA leads to an enhanced p300-mediated histone lysine crotonylation compared to histone lysine acetylation (Sabari et al., 2015). This observation supports the hypothesis that an increase in the ratio of Acyl-CoA/acetyl-CoA (achieved by decreased or increased concentrations of acetyl-CoA and acyl-CoA, respectively) promotes the chance of acyl-CoA to compete with acetyl-CoA to be utilized by HATs such as p300 to modify histones with non-acetyl acyl-CoAs. This exchange of acetylation to crotonylation could lead to enhanced gene expression levels suggesting that a dynamic exchange in histone modifications, such as oscillations between different acylations on a given lysine residue, can have functional consequences.

All these observations favor the hypothesis that the occurrence of differential lysine acylations, capable of stimulating gene expression to a higher degree than lysine acetylation, would depend on the availability of acyl-CoAs, which is itself regulated metabolically.

- Histone modifications alternation could impact gene targeting by chromatin binding factors

To understand how PTMs exert their function on chromatin and find their mechanism of action, it is essential to find PTM's "writers", "readers", and "erasers". As mentioned before, in addition to lysine acetylation, p300 could mediate multiple lysine acylations including lysine crotonylation, propionylation, butyrylation, and β -hydroxybutyrylation (Kaczmarska et al., 2016).

Flynn et al tested the ability of a series of human bromodomains to recognize some of the new histone acyl modifications. Although most of the acetyl-lysine bromodomain readers bind to propionyl lysines, the increasing length of acyl groups in butyryl- and crotonyl- lysines restricts the interaction of most of the assayed bromodomain readers. While BRD9 and CECR2 bind to Kbu strongly, the second bromodomain of TAF1 is the only bromodomain reader of Kcr (Flynn et al., 2015). Interestingly several very recent investigations have identified the YEATS domain as a reader of histone Kcr (Andrews et al., 2016; Li et al., 2016; Zhang et al. 2016).

Our immunofluorescence assay using specific antibodies against H4K5ac, H4K5bu, H4K8ac, and H4K8bu showed a specific differential labelling of acetylation and butyrylation of H4K5 and H4K8 in late spermatogenic cells compared to earlier male germ cells. Therefore we made the hypothesis that butyrylation might have special effects on Brdt's function. A combined analysis of our previous Brdt ChIP-seq data and of our present PTMs ChIP-seq data performed on male germ cells showed that Brdt is bound to the same TSS regions, corresponding to highly active genes, marked by both acetylation and butyrylation on K5 and K8 of histone H4 (H4K5acbu and H4K8acbu correlate with Brdt peaks at gene TSSs). Therefore the next question was whether Brdt could bind to H4K5buK8bu as well as to different combinations of acetylation and butyrylation.

Our pull-down assay based on different possible combinations of acetylation and butyrylation at H4K5K8 revealed that the butyrylation of K5 prevents the binding of Brdt. Indeed our structural and functional analyses show that Brdt's binding to H4K5 depends on the nature of the PTM modifying the lysine 5 of histone H4: while butyrylation inhibits this binding, acetylation supports it. Therefore a dynamic exchange of acetylation and butyrylation could contribute to the genome reorganization mediated by Brdt. In the late phases of spermatogenesis, the inability of Brdt to bind butyrylated histones could delay histone removal, implying that the differential timing of histone removal in various genomic regions could be determined by the nature of the acyl-CoA modified histone lysine residues (Goudarzi et al., 2016).

These data also suggest that the replacement of nucleosome-containing butyrylated histone H4 by protamine might not involve the intermediate step of TPs incorporation to chromatin

globally observed in the process of histone-to-Prm exchange, because of Brdt's inability in recognizing butyrylated histone. According to previous reports showing that TP1/TP2 deficient cells have less ability to compact their genome, we speculate that the regions containing butyrylated histones would not undergo compaction in a similar manner as the regions containing acetylated histone in mature spermatozoa ([Gaucher et al., 2010](#)).

The ability of Brdt to recognize H4K5acK8bu but not H4K5buK8ac raises the possibility that bromodomain readers might tolerate combinations of acetyl- and non-acetyl modifications. These observations encouraged other groups to search for possible readers of non-acetyl modifications. They found that the YEATS domain, previously reported as a Kac reader actually prefers binding Kcr, Kbu, and Kpr than Kac, with and shows the highest affinity for Kcr ([Li et al., 2014](#); [Andrews et al., 2016](#); [Zhao et al., 2016](#); [Zhang et al., 2016](#)).

The fact that the same lysine residue can be modified with different acyl-CoAs and that the new histone PTMs (Kcr, Kbu, Kbhb, and Khib) present features different from acetylation, allow us to propose that differential histone acylations (replacement of acetyl with other acyl-groups) would have specific functional consequences.

- CBP/p300 is involved in establishing the male epigenome and the corresponding gene expression programs

Although the hyperacetylation wave occurring during the transition of round to elongating spermatids has been shown to drive histone replacement and the male genome reorganization, how this histone hyperacetylation is established is an issue that remains unclear. To answer this question we focused our attention on CBP and p300. Our data show that the double knockout mice with slightly reduced Cbp and p300 mRNA levels in post-meiotic cells are fertile with normal spermatogenesis progression and hyperacetylation wave. However these investigations highlighted a regulated CBP/p300-dependent gene expression program activated in late male haploid cells just prior to the shut down of gene expression.

A whole genome transcriptomic analysis performed in round and elongating spermatids comparing wild type and mutant mice revealed that the expression of a small group of genes was affected by this partial depletion of cbp and p300 in elongating spermatids and that these affected genes were mainly involved in metabolic remodeling. In this study we also discovered a group of genes which are highly expressed in pre-meiotic cells, repressed in spermatocytes and early post-meiotic cells and then become reactivated in elongating spermatids, just prior to the global shut-down of transcription ([Boussouar et al., 2014](#)).

Although this work could not highlight a specific role for CBP and p300 in controlling histone hyperacetylation, it showed that CBP and p300 control a specific gene expression program in late spermatogenic cells. It would be very interesting to use these mice and measure by ChIP-seq the alteration of acetylation and butyrylation in a position and stage-specific manner. This would give us an information on how CBP and p300 regulate the differential histone acylations in a completely physiological setting.

- The HAT p300 catalyzes differential histone acylation with the ability to discriminate between acyl-CoAs

It has been shown that p300 can use a wide range of acyl-CoAs to modify histone lysine residues ([Chen et al., 2007](#); [Cheng et al., 2009](#); [Sabari et al., 2015](#)). Our investigation also shows that p300 could use acetyl-CoA or butyryl-CoA to randomly modify a specific position on histones *in vitro*. However our structural work on p300 shows that p300 largely prefers acetyl-CoA over longer CoA derivatives. This observation suggests that *in vivo* either p300 acquires a co-factor enhancing its acyl-transferase activity or non-acetyl acyl-transferase activity is driven by other type of enzymes.

In addition, these studies also indicate the importance of the residues involved in the acyltransferase activity of p300. The mutation of Ileu1395 located in the hydrophobic back pocket of p300 to Met led to a gain-of-function butyryltransferase activity of p300 whereas Ileu1395 mutated to bulkier amino acids such as Phe has an opposite effect. This observation implies that Ileu1395 plays an important role in controlling the entry of acyl-CoAs into the hydrophobic back pocket of p300 ([Kaczmarska et al., 2016](#)).

These studies also provide a hint on how a co-factor could modify the structure of p300 to facilitate histone acylation over acetylation

General Conclusion

General conclusion and future perspectives:

The work performed in the frame of my thesis highlights the important role of a histone modification, butyrylation, during male genome reprogramming and supports the hypothesis that stage-specific gene expression during spermatogenesis could be controlled by the cell metabolism context. A detailed investigation of butyrylation in parallel with acetylation revealed the functional significance of histone butyrylation and showed how it could be distinguished from and compete with the well-studied histone acetylation.

Our study supports the idea that the metabolic status of the cell could determine the nature and level of histone modifications. In particular, it shows a critical role for H4 K5 butyrylation on transcription with a special role in modulating the interaction of chromatin binding factors such as Brdt. We found that H4K5bu interferes with Brdt binding and could therefore mediate its dynamic turnover in a specific chromatin region.

Recent work has shown that an increased histone crotonylation, observed in response to a crotonate treatment in the cell or resulting from a p300-catalyzed histone crotonylation, stimulates transcription to a greater degree than histone acetylation. This differential histone acylations could have important impacts on chromatin function through the control of the interaction of chromatin binding factors as well as on gene expression programs.

We hypothesize that in the active dynamic regions of chromatin, histone acetyltransferases (HATs), possibly p300 and/or acyltransferases, use available acyl-CoAs (in the present study acetyl-CoA and butyryl-CoA) to modify histone H4 tails at K5 and K8, and that HDAC activity contributes to the dynamic exchange of these marks at active gene TSSs which can affect Brdt's affinity toward chromatin. This could be true for other bromodomain-containing factors.

An important conclusion from this work is that an oscillation between acetylation and butyrylation (or other long chain fatty acids) on specific positions on histone H4, especially at the position K5, is important for the turnover of Brdt (or other bromodomain factors) binding and hence should contribute to the dynamic nature of active gene TSSs.

Taken together, the study presented here and those in the literature on histone lysine acylations link cell metabolism to chromatin function and gene expression. These investigations open the door to the new concept that any changes in cell metabolism and in the balance of acyl metabolites could trigger changes in chromatin functions and diseases by altering the type of histone PTM used ([Dutta et al., 2016](#)).

Future works are needed to answer the following open questions:

- 1) What are precisely all the regulatory pathways of cellular butyryl-CoA production? 2) Are there any specific enzymes other than the regular HATs that manage deposition of non-acetyl acyl-CoAs on lysine residues?
- 3) Could a local increase in the concentration of acyl-CoAs drive lysine acylations in a non-enzymatic manner?
- 4) Could lysine acylations alterations contribute to malignant transformation of cells?

These questions hit a very hot topic in biology and any answer found to any of these questions should highly impact our understanding of epigenetics.

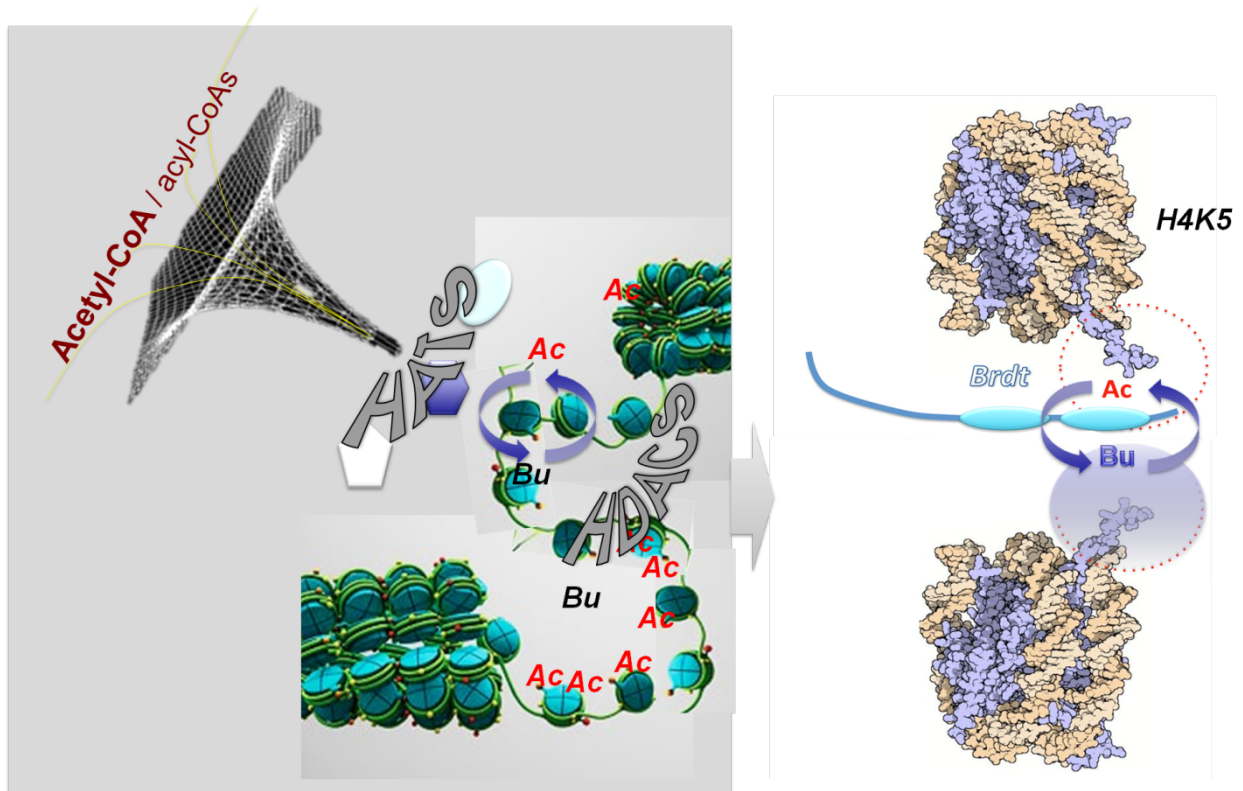


Figure working model: Dynamic exchange of histone acylations. Histone acetyl transferases can use Acetyl-CoA and any available acyl-CoAs to modify histones randomly. HDACs can remove these histone modifications. The turnover of histone modifications increases the dynamics of chromatin binding proteins such as Brdt that has a very important role in controlling male gene expression profile.

References

Akama, K., Kondo, M., Sato, H., and Nakano, M. (1999). Transition protein 4 from boar late spermatid nuclei is a topological factor that stimulates DNA-relaxing activity of topoisomerase I. *FEBS Lett* *442*, 189-192.

Akama, K., Sato, H., Hasegawa, S., Shimada, I., and Nakano, M. (1998). Transition protein 1 from boar late spermatid nuclei having DNA-melting activity is a dimeric protein. *Biochem Mol Biol Int* *44*, 315-323.

Alberts, B. Johnson, A. Lewis, J. Raff, M. Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell*, 4th edition.

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* *9*, 923-937.

Andrews, F.H., Shinsky, S.A., Shanle, E.K., Bridgers, J.B., Gest, A., Tsun, I.K., Krajewski, K., Shi, X., Strahl, B.D., and Kutateladze, T.G. (2016). The Taf14 YEATS domain is a reader of histone crotonylation. *Nat Chem Biol* *12*, 396-398.

Angelov, D., Molla, A., Perche, P.Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P., and Dimitrov, S. (2003). The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol Cell* *11*, 1033-1041.

Ausio, J., and Abbott, D.W. (2002). The many tales of a tail: carboxyl-terminal tail heterogeneity specializes histone H2A variants for defined chromatin function. *Biochemistry* *41*, 5945-5949.

Baarends, W.M., Hoogerbrugge, J.W., Roest, H.P., Ooms, M., Vreeburg, J., Hoeijmakers, J.H., and Grootegeed, J.A. (1999). Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. *Dev Biol* *207*, 322-333.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res* *21*, 381-395.

Bao, X., Wang, Y., Li, X., Li, X.M., Liu, Z., Yang, T., Wong, C.F., Zhang, J., Hao, Q., and Li, X.D. (2014). Identification of 'erasers' for lysine crotonylated histone marks using a chemical proteomics approach. *Elife* *3*.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* *129*, 823-837.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. *Science* *301*, 1090-1093.

Bonisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res* *40*, 10719-10741.

Boussouar, F., Goudarzi, A., Buchou, T., Shiota, H., Barral, S., Debernardi, A., Guardiola, P., Brindle, P., Martinez, G., Arnoult, C., *et al.* (2014). A specific CBP/p300-dependent gene expression programme drives the metabolic remodelling in late stages of spermatogenesis. *Andrology* *2*, 351-359.

Boussouar, F., Rousseaux, S., and Khochbin, S. (2008). A new insight into male genome reprogramming by histone variants and histone code. *Cell Cycle* *7*, 3499-3502.

Braun, R.E. (2001). Packaging paternal chromosomes with protamine. *Nat Genet* *28*, 10-12.

Braun, R.E., Peschon, J.J., Behringer, R.R., Brinster, R.L., and Palmiter, R.D. (1989). Protamine 3'-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. *Genes Dev* *3*, 793-802.

Breiling, A., and Lyko, F. (2015). Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin* *8*, 24.

Burgess, R.J., and Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. *Nat Struct Mol Biol* *20*, 14-22.

Caron, C., Pivot-Pajot, C., van Grunsven, L.A., Col, E., Lestrat, C., Rousseaux, S., and Khochbin, S. (2003). Cdy1: a new transcriptional co-repressor. *EMBO Rep* 4, 877-882.

Caron, N., Veilleux, S., and Boissonneault, G. (2001). Stimulation of DNA repair by the spermatidal TP1 protein. *Mol Reprod Dev* 58, 437-443.

Celeste, A., Fernandez-Capetillo, O., Kruhlak, M.J., Pilch, D.R., Staudt, D.W., Lee, A., Bonner, R.F., Bonner, W.M., and Nussenzweig, A. (2003). Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. *Nat Cell Biol* 5, 675-679.

Chadwick, B.P., and Willard, H.F. (2001). A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J Cell Biol* 152, 375-384.

Chen, H.Y., Sun, J.M., Zhang, Y., Davie, J.R., and Meistrich, M.L. (1998). Ubiquitination of histone H3 in elongating spermatids of rat testes. *J Biol Chem* 273, 13165-13169.

Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C., Falck, J.R., Peng, J., Gu, W., and Zhao, Y. (2007). Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* 6, 812-819.

Cheng, Y.C. (2008). Molecular Mechanism in spermatogenesis.

Cheng, Z., Tang, Y., Chen, Y., Kim, S., Liu, H., Li, S.S., Gu, W., and Zhao, Y. (2009). Molecular characterization of propionyllysines in non-histone proteins. *Mol Cell Proteomics* 8, 45-52.

Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M., and Allis, C.D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol Cell* 5, 905-915.

Cho, C., Willis, W.D., Goulding, E.H., Jung-Ha, H., Choi, Y.C., Hecht, N.B., and Eddy, E.M. (2001). Haploinsufficiency of protamine-1 or -2 causes infertility in mice. *Nat Genet* 28, 82-86.

Colak, G., Xie, Z., Zhu, A.Y., Dai, L., Lu, Z., Zhang, Y., Wan, X., Chen, Y., Cha, Y.H., Lin, H., *et al.* (2013). Identification of lysine succinylation substrates and the succinylation regulatory enzyme CobB in *Escherichia coli*. *Mol Cell Proteomics* 12, 3509-3520.

Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* 11, 1037-1043.

Cutter, A.R., and Hayes, J.J. (2015). A brief review of nucleosome structure. *FEBS Lett* 589, 2914-2922.

Dai, L., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., Debernardi, A., Buchou, T., Rousseaux, S., Jin, F., *et al.* (2014). Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat Chem Biol* 10, 365-370.

Dancy, B.M., and Cole, P.A. (2015). Protein lysine acetylation by p300/CBP. *Chem Rev* 115, 2419-2452.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* 319: 1097-1113.

De Jonge, C., Barratt, C. (2006). The Sperm Cell. Production, Maturation, Fertilization, Regeneration.

de Rooij, D.G. (1998). Stem cells in the testis. *Int J Exp Pathol* 79, 67-80.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399, 491-496.

Domenjoud, L., Kremling, H., Burfeind, P., Maier, W.M., and Engel, W. (1991). On the expression of protamine genes in the testis of man and other mammals. *Andrologia* 23, 333-337.

Downs, J.A., Lowndes, N.F., and Jackson, S.P. (2000). A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* 408, 1001-1004.

Dutta, A., Abmayr, S.M., and Workman, J.L. (2016). Diverse Activities of Histone Acylations Connect Metabolism to Chromatin Function. *Mol Cell* 63, 547-552.

Eddy, E.M., and O'Brien, D.A. (1998). Gene expression during mammalian meiosis. *Curr Top Dev Biol* 37, 141-200.

Falkenberg, K.J., and Johnstone, R.W. (2014). Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* 13, 673-691.

Fan, J.Y., Gordon, F., Luger, K., Hansen, J.C., and Tremethick, D.J. (2002). The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nat Struct Biol* 9, 172-176.

Felsenfeld, G. (2014). A brief history of epigenetics. *Cold Spring Harb Perspect Biol* 6.

Feltus, F.A., Lee, E.K., Costello, J.F., Plass, C., and Vertino, P.M. (2006). DNA motifs associated with aberrant CpG island methylation. *Genomics* 87, 572-579.

Filipescu, D., Muller, S., and Almouzni, G. (2014). Histone H3 variants and their chaperones during development and disease: contributing to epigenetic control. *Annu Rev Cell Dev Biol* 30, 615-646.

Fink, M., Imholz, D., and Thoma, F. (2007). Contribution of the serine 129 of histone H2A to chromatin structure. *Mol Cell Biol* 27, 3589-3600.

Flynn, E.M., Huang, O.W., Poy, F., Oppikofer, M., Bellon, S.F., Tang, Y., and Cochran, A.G. (2015). A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. *Structure* 23, 1801-1814.

Gaucher, J., Boussouar, F., Montellier, E., Curtet, S., Buchou, T., Bertrand, S., Hery, P., Jounier, S., Depaux, A., Vitte, A.L., *et al.* (2012). Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* 31, 3809-3820.

Gaucher, J., Reynoird, N., Montellier, E., Boussouar, F., Rousseaux, S., and Khochbin, S. (2010). From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 277, 599-604.

Gautier, T., Abbott, D.W., Molla, A., Verdel, A., Ausio, J., and Dimitrov, S. (2004). Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* 5, 715-720.

Gebara, M.M., and McCarrey, J.R. (1992). Protein-DNA interactions associated with the onset of testis-specific expression of the mammalian Pkg-2 gene. *Mol Cell Biol* 12, 1422-1431.

Gilbert, N., Gilchrist, S., and Bickmore, W.A. (2005). Chromatin organization in the mammalian nucleus. *Int Rev Cytol* 242, 283-336.

Godmann, M., Auger, V., Ferraroni-Aguiar, V., Di Sauro, A., Sette, C., Behr, R., and Kimmins, S. (2007). Dynamic regulation of histone H3 methylation at lysine 4 in mammalian spermatogenesis. *Biol Reprod* 77, 754-764.

Gonzalez-Romero, R., Mendez, J., Ausio, J., and Eirin-Lopez, J.M. (2008). Quickly evolving histones, nucleosome stability and chromatin folding: all about histone H2A.Bbd. *Gene* 413, 1-7.

Goudarzi, A., Shiota, H., Rousseaux, S., and Khochbin, S. (2014). Genome-scale acetylation-dependent histone eviction during spermatogenesis. *J Mol Biol* 426, 3342-3349.

Goudarzi, A., Zhang, D., Huang, H., Barral, S., Kwon, O.K., Qi, S., Tang, Z., Buchou, T., Vitte, A.L., He, T., *et al.* (2016). Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters. *Mol Cell* 62, 169-180.

Govin, J., Caron, C., Lestrat, C., Rousseaux, S., and Khochbin, S. (2004). The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur J Biochem* 271, 3459-3469.

Govin, J., Dorsey, J., Gaucher, J., Rousseaux, S., Khochbin, S., and Berger, S.L. (2010). Systematic screen reveals new functional dynamics of histones H3 and H4 during gametogenesis. *Genes Dev* 24, 1772-1786.

Govin, J., Escoffier, E., Rousseaux, S., Kuhn, L., Ferro, M., Thevenon, J., Catena, R., Davidson, I., Garin, J., Khochbin, S., *et al.* (2007). Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J Cell Biol* 176, 283-294.

Green, G.R., Balhorn, R., Poccia, D.L., and Hecht, N.B. (1994). Synthesis and processing of mammalian protamines and transition proteins. *Mol Reprod Dev* 37, 255-263.

Grimes, S.R., Jr., Meistrich, M.L., Platz, R.D., and Hnilica, L.S. (1977). Nuclear protein transitions in rat testis spermatids. *Exp Cell Res* 110, 31-39.

Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., *et al.* (2013). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci* 17, 215-222.

Haberland, M., Montgomery, R.L., and Olson, E.N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet* 10, 32-42.

Handel, M.A. (2004). The XY body: a specialized meiotic chromatin domain. *Exp Cell Res* 296, 57-63.

Hazzouri, M., Pivot-Pajot, C., Faure, A.K., Usson, Y., Pelletier, R., Sele, B., Khochbin, S., and Rousseaux, S. (2000). Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur J Cell Biol* 79, 950-960.

Hogarth, C.A., and Griswold, M.D. (2010). The key role of vitamin A in spermatogenesis. *J Clin Invest* 120, 956-962.

Hoyer-Fender, S. (2003). Molecular aspects of XY body formation. *Cytogenet Genome Res* 103, 245-255.

Hublitz, P., Albert, M., and Peters, A.H. (2009). Mechanisms of transcriptional repression by histone lysine methylation. *Int J Dev Biol* 53, 335-354.

Jacob, Y., Bergamin, E., Donoghue, M.T., Mongeon, V., LeBlanc, C., Voigt, P., Underwood, C.J., Brunzelle, J.S., Michaels, S.D., Reinberg, D., *et al.* (2014). Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* 343, 1249-1253.

Jan, S.Z., Hamer, G., Repping, S., de Rooij, D.G., van Pelt, A.M., and Vormer, T.L. (2012). Molecular control of rodent spermatogenesis. *Biochim Biophys Acta* 1822, 1838-1850.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* 41, 941-945.

Jost, K.L., Bertulat, B., and Cardoso, M.C. (2012). Heterochromatin and gene positioning: inside, outside, any side? *Chromosoma* 121, 555-563.

Kaczmarzka, Z., Ortega, E., Goudarzi, A., Huang, H., Kim, S., Marquez, J.A., Zhao, Y., Khochbin, S., Panne, D., (2016). Structure of the p300 in complex with Acyl-CoA variants: implications for metabolic control of chromatin modification. In Press.

Kaji, K., and Kudo, A. (2004). The mechanism of sperm-oocyte fusion in mammals. *Reproduction* 127, 423-429.

Kalocsay, M., Hiller, N.J., and Jentsch, S. (2009). Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. *Mol Cell* 33, 335-343.

Karukurichi, K.R., Wang, L., Uzasci, L., Manlandro, C.M., Wang, Q., and Cole, P.A. (2010). Analysis of p300/CBP histone acetyltransferase regulation using circular permutation and semisynthesis. *J Am Chem Soc* 132, 1222-1223.

Kimmins, S., and Sassone-Corsi, P. (2005). Chromatin remodelling and epigenetic features of germ cells. *Nature* 434, 583-589.

Kistler, W.S., Noyes, C., Hsu, R., and Henrikson, R.L. (1975). The amino acid sequence of a testis-specific basic protein that is associated with spermatogenesis. *J Biol Chem* 250, 1847-1853.

Kleene, K.C. (1996). Patterns of translational regulation in the mammalian testis. *Mol Reprod Dev* 43, 268-281.

Kremling, H., Luerksen, H., Adham, I.M., Klemm, U., Tsaousidou, S., and Engel, W. (1989). Nucleotide sequences and expression of cDNA clones for boar and bull transition protein 1 and its evolutionary conservation in mammals. *Differentiation* 40, 184-190.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324, 929-930.

Krude, T. (1995). Chromatin. Nucleosome assembly during DNA replication. *Curr Biol* 5, 1232-1234.

Laberge, R.M., and Boissonneault, G. (2005). On the nature and origin of DNA strand breaks in elongating spermatids. *Biol Reprod* 73, 289-296.

Latrick, C.M., Marek, M., Ouarrhni, K., Papin, C., Stoll, I., Ignatyeva, M., Obri, A., Ennifar, E., Dimitrov, S., Romier, C., *et al.* (2016). Molecular basis and specificity of H2A.Z-H2B recognition and deposition by the histone chaperone YL1. *Nat Struct Mol Biol* 23, 309-316.

Lawrence, M., Daujat, S., and Schneider, R. (2015). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* 32, 42-56.

Leduc, F., Maquennehan, V., Nkoma, G.B., and Boissonneault, G. (2008). DNA damage response during chromatin remodeling in elongating spermatids of mice. *Biol Reprod* 78, 324-332.

Lee, J.S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. *Cell* 142, 682-685.

Levesque, D., Veilleux, S., Caron, N., and Boissonneault, G. (1998). Architectural DNA-binding properties of the spermatidal transition proteins 1 and 2. *Biochem Biophys Res Commun* 252, 602-609.

Lewin, B. Krebs, J.E. (2015). LEWIN'S CELL.

Li, A., Maffey, A.H., Abbott, W.D., Conde e Silva, N., Prunell, A., Siino, J., Churikov, D., Zalensky, A.O., and Ausio, J. (2005). Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). *Biochemistry* 44, 2529-2535.

Li, Y., Sabari, B.R., Panchenko, T., Wen, H., Zhao, D., Guan, H., Wan, L., Huang, H., Tang, Z., Zhao, Y., *et al.* (2016). Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Mol Cell* 62, 181-193.

Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., Ren, Y., Jin, Q., Dent, S.Y., Li, W., *et al.* (2014). AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell* 159, 558-571.

Liu, B., Lin, Y., Darwanto, A., Song, X., Xu, G., and Zhang, K. (2009). Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells. *J Biol Chem* 284, 32288-32295.

Lu, L.Y., Wu, J., Ye, L., Gavrilina, G.B., Saunders, T.L., and Yu, X. (2010). RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Dev Cell* 18, 371-384.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.

Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). Chromatin as dynamic 10-nm fibers. *Chromosoma* 123, 225-237.

Mali, P., Kaipia, A., Kangasniemi, M., Toppari, J., Sandberg, M., Hecht, N.B., and Parvinen, M. (1989). Stage-specific expression of nucleoprotein mRNAs during rat and mouse spermiogenesis. *Reprod Fertil Dev* 1, 369-382.

Marino-Ramirez, L., Kann, M.G., Shoemaker, B.A., and Landsman, D. (2005). Histone structure and nucleosome stability. *Expert Rev Proteomics* 2, 719-729.

Marques, M., Laflamme, L., Gervais, A.L., and Gaudreau, L. (2010). Reconciling the positive and negative roles of histone H2A.Z in gene transcription. *Epigenetics* 5, 267-272.

McNally, M.A., and Hartman, A.L. (2012). Ketone bodies in epilepsy. *J Neurochem* 121, 28-35.

Meistrich, M.L., Mohapatra, B., Shirley, C.R., and Zhao, M. (2003). Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* 111, 483-488.

Meyer-Ficca, M.L., Scherthan, H., Burkle, A., and Meyer, R.G. (2005). Poly(ADP-ribosyl)ation during chromatin remodeling steps in rat spermiogenesis. *Chromosoma* 114, 67-74.

Montellier, E., Boussouar, F., Rousseaux, S., Zhang, K., Buchou, T., Fenaille, F., Shiota, H., Debernardi, A., Hery, P., Curtet, S., *et al.* (2013). Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev* 27, 1680-1692.

Montellier, E., Rousseaux, S., Zhao, Y., and Khochbin, S. (2012). Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *Bioessays* 34, 187-193.

Moriniere, J., Rousseaux, S., Steuerwald, U., Soler-Lopez, M., Curtet, S., Vitte, A.L., Govin, J., Gaucher, J., Sadoul, K., Hart, D.J., *et al.* (2009). Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 461, 664-668.

Musselman, C.A., Lalonde, M.E., Cote, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* 19, 1218-1227.

Narlikar, G.J., Sundaramoorthy, R., and Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* 154, 490-503.

Noble, D. (2015). Conrad Waddington and the origin of epigenetics. *J Exp Biol* 218, 816-818.

Nottke, A., Colaiacovo, M.P., and Shi, Y. (2009). Developmental roles of the histone lysine demethylases. *Development* 136, 879-889.

Ooi, S.L., and Henikoff, S. (2007). Germline histone dynamics and epigenetics. *Curr Opin Cell Biol* 19, 257-265.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep* 10, 1235-1241.

Page, S.L., and Hawley, R.S. (2004). The genetics and molecular biology of the synaptonemal complex. *Annu Rev Cell Dev Biol* 20, 525-558.

Papoutsopoulou, S., Nikolakaki, E., Chalepakis, G., Kruft, V., Chevaillier, P., and Giannakouros, T. (1999). SR protein-specific kinase 1 is highly expressed in testis and phosphorylates protamine 1. *Nucleic Acids Res* 27, 2972-2980.

Park, J., Chen, Y., Tishkoff, D.X., Peng, C., Tan, M., Dai, L., Xie, Z., Zhang, Y., Zwaans, B.M., Skinner, M.E., *et al.* (2013). SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways. *Mol Cell* 50, 919-930.

Perche, P.Y., Vourc'h, C., Konecny, L., Souchier, C., Robert-Nicoud, M., Dimitrov, S., and Khochbin, S. (2000). Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. *Curr Biol* 10, 1531-1534.

Piatti, P., Zeilner, A., and Lusser, A. (2011). ATP-dependent chromatin remodeling factors and their roles in affecting nucleosome fiber composition. *Int J Mol Sci* 12, 6544-6565.

Pivot-Pajot, C., Caron, C., Govin, J., Vion, A., Rousseaux, S., and Khochbin, S. (2003). Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* 23, 5354-5365.

Pougovkina, O., Te Brinke, H., Wanders, R.J., Houten, S.M., and de Boer, V.C. (2014). Aberrant protein acylation is a common observation in inborn errors of acyl-CoA metabolism. *J Inherit Metab Dis* 37, 709-714.

Quivy, J.P., Grandi, P., and Almouzni, G. (2001). Dimerization of the largest subunit of chromatin assembly factor 1: importance in vitro and during *Xenopus* early development. *EMBO J* 20, 2015-2027.

Rathke, C., Baarends, W.M., Awe, S., and Renkawitz-Pohl, R. (2013). Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta* 1839, 155-168.

Rathke, C., Baarends, W.M., Jayaramaiah-Raja, S., Bartkuhn, M., Renkawitz, R., and Renkawitz-Pohl, R. (2007). Transition from a nucleosome-based to a protamine-based chromatin configuration during spermiogenesis in *Drosophila*. *J Cell Sci* 120, 1689-1700.

Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* 9, 1091-1100.

Rousseaux, S., and Khochbin, S. (2015). Histone Acylation beyond Acetylation: Terra Incognita in Chromatin Biology. *Cell J* 17, 1-6.

Russell, L.D., Ettlin, R.A., Sinha-Hikim, A.P., Clegg, E.D. (1990). Histological and histopathological evaluation of the testis. Cache River Press, Florida.

Ryan, C.M., Kindle, K.B., Collins, H.M., and Heery, D.M. (2010). SUMOylation regulates the nuclear mobility of CREB binding protein and its association with nuclear bodies in live cells. *Biochem Biophys Res Commun* 391, 1136-1141.

Sabari, B.R., Tang, Z., Huang, H., Yong-Gonzalez, V., Molina, H., Kong, H.E., Dai, L., Shimada, M., Cross, J.R., Zhao, Y., *et al.* (2015). Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell* 58, 203-215.

Saksouk, N., Simboeck, E., and Dejardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* 8, 3.

Sarcinella, E., Zuzarte, P.C., Lau, P.N., Draker, R., and Cheung, P. (2007). Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol* 27, 6457-6468.

Sasaki, K., Ito, T., Nishino, N., Khochbin, S., and Yoshida, M. (2009). Real-time imaging of histone H4 hyperacetylation in living cells. *Proc Natl Acad Sci U S A* 106, 16257-16262.

Sassone-Corsi, P. (2002). Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science* 296, 2176-2178.

Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* 76, 75-100.

Shima, J.E., McLean, D.J., McCarrey, J.R., and Griswold, M.D. (2004). The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol Reprod* 71, 319-330.

Shinagawa, T., Takagi, T., Tsukamoto, D., Tomaru, C., Huynh, L.M., Sivaraman, P., Kumarevel, T., Inoue, K., Nakato, R., Katou, Y., *et al.* (2014). Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell* 14, 217-227.

Sin, H.S., Barski, A., Zhang, F., Kartashov, A.V., Nussenzweig, A., Chen, J., Andreassen, P.R., and Namekawa, S.H. (2012). RNF8 regulates active epigenetic modifications and escape gene activation from inactive sex chromosomes in post-meiotic spermatids. *Genes Dev* 26, 2737-2748.

Singh, J., and Rao, M.R. (1988). Interaction of rat testis protein, TP, with nucleosome core particle. *Biochem Int* 17, 701-710.

Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J. (2011). A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat Struct Mol Biol* 19, 25-30.

Soshnev, A.A., Josefowicz, S.Z., and Allis, C.D. (2016). Greater Than the Sum of Parts: Complexity of the Dynamic Epigenome. *Mol Cell* 62, 681-694.

Steger, K. (1999). Transcriptional and translational regulation of gene expression in haploid spermatids. *Anat Embryol (Berl)* 199, 471-487.

Steger, K. (2001). Haploid spermatids exhibit translationally repressed mRNAs. *Anat Embryol (Berl)* 203, 323-334.

Steger, K., Failing, K., Klonisch, T., Behre, H.M., Manning, M., Weidner, W., Hertle, L., Bergmann, M., and Kliesch, S. (2001). Round spermatids from infertile men exhibit decreased protamine-1 and -2 mRNA. *Hum Reprod* 16, 709-716.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.

Sun, W.J., Zhou, X., Zheng, J.H., Lu, M.D., Nie, J.Y., Yang, X.J., and Zheng, Z.Q. (2012). Histone acetyltransferases and deacetylases: molecular and clinical implications to gastrointestinal carcinogenesis. *Acta Biochim Biophys Sin (Shanghai)* 44, 80-91.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* 7, 1121-1124.

Szenker, E., Ray-Gallet, D., and Almouzni, G. (2011). The double face of the histone variant H3.3. *Cell Res* 21, 421-434.

Tachiwana, H., Kagawa, W., Osakabe, A., Kawaguchi, K., Shiga, T., Hayashi-Takanaka, Y., Kimura, H., and Kurumizaka, H. (2010). Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* 107, 10454-10459.

Takeda, N., Yoshinaga, K., Furushima, K., Takamune, K., Li, Z., Abe, S., Aizawa, S., and Yamamura, K. (2016). Viable offspring obtained from Prm1-deficient sperm in mice. *Sci Rep* 6, 27409.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., *et al.* (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016-1028.

Tan, M., Peng, C., Anderson, K.A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., Zhang, Y., *et al.* (2014). Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. *Cell Metab* 19, 605-617.

Thambirajah, A.A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A.H., and Ausio, J. (2006). H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. *J Biol Chem* 281, 20036-20044.

Thatcher, T.H., and Gorovsky, M.A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res* 22, 174-179.

Thiriet, C., and Hayes, J.J. (2005). Chromatin in need of a fix: phosphorylation of H2AX connects chromatin to DNA repair. *Mol Cell* 18, 617-622.

Timmermann, S., Lehrmann, H., Polesskaya, A., and Harel-Bellan, A. (2001). Histone acetylation and disease. *Cell Mol Life Sci* 58, 728-736.

Trasler, J.M. (2009). Epigenetics in spermatogenesis. *Mol Cell Endocrinol* 306, 33-36.

Tropberger, P., and Schneider, R. (2013). Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat Struct Mol Biol* 20, 657-661.

Turner, B.M. (2000). Histone acetylation and an epigenetic code. *Bioessays* 22, 836-845.

Turner, J.M., Aprelikova, O., Xu, X., Wang, R., Kim, S., Chandramouli, G.V., Barrett, J.C., Burgoyne, P.S., and Deng, C.X. (2004). BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation. *Curr Biol* 14, 2135-2142.

Turner, J.M., Mahadevaiah, S.K., Ellis, P.J., Mitchell, M.J., and Burgoyne, P.S. (2006). Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids. *Dev Cell* 10, 521-529.

vanWert, J.M., Wolfe, S.A., and Grimes, S.R. (1996). Binding of nuclear proteins to a conserved histone H1t promoter element suggests an important role in testis-specific transcription. *J Cell Biochem* 60, 348-362.

Varga-Weisz, P. (2001). ATP-dependent chromatin remodeling factors: nucleosome shufflers with many missions. *Oncogene* 20, 3076-3085.

Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. In *Nat Rev Mol Cell Biol*, pp. 178-189.

Vignali, M., Hassan, A.H., Neely, K.E., and Workman, J.L. (2000). ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* 20, 1899-1910.

Voigt, P., and Reinberg, D. (2011). Histone tails: ideal motifs for probing epigenetics through chemical biology approaches. *Chembiochem* 12, 236-252.

Vollmuth, F., and Geyer, M. (2010). Interaction of propionylated and butyrylated histone H3 lysine marks with Brd4 bromodomains. *Angew Chem Int Ed Engl* 49, 6768-6772.

Weinert, B.T., Scholz, C., Wagner, S.A., Iesmantavicius, V., Su, D., Daniel, J.A., and Choudhary, C. (2013). Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. *Cell Rep* 4, 842-851.

Weinhold, B. (2006). Epigenetics: the science of change. *Environ Health Perspect* 114, A160-167.

Wellen, K.E., Hatzivassiliou, G., Sachdeva, U.M., Bui, T.V., Cross, J.R., and Thompson, C.B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* 324, 1076-1080.

Widom, J., and Klug, A. (1985). Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. *Cell* 43, 207-213.

Williams, S.P., Athey, B.D., Muglia, L.J., Schappe, R.S., Gough, A.H., and Langmore, J.P. (1986). Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys J* 49, 233-248.

Witt, O., Albig, W., and Doenecke, D. (1996). Testis-specific expression of a novel human H3 histone gene. *Exp Cell Res* 229, 301-306.

Woodcock, C.L., and Ghosh, R.P. (2010). Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol* 2, a000596.

Workman, L.J., Abmayr, S.M. (2013). Fundamentals of chromatin.

Wu, J.Y., Ribar, T.J., Cummings, D.E., Burton, K.A., McKnight, G.S., and Means, A.R. (2000). Spermiogenesis and exchange of basic nuclear proteins are impaired in male germ cells lacking Camk4. *Nat Genet* 25, 448-452.

Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y. (2012). Lysine succinylation and lysine malonylation in histones. *Mol Cell Proteomics* 11, 100-107.

Xie, Z., Zhang, D., Chung, D., Tang, Z., Huang, H., Dai, L., Qi, S., Li, J., Colak, G., Chen, Y., *et al.* (2016). Metabolic Regulation of Gene Expression by Histone Lysine beta-Hydroxybutyrylation. *Mol Cell* 62, 194-206.

Xu, W., and Cooper, G.M. (1995). Identification of a candidate c-mos repressor that restricts transcription of germ cell-specific genes. *Mol Cell Biol* 15, 5369-5375.

Yan, W., and McCarrey, J.R. (2009). Sex chromosome inactivation in the male. *Epigenetics* 4, 452-456.

Yang, P., Wu, W., and Macfarlan, T.S. (2014). Maternal histone variants and their chaperones promote paternal genome activation and boost somatic cell reprogramming. *Bioessays* 37, 52-59.

Yoh, S.M., Cho, H., Pickle, L., Evans, R.M., and Jones, K.A. (2007). The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev* 21, 160-174.

Youdell, M.L., Kizer, K.O., Kisseleva-Romanova, E., Fuchs, S.M., Duro, E., Strahl, B.D., and Mellor, J. (2008). Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol Cell Biol* 28, 4915-4926.

Yu, Y.E., Zhang, Y., Unni, E., Shirley, C.R., Deng, J.M., Russell, L.D., Weil, M.M., Behringer, R.R., and Meistrich, M.L. (2000). Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. *Proc Natl Acad Sci U S A* 97, 4683-4688.

Zhang, K., Chen, Y., Zhang, Z., and Zhao, Y. (2009). Identification and verification of lysine propionylation and butyrylation in yeast core histones using PTMap software. *J Proteome Res* 8, 900-906.

Zhang, Q., Zeng, L., Zhao, C., Ju, Y., Konuma, T., and Zhou, M.M. (2016). Structural Insights into Histone Crotonyl-Lysine Recognition by the AF9 YEATS Domain. *Structure* 24, 1606-1612.

Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., and Zhao, Y. (2011). Identification of lysine succinylation as a new post-translational modification. *Nat Chem Biol* 7, 58-63.

Zhao, D., Guan, H., Zhao, S., Mi, W., Wen, H., Li, Y., Zhao, Y., Allis, C.D., Shi, X., and Li, H. (2016). YEATS2 is a selective histone crotonylation reader. *Cell Res* 26, 629-632.

Zhao, M., Shirley, C.R., Hayashi, S., Marcon, L., Mohapatra, B., Suganuma, R., Behringer, R.R., Boissonneault, G., Yanagimachi, R., and Meistrich, M.L. (2004a). Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. *Genesis* 38, 200-213.

Zhao, M., Shirley, C.R., Mounsey, S., and Meistrich, M.L. (2004b). Nucleoprotein transitions during spermiogenesis in mice with transition nuclear protein Tnp1 and Tnp2 mutations. *Biol Reprod* 71, 1016-1025.

Zhao, M., Shirley, C.R., Yu, Y.E., Mohapatra, B., Zhang, Y., Unni, E., Deng, J.M., Arango, N.A., Terry, N.H., Weil, M.M., *et al.* (2001). Targeted disruption of the transition protein 2 gene affects sperm chromatin structure and reduces fertility in mice. *Mol Cell Biol* 21, 7243-7255.

Zhong, J., Peters, A.H., Lee, K., and Braun, R.E. (1999). A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. *Nat Genet* 22, 171-174.

Annexes

Research Highlights

Highlights from the latest articles in epigenomics



Transgenerational inheritance of chromatin states

Evaluation of: Raychaudhuri N, Dubruille R, Orsi GA *et al.* Transgenerational propagation and quantitative maintenance of paternal centromeres depends on Cid/Cenp-A presence in *Drosophila* sperm. *PLoS Biol.* 10(12), e1001434 (2012).

The genome-wide histone-to-protamine transition in the post-meiotic phases of spermatogenesis in many organisms, including *Drosophila*, involves one of the most spectacular genome reorganizations [1], derived by mechanisms that are set in the preceding stages [2]. This transition in genome packaging and, most importantly, the disappearance of histones, raises an essential issue relative to the nature of the final mature male gamete epigenome and the specific associated elements that are transmitted to the embryo.

Early investigations have revealed that in mammals, mature sperm cells partially retain elements of chromatin. For instance, in humans, it is believed that between 5 and 15% of histones survive histone-to-protamine exchange. The recent genome-wide mapping of residual nucleosomes and histone marks in human sperm cells revealed the existence of a possible sperm epigenome organized by the selective association of defined genomic regions – that is, developmental and imprinted genes, with canonical histones [3,4]. These studies also imply that these specific organizational states of the male genome would be inherited by the embryo and could have a functional impact on gene expression during early development. However, up to now, no experiments had been designated to test this hypothesis and no data were available to support this fundamental claim.

Among the somatic chromatin elements that were considered to escape histone

removal was the centromeric histone H3 variant, Cenp-A [5]. However, in this case there was no data explaining the reason behind the retention of this specific histone. Very recently, a breakthrough came out of the study of the male germ cell Cenp-A homolog, Cid, in *Drosophila*, whose male germ cells also undergo a genome-wide histone-to-protamine replacement [6]. These data show not only that the *Drosophila* Cenp-A (Cid) almost completely escapes histone-to-protamine replacement, but also that its maintenance in the sperm is of critical importance. Indeed, the authors found that Cid elimination from sperm results in the loss of paternal chromosomes during the initial syncytial cycles of early embryogenesis, hence demonstrating that the Cid inherited along with the male genome is indispensable for the paternal genome contribution to the embryo, ensuring its propagation. This constitutes the first demonstration of the essential role of a retained histone variant in the male gametes, and may also shed light on the possible importance of other histone variant maintenance in the sperm.

References

- 1 Gaucher J, Reynoird N, Montellier E *et al.* From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J.* 277(3), 599–604 (2010).
- 2 Gaucher J, Boussouar F, Montellier E *et al.* Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J.* 31(19), 3809–3820 (2012).
- 3 Brykczynska U, Hisano M, Erkek S *et al.* Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat. Struct. Mol. Biol.* 17(6), 679–687 (2010).
- 4 Hammoud SS, Nix DA, Zhang H *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 460(7254), 473–478 (2009).

Hitoshi Shiota¹, Afsaneh Goudarzi¹, Sophie Rousseaux¹ & Saadi Khochbin^{*1}

¹INSERM, U823, Université Joseph Fourier – Grenoble 1, Institut Albert Bonniot, Grenoble, F-38700, France

*Author for correspondence: khochbin@ujf-grenoble.fr

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.



- 5 Palmer DK, O'Day K, Margolis RL. The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. *Chromosoma* 100(1), 32–36 (1990).
- 6 Raychaudhuri N, Dubruielle R, Orsi GA *et al.* Transgenerational propagation and quantitative maintenance of paternal centromeres depends on Cid/Cenp-A presence in *Drosophila* sperm. *PLoS Biol.* 10(12), e1001434 (2012).

Shielding Cenp-A removal in post-meiotic male cells

Evaluation of: Dunleavy EM, Beier NL, Gorgescu W *et al.* The cell cycle timing of centromeric chromatin assembly in *Drosophila* meiosis is distinct from mitosis yet requires CAL1 and CENP-C. *PLoS Biol.* 10(12), e1001460 (2012).

The focus of this particular discussion is the discovery by Dunleavy and colleagues of a significant Cid assembly in male post-meiotic cells in addition to that observed during meiosis [1]. It is of note that, as mentioned in the above discussion, in *Drosophila*, like in many mammalian species, the post-meiotic haploid cells, named spermatids, prepare for a genome-wide histone replacement [2]. It is therefore of interest that Cid that escapes histone removal [1,3] is at least partially assembled in haploid cells prior to histone displacement. There is another example of such an incorporation of a histone variant before the process of general histone eviction, observed recently during mouse spermatogenesis, which concerns a H2A variant [4]. Indeed, a testis-specific H2A variant, H2AL2, was found to be incorporated into chromatin just before the accumulation of protamines. Remarkably, with regard to Cenp-A/Cid, this H2AL2 is also retained in the mature

sperm and specifically organizes the pericentric regions [4].

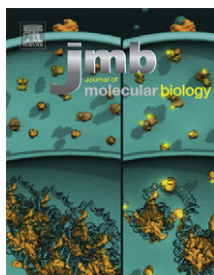
It is therefore possible that the observed assembly of Cid in nonreplicative haploid spermatids is associated with the assembly of a histone-like or a specific DNA-packaging protein on the pericentric regions. Similarly, the assembly of H2AL2 in mouse spermatids could also be associated with that of Cenp-A. This would imply that, in mammals as well as in the fly sperms, centromeric and pericentric regions would both be associated with specific DNA-associated proteins. If this hypothesis is verified, the next question would be whether the repackaging of pericentric regions could have a role in protecting Cid/Cenp-A against histone removal. Since the mouse H2AL2 present in pericentric regions resists histone displacement, one can propose that its deposition around centromeres could 'shield' Cenp-A against the wave of nucleosomes disassembly affecting the rest of the genome.

In the light of these discussions, one could also speculate not only that centromeric histones could be inherited and are indispensable to the male genome propagation [1,3], but also that pericentric regions bear specific chromatin organization of equally important functions. However, in

contrast to Cid, the possible protective role of pericentric chromatin should be critical only in sperms since, after fertilization in the mouse, H2AL2 is rapidly removed and replaced by maternal histones [5].

References

- 1 Dunleavy EM, Beier NL, Gorgescu W *et al.* The cell cycle timing of centromeric chromatin assembly in *Drosophila* meiosis is distinct from mitosis yet requires CAL1 and CENP-C. *PLoS Biol.* 10(12), e1001460 (2012).
- 2 Gaucher J, Reynoird N, Montellier E *et al.* From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J.* 277(3), 599–604 (2010).
- 3 Raychaudhuri N, Dubruielle R, Orsi GA *et al.* Transgenerational propagation and quantitative maintenance of paternal centromeres depends on cid/cenp-a presence in *Drosophila* sperm. *PLoS Biol.* 10(12), e1001434 (2012).
- 4 Govin J, Escoffier E, Rousseaux S *et al.* Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J. Cell Biol.* 176(3), 283–294 (2007).
- 5 Wu F, Caron C, De Robertis C *et al.* Testis-specific histone variants H2AL1/2 rapidly disappear from paternal heterochromatin after fertilization. *J. Reprod. Dev.* 54(6), 413–417 (2008).



Genome-Scale Acetylation-Dependent Histone Eviction during Spermatogenesis

Afsaneh Goudarzi, Hitoshi Shiota, Sophie Rousseaux and Saadi Khochbin

Institut National de la Santé et de la Recherche Médicale U823 and Université Grenoble Alpes Institut Albert Bonniot, Grenoble F-38700, France

Correspondence to Saadi Khochbin: khochbin@ujf-grenoble.fr

<http://dx.doi.org/10.1016/j.jmb.2014.02.023>

Edited by S. Khorasanizadeh

Abstract

A genome-wide histone hyperacetylation is known to occur in the absence of transcription in haploid male germ cells, spermatids, before and during the global histone eviction and their replacement by non-histone DNA-packaging proteins. Although the occurrence of this histone hyperacetylation has been correlated with histone removal for a long time, the underlying mechanisms have remained largely obscure. Important recent discoveries have not only shed light on how histone acetylation could drive a subsequent transformation in genome organization but also revealed that the associated nucleosome dismantlement is a multi-step process, requiring the contribution of histone variants, critical destabilizing histone modifications and chromatin readers, including Brdt, working together to achieve the full packaging of the male genome, indispensable for the propagation of life.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

In both female and male organisms, germ cells become distinct early in life and pursue a specific differentiation process with the aim of generating cells that will survive the parent organisms and perpetuate the species. Within this setting, male germ cells need to develop very unique cellular functions to fulfil their final task, which, in their particular case, includes the transport the genome out of the parent organism. This genome delivery “mission” of the male germ cells imposes specific constraints to their genome that are not seen in any other cells in an organism. Although both male and female germ cells undergo meiosis to end up with a single-copy genome, in the specific case of male germ cells, meiosis also tolls the bells for the start of what can be considered as the most spectacular and large-scale known chromatin remodeling. In mammals, this chromatin remodeling is characterized by an almost genome-wide histone removal and their step-wise replacement by small basic proteins. During this process, histones are first replaced by transition proteins (TPs), which are themselves replaced by protamines (Prms) [1,2]. Prms ensure

the full protection of the genome against the harsh environment encountered by spermatozoa during their journey to reach the female gametes.

These dramatic changes within the essence of genome organization require the action of specialized factors that are themselves expressed exclusively in the male germ cells. They are encoded by groups of genes expressed in successive waves as soon as the progenitor spermatogenic cells, spermatogonia, become committed into spermatogenic differentiation and give rise to spermatocytes entering their first meiotic division [3–6]. During meiosis and following the generation of the haploid post-meiotic cells, named spermatids, a particular gene expression program is turned on, leading to the production of specific factors that will target the genome and drive its transformation into an inert transportable package. Some elements of these new genome organizations have been known for a long time. They include DNA-packaging proteins, histone variants, TPs and Prms that are expressed and that act sequentially following a precise and well-established timing [1]. However, critical information on specific factors managing these elements is still missing. The molecular basis of the precise function

of specific histone variants in haploid cells and the machineries removing histones and assembling TPs and Prms are awaiting to be defined. Additionally, it is also important to understand how, in the context of general histone eviction, some histones escape replacement and are finally retained in the mature spermatozoa [7–9].

The situation has changed in the past few years with the publication of several key investigations that now shed light on how a nucleosome-based genome loses its fundamental organizing structural unit and adopts a new packaging principle, which is apt to be recognized and taken in charge by the maternal genome reprogramming factors in the egg. These new findings will be described and critically commented here, and currently available working models for the molecular basis of histone-to-protamine transition will be presented and discussed.

Does Histone Hyperacetylation Signal Histone Removal?

In different species, the occurrence of a genome-wide histone hyperacetylation has been observed before and during the disappearance of histones and their replacement by TPs (for review, see Refs. [10] and [11]). These correlative observations and the conserved nature of the association between histone acetylation and removal in different species brought the idea that there could be a causal relationship between the two events. Following this hypothesis, early investigations revealed that histone hyperacetylation could facilitate histone eviction *in vitro* upon the mere incubation of hyperacetylated chromatin with Prms [12–14]. However, although these experiments revealed how histone hyperacetylation could be associated with an enhanced histone exchange, the underlying *in vivo* mechanisms are unlikely to follow such a simple process since, in somatic cells, both the assembly and removal of histones involve specialized machineries [15,16].

Following the discovery of bromodomains as a structural determinant capable of specifically recognizing an acetylated histone tail in 1999 [17], there emerged the hypothesis that putative testis-specific bromodomain-containing factors could in fact recognize hyperacetylated histones and mediate their removal [18]. An analysis of public sequence database for testis-specific genes encoding bromodomain-containing proteins pointed to Brdt, a double-bromodomain-containing factor of BET family, exclusively expressed in spermatogenic cells, as a promising candidate [18]. The early functional investigations of Brdt revealed its capacity to specifically act on hyperacetylated chromatin both *in vitro* and *in vivo* [18,19]. This hypothesis received additional supports from the observation of the phenotype of mice expressing a truncated form of

Brdt deleted for its first bromodomain (BD1). Indeed, *Brdt*^{ΔBD1/ΔBD1} male mice were infertile with dramatic defects observed in elongating spermatids [20] at the time when histone hyperacetylation is normally expected to occur. Although the first published analyses of these mice did not consider histone hyperacetylation and their removal in spermatogenic cells [20], the reported blockade of spermatogenesis at the elongation step designated Brdt as a critical factor in managing the biology of elongating spermatids.

The strongest support for the hypothesis presented above, involving Brdt in the removal of hyperacetylated histones, came from the structural characterization of Brdt's two bromodomains. These studies showed that Brdt's BD1 histone binding requires simultaneous acetylation of K5 and K8 of histone H4 tail [21]. Remarkably, the conclusions of these structural studies were confirmed *in vivo* using an original Förster resonance energy transfer probe encompassing Brdt's two bromodomains fused N-terminal to histone H4. This approach demonstrated that the acetylation-dependent interaction of Brdt with histone H4 tail requires the presence of acetylated lysines at both positions 5 and 8 of H4 also in an *in vivo* setting [22].

The abovementioned structural studies, besides describing the first example of a combinatorial reading of histone marks by a single domain, also revealed that Brdt's BD1 is the first identified module specific for the recognition of hyperacetylated histone H4. Indeed, the simultaneous acetylation of H4K5 and K8 is considered as an indicator of the full acetylation of the four acetyl-acceptor lysines of the histone H4 N-terminal tail [23–25].

Taking into account these structural data on the recognition of hyperacetylated histone H4 by Brdt and the fact that a genome-wide histone H4 hyperacetylation occurs only in elongating spermatids, one could therefore easily predict an outstanding action of Brdt on hyperacetylated histones in elongating spermatids, in agreement with the reported impairment of spermatogenesis that precisely occurs in these cells in the absence of Brdt's BD1 [21].

The investigation of histone removal and the assembly of TPs and Prms in elongating spermatids expressing the BD1-less Brdt mutant finally demonstrated the specific involvement of Brdt in the process of histone-to-Prm exchange. Indeed, in these cells, although histone hyperacetylation occurs normally and TP and Prms are synthesized at the expected timing, the removal of hyperacetylated histones does not occur and TPs and Prms are not assembled and remain in the perinuclear regions of the cells [5].

These studies also involving other mouse models for Brdt, including a complete deletion of *Brdt* gene and the use of a dominant-negative non-functional Brdt, also showed that Brdt is a multifunctional factor and that, depending on the spermatogenic cell types,

its activities range from transcriptional regulation in early stages [5,26] to the discussed exchange of histones with TPs in elongating spermatids [5].

Mechanisms Underlying Acetylated Histone Eviction

All the arguments and data described above strongly support the role of Brdt, mainly its first bromodomain, in the exchange of acetylated histones by TPs in elongating spermatids. However, these data, besides revealing the ability of Brdt to recognize hyperacetylated histone H4 [21,22] and the requirement of its first bromodomain to mediate histone-to-TP exchange [5], did not show how Brdt is actually managing the exchange of nucleosomal histones.

From the early functional data on Brdt, it is possible to propose a working model explaining how Brdt is in fact controlling histone removal and their exchange with TPs. Indeed, a work published in 2003 already showed that Brdt is capable of inducing a dramatic acetylation-dependent chromatin compaction [18] leading to the release and degradation of free histones (in this case, it was shown for linker histone H1^o) in the nucleoplasm after chromatin compaction [19]. These studies also showed that the chromatin-compacting activity of Brdt requires not only the bromodomains but also the regions flanking the two bromodomains [18]. Based on these data, a working model can be proposed where Brdt and the acetylation-dependent chromatin compaction would constitute a step-wise process initially involving the recognition and binding of hyperacetylated chromatin by Brdt, followed by Brdt–Brdt interactions, finally leading to the cross-link of adjacent nucleosomes and chromatin compaction. In support of this model, the acetylation-dependent Brdt–Brdt interaction was demonstrated using a Förster resonance energy transfer approach [5].

The question underlying this observation is how does this Brdt-mediated chromatin compaction induce histone eviction?

One possibility is that the nucleosomes, already highly destabilized by the incorporation of histone variants and histone post-translational modifications (PTMs), would “squeeze out” their histones through the compacting action of Brdt. The evicted Brdt-bound histones could then be taken in charge by chaperons and a degradation machinery, while TP-associated chaperones would assemble the new DNA-packaging structures.

However, this scheme does not give any indication about the fate of hyperacetylated histones. A probable step following the Brdt-dependent exchange of acetylated histones by TPs could be a direct degradation of these acetylated histones by a specialized 20S testis-specific proteasome, associated with the proteasome activator PA200, named

“spermatoproteasome”. Indeed, a recently published report proposes that the PA200 regulatory module is capable of directly recognizing acetylated histones through a bromodomain-like module and directing their ubiquitin-independent degradation [27].

However, this interpretation is challenged by the observation from another published work clearly indicating that the spermatoproteasome cannot degrade acetylated nucleosomal histones in the elongating spermatid chromatin. Indeed, considering the data discussed above on the role of Brdt's first bromodomain, it clearly appears that, in the absence of Brdt's BD1, despite the occurrence of histone hyperacetylation, histone removal is impaired and cells retain acetylated histones until their death [5]. Additionally, histone removal is also severely disturbed in elongating spermatids expressing a modified form of the testis-specific histone H2B variant, TH2B. In this latter case also, histone hyperacetylation occurs in elongating spermatids in the right timing and in the expected proportion, but again, histone removal does not follow their hyperacetylation [6]. Therefore, a possible explanation for these observations is that acetylated histones should first be removed and then passed onto the spermatoproteasome for degradation.

It is also important to note that the possibility of the ubiquitin-dependent degradation of core histones in spermatids could not be ruled out since previous studies have demonstrated the occurrence of mono-ubiquitination and polyubiquitination of histones H2B, H3 and H3T in the rat elongating spermatids [28]. Along this line, investigators have also identified a testis-specific E2 (UBC4-testis) that may work with an E3 named E3^{histone} because of its ability to ubiquitinate core histones, but the involvement of this machinery in histone degradation in elongating spermatids was not demonstrated [29–31].

Additionally, both in mice lacking PA200 and in drosophila lacking a proteasome subunit, Prosa6 T, only a delay in histone removal is observed. Indeed, in the absence of PA200 condensing spermatids devoid of histones could be observed [27], and in the elongating spermatids of Prosa6 T-less drosophila, Prm assembly finally takes place [32], strongly suggesting that a proteasome-independent histone degradation system should also be operational in spermatogenic cells and ensure the degradation of the major part of the removed histones. Altogether, it seems clear that histone acetylation is *a priori* a signal for histone removal and histone degradation could be a consequence of this process.

Histone Degradation Requires Facilitated Histone Removal

In at least two cases, a normal histone hyperacetylation occurs despite an impairment of histone

removal in elongating spermatids, which strongly suggests that histone hyperacetylation by itself is not enough to mediate the subsequent nucleosome dismantlement. Indeed, as mentioned above, in the absence of Brdt's BD1 [5] or in the presence of a C-terminally modified TH2B [6], histone hyperacetylation occurs without inducing any subsequent histone removal.

These observations strongly suggest that acetylation-dependent histone removal is a multi-step process and involves various and independent factors. Based on the literature, it can be argued that at least two independent events cooperate to mediate the nucleosome disassembly, which is itself also a multi-step process.

One of the first determinants in nucleosome dismantlement in elongating spermatids is a prior drastic destabilization of nucleosomes. Spermatogenic cells express almost all the histone variants encoded by the genome including a variety of testis-specific members exclusively expressed in spermatogenic cells [11,33]. Several reports investigating the biochemical properties of these testis-specific variants encompassing H2A, H2B and H3 variants all support their capacity to create a nucleosome that is less stable than the ones containing canonical histones. Indeed, data on testis-specific histones H3T, TH2B and H2AL2 and H2Abbd show that the presence of these histones in a nucleosome is associated with increased nucleosome instability [33–36]. Since these histones are massively synthesized and incorporated in stages preceding histone removal, one could argue that their synthesis and assembly in chromatin is a requisite for a subsequent nucleosome disassembly. A recent publication reporting the genetic manipulation of the testis-specific H2B variant, TH2B, adding a C-terminal extension to the protein, revealed that, despite the occurrence of histone hyperacetylation and the presence of functional Brdt, histone removal is impaired in elongating spermatids. This report demonstrated that this interference with histone removal is specific to elongating spermatids, since the same TH2B molecule could be normally assembled and later removed in early embryos [6].

The same report also reveals another mechanism that presides at the dismantlement of nucleosomes in spermatogenic cells, involving PTMs at critical positions in histones, which directly affect the stability of nucleosomes. Indeed, in the total absence of TH2B, the canonical H2B can compensate for the lack of TH2B and histone replacement normally occurs, but the nucleosomes undergoing this disassembly contain histones that bear specific PTMs at several critical histone–DNA and histone–histone interaction points. These histone PTMs would confer to the nucleosomes the instability that is normally associated with TH2B recruitment [6].

The uncovering of this mechanism also argues in favor of the existence of other nucleosome-destabilizing mechanisms involving machineries capable of modifying histones within the histone-fold regions [37]. Therefore, it is expected that, in elongating spermatids, in parallel with the well-known acetylations of the histone tails, lysines present in the core regions of histones also show an increased acetylation aiming at inducing increased nucleosome instability. Such situations have already been described in somatic cells, where acetylation of critical histone-fold lysines, mainly H3K56 and H3K122, has clearly been shown to induce nucleosome instability [38]. Although the occurrence of H3K122 and H3K56 acetylation has not been studied in spermatogenic cells, a very recent investigation of a newly discovered histone H3K64 acetylation, also inducing nucleosome instability and facilitating histone eviction, revealed its genome-wide occurrence at the time of histone exchange in elongating spermatids [39].

In summary, taking into account all the data discussed above, it is possible to propose the following temporal scheme for histone removal in elongating spermatids. First, canonical histones are replaced at a very large scale by testis-specific histone variants capable of creating less stable nucleosomes [6,33]. Second, a genome-wide histone hyperacetylation occurs with two distinct objectives: tail acetylation, which provides a surface for the recruitment of Brdt [5], and histone-fold acetylation, to further destabilize the nucleosomes [38]. The recognition of acetylated histone tails by Brdt would then initiate the process of histone removal, which can only occur if it has been previously “prepared” by the combined destabilizing actions of histone variants and acetylation of histone-fold regions. The removed histones could then eventually be degraded following the action of the spermatoproteasome and other yet unknown proteolytic machineries (Fig. 1).

Burning Questions

In the light of the literature and especially of the data discussed above, histone hyperacetylation in elongating spermatids appears as a determinant event in the process of histone-to-Prm transition. An important unsolved question is whether this histone hyperacetylation is an active process or rather an indirect consequence of prior events.

Although today there is no clear response to this question, some hints from published reports could support the proposition of a working model.

The simplest explanation for the increase in histone acetylation could merely be a decrease in the amount of cellular histone deacetylase (HDAC) at the time of spermatid elongation. This is supported

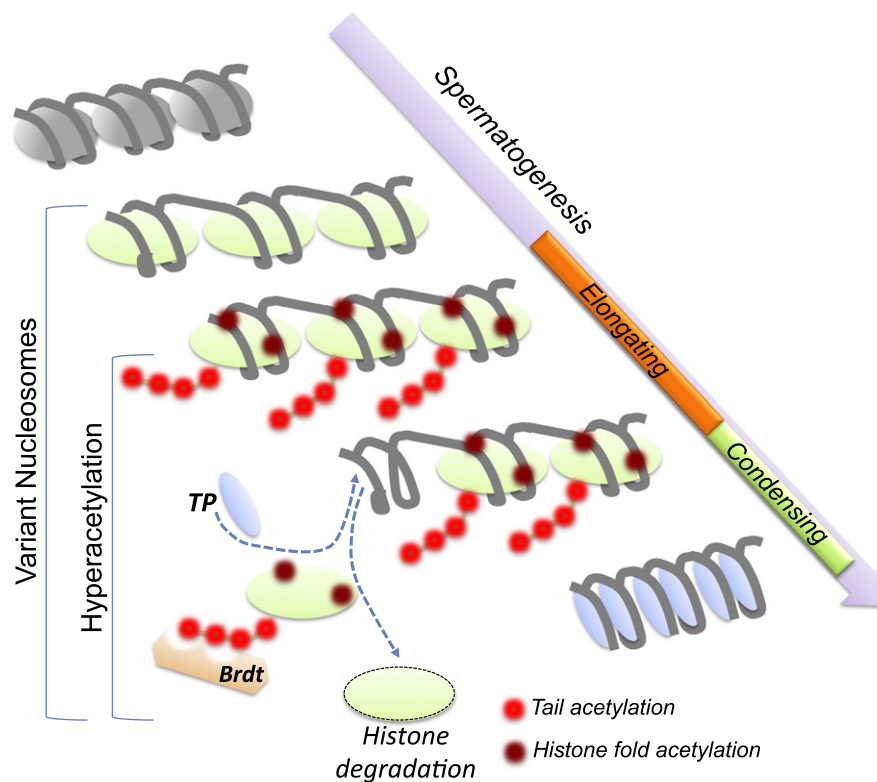


Fig. 1. A working model for the histone-to-transition proteins exchange in elongating spermatids. Global incorporation of various H2A, H2B and H3 histone variants creates highly unstable nucleosomes, which then undergo hyperacetylation both at their accessible N-terminal tails and within their histone-fold regions. Acetylation at critical lysines, such as H3K122 and H3K64, further destabilizes the nucleosomes, while tail acetylation generates a platform for the recruitment of Brdt. Brdt could then start the process of histone eviction and replacement by TPs. Evicted acetylated histones would then be degraded through the action of spermatoproteasome or/and yet unknown proteolytic systems.

by the observation of a dramatic decrease in the amounts of HDAC1 and HDAC2 observed by Western blot in fractionated elongating mouse spermatids [40]. The observed increase in histone acetylation in these cells could therefore be largely explained by the loss of the enzymes removing the acetyl group in these cells. Another argument in favor of a major role for HDACs in this process is the fact that histone hyperacetylation could be induced in round spermatids by a simple treatment of a spermatogenic cell suspension with the HDAC inhibitor TSA. Indeed, this treatment induced a dramatic histone hyperacetylation in round spermatids, which normally have hypoacetylated chromatin, whereas no remarkable change in acetylation could be observed once histone hyperacetylation starts in elongating cells [41], suggesting that there are no more HDACs to inhibit.

In the literature, some data also support the involvement of specific histone acetyltransferases (HATs) in the process of histone hyperacetylation. However, almost all the corresponding data are controversial and do not support the active involvement of a particular HAT in histone hyperacetylation.

First, a work published about a decade ago, identifying Cdy1 as a new HAT, proposed its involvement in the occurrence of histone hyperacetylation in elongating spermatid [42]. However, the HAT activity of Cdy1 specifically tested by other researchers could not be confirmed [40] and, since then, no other report on a HAT activity associated with Cdy1 has been published.

The study of the phenotype associated with the invalidation of *RNF8*, a gene encoding a histone E3 ubiquitin ligase, also led to the observation of a defective histone removal associated with a lack of histone hyperacetylation. The authors then attributed this lack of histone acetylation to a down-regulation of the Mof HAT, a known H4K16 HAT, in the *RNF8*^{-/-} spermatogenic cells, and suggested that this HAT could be a major player in mediating histone hyperacetylation [43]. However, here again using the same *RNF8*^{-/-} mice strain, another group did not observe any abnormality in histone H4K16 acetylation or in the process of histone-to-Prm exchange [44]. Finally, the analysis of the expression of two ubiquitous and generalist HATs, CBP and p300, in spermatogenic cells showed their important down-regulation in post-

meiotic cells, especially in elongating spermatids, when histone hyperacetylation occurs. Additionally, their further partial knockdown in post-meiotic cells by the targeted deletion of the corresponding genes did not reveal any alterations in the occurrence of histone acetylation in elongating spermatids [3].

All these observations suggest that the occurrence of histone hyperacetylation could only be a consequence of a decrease in the amount of cellular HDACs occurring prior to the exchange of histones by TPs and that cellular HATs present in these cells could ensure this histone hyperacetylation, without a need for additional and specific HATs.

An intriguing observation remains however the fact that, in elongating spermatids, acetylation is not the only PTM that globally affects histones prior to their removal. Indeed, phosphorylation of H2A.X [45], H2A ubiquitination [46] and histone poly-ADP ribosylation [47] are also reported to occur at almost the same timing as histone acetylation. Surprisingly, the discovery of new histone PTMs and the investigation of their occurrence in spermatogenic cells again pointed to elongating spermatids as a particular cell population showing a dramatic enhancement of these new modifications [48]. A question therefore arises: why do nearly all the histone PTMs investigated so far suddenly show an enhanced occurrence on the chromatin in elongating spermatids? As for histone acetylation, which increases as HDACs decrease, it is likely that these other PTMs are increased following prior events, such as the down-regulation of the enzymes involved in their removal from histones. Additionally, the opening of chromatin and histone removal may provide a window of opportunity for all the present histone PTM machineries to efficiently act on histones before their degradation.

Another important question is the functional implication of these genome-wide histone modifications occurring in parallel with acetylation. Some of these reported histone modifications could also be associated with the process of nucleosome dismantlement and histone eviction. For instance, histone poly-ADP ribosylation and H2A.X phosphorylation could be linked to the cell response to naturally occurring double-strand DNA break that accompanies histone eviction [45,49]. One can also question the role of the newly identified histone PTMs but an answer should await more investigations on the functional significance of these new PTMs.

Concluding Remarks

Histones have been largely studied as major elements in the control of gene expression and support for epigenetic mechanisms. However, de-

spite being very abundant cellular proteins and a focus of intense investigations, their metabolism is not very well understood. Although the synthesis of the canonical forms is coupled to DNA replication and follows well-defined rules, their degradation and the machineries involved have only recently become known [50–52].

A complicated matter in these studies is perhaps that histone turnover must be highly region dependent and variant dependent and that histone half-life in active chromatin and heterochromatin should be very different [53]. The almost genome-scale histone removal in elongating spermatids provides therefore a unique model to tackle mechanisms controlling histone turnover in cells. It could also guide our understanding of histone degradation in somatic cells. Indeed, the definition of the spermatoproteasome in spermatogenic cells, especially the role of the PA200 subunit in the recognition and direct degradation of acetylated histones, allowed the authors to show the occurrence of a similar mechanism in somatic cells degrading histones during DNA repair [27]. However, as discussed above, this particular system does not seem to degrade nucleosomal histones and a collaboration with histone removal machineries is probably necessary.

Finally, although we are advancing in the understanding of the genome-wide histone eviction in spermatogenic cells, it is not clear how histones associated with specific genomic loci could escape removal [7–9]. A recent study established a tight relationship between the histone retention and non-methylated CpG islands and histone H3 variants and their methylation at K4 and K27 [54]. However, the question remains on why factors such as Brdt and the spermatoproteasome do not act on these nucleosomes. One possibility is that histone H4 in these retained nucleosomes does not bear acetylation at K5 or K8 or both, considering that their simultaneous acetylation is absolutely necessary for Brdt binding [21]. A hint from the literature suggests that this may well be the case. Indeed, the analysis of H4 acetylation in the retained human mature sperm histones revealed hypoacetylation of H4K5 [55]. In the light of this observation, it could be proposed that regions escaping full acetylation of nucleosomal histone H4 may become protected against histone removal simply because they do not bear the code necessary for Brdt recognition.

In summary, although recent investigations brought an unprecedented insight into the molecular basis presiding at the dramatic transitions in the male genome packaging and the mechanisms discovered also shed some lights into the mechanisms underlying histone turnover in somatic cells, hundreds of questions still remain unanswered and years of investigations will be necessary to finally unravel the mysteries of the male genome packing.

Acknowledgements

Work in S.K. laboratory is supported by Agence Nationale de la Recherche EpiSperm2 grant and its cancer-oriented developments by Institut National Du Cancer. A.G. and H.S. are recipients of a Marie Curie Initial Training Network funded by European Commission (FP7-PEOPLE-2011-ITN and PITN-GA-289880).

Conflicts of Interest: All authors declare that they have no conflicts of interest.

Received 15 January 2014;

Received in revised form 26 February 2014;

Accepted 26 February 2014

Available online 5 March 2014

Keywords:

BET bromodomain;
acetylation;
Brdt;
histone variants;
histone eviction

Abbreviations used:

PTM, post-translational modification; HDAC, histone deacetylase; HAT, histone acetyltransferase.

References

- [1] Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S. From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 2009;277:599–604.
- [2] Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R. Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta* 2013. <http://dx.doi.org/10.1016/j.bbagr.2013.08.004> [Epub ahead of print].
- [3] Boussouar F, Goudarzi A, Buchou T, Shiota H, Barral S, Debernardi A, et al. A specific CBP/p300-dependent gene expression program drives the metabolic remodelling in late stages of spermatogenesis. *Andrology* 2014. <http://dx.doi.org/10.1111/j.2047-2927.2014.00184.x> [Epub ahead of print] [PMID: 24522976].
- [4] Chalmel F, Lardenois A, Primig M. Toward understanding the core meiotic transcriptome in mammals and its implications for somatic cancer. *Ann N Y Acad Sci* 2007;1120:1–15.
- [5] Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, Bertrand S, et al. Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* 2012;31:3809–20.
- [6] Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, et al. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev* 2013;27:1680–92.
- [7] Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, et al. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. *Genome Res* 2009;19:1338–49.
- [8] Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, et al. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol* 2010;17:679–87.
- [9] Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 2009;460:473–8.
- [10] Boussouar F, Rousseaux S, Khochbin S. A new insight into male genome reprogramming by histone variants and histone code. *Cell Cycle* 2008;7:3499–502.
- [11] Govin J, Caron C, Lestrat C, Rousseaux S, Khochbin S. The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur J Biochem* 2004;271:3459–69.
- [12] Marushige K, Marushige Y, Wong TK. Complete displacement of somatic histones during transformation of spermatid chromatin: a model experiment. *Biochemistry* 1976;15:2047–53.
- [13] Oliva R, Bazett-Jones D, Mezquita C, Dixon GH. Factors affecting nucleosome disassembly by protamines *in vitro*. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. *J Biol Chem* 1987;262:17016–25.
- [14] Oliva R, Mezquita C. Marked differences in the ability of distinct protamines to disassemble nucleosomal core particles *in vitro*. *Biochemistry* 1986;25:6508–11.
- [15] Burgess RJ, Zhang Z. Histone chaperones in nucleosome assembly and human disease. *Nat Struct Mol Biol* 2013;20:14–22.
- [16] Ray-Gallet D, Almouzni G. Nucleosome dynamics and histone variants. *Essays Biochem* 2010;48:75–87.
- [17] Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 1999;399:491–6.
- [18] Pivot-Pajot C, Caron C, Govin J, Vion A, Rousseaux S, Khochbin S. Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* 2003;23:5354–65.
- [19] Govin J, Lestrat C, Caron C, Pivot-Pajot C, Rousseaux S, Khochbin S. Histone acetylation-mediated chromatin compaction during mouse spermatogenesis. *Ernst Schering Res Found Workshop* 2006:155–72.
- [20] Shang E, Nickerson HD, Wen D, Wang X, Wolgemuth DJ. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development* 2007;134:3507–15.
- [21] Moriniere J, Rousseaux S, Steuerwald U, Soler-Lopez M, Curtet S, Vitte AL, et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 2009;461:664–8.
- [22] Sasaki K, Ito T, Nishino N, Khochbin S, Yoshida M. Real-time imaging of histone H4 hyperacetylation in living cells. *Proc Natl Acad Sci U S A* 2009;106:16257–62.
- [23] Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, et al. Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 2007;282:7641–55.
- [24] Thorne AW, Kmiecik D, Mitchelson K, Sautiere P, Crane-Robinson C. Patterns of histone acetylation. *Eur J Biochem* 1990;193:701–13.
- [25] Zhang K, Williams KE, Huang L, Yau P, Siino JS, Bradbury EM, et al. Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol Cell Proteomics* 2002;1:500–8.

- [26] Berkovits BD, Wang L, Guarnieri P, Wolgemuth DJ. The testis-specific double bromodomain-containing protein BRDT forms a complex with multiple spliceosome components and is required for mRNA splicing and 3'-UTR truncation in round spermatids. *Nucleic Acids Res* 2012;40:7162–75.
- [27] Qian MX, Pang Y, Liu CH, Haratake K, Du BY, Ji DY, et al. Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* 2013;153:1012–24.
- [28] Chen HY, Sun JM, Zhang Y, Davie JR, Meistrich ML. Ubiquitination of histone H3 in elongating spermatids of rat testes. *J Biol Chem* 1998;273:13165–9.
- [29] Bedard N, Hingamp P, Pang Z, Karaplis A, Morales C, Trasler J, et al. Mice lacking the UBC4-testis gene have a delay in postnatal testis development but normal spermatogenesis and fertility. *Mol Cell Biol* 2005;25:6346–54.
- [30] Liu Z, Miao D, Xia Q, Hermo L, Wing SS. Regulated expression of the ubiquitin protein ligase, E3(Histone)/LASU1/Mule/ARF-BP1/HUWE1, during spermatogenesis. *Dev Dyn* 2007;236:2889–98.
- [31] Liu Z, Oughtred R, Wing SS. Characterization of E3^{Histone}, a novel testis ubiquitin protein ligase which ubiquitinates histones. *Mol Cell Biol* 2005;25:2819–31.
- [32] Zhong L, Belote JM. The testis-specific proteasome subunit Prospha6T of *D. melanogaster* is required for individualization and nuclear maturation during spermatogenesis. *Development* 2007;134:3517–25.
- [33] Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thevenon J, et al. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J Cell Biol* 2007;176:283–94.
- [34] Gautier T, Abbott DW, Molla A, Verdel A, Ausio J, Dimitrov S. Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* 2004;5:715–20.
- [35] Li A, Maffey AH, Abbott WD, Conde e Silva N, Prunell A, Siino J, et al. Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). *Biochemistry* 2005;44:2529–35.
- [36] Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T, Hayashi-Takanaka Y, et al. Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* 2010;107:10454–9.
- [37] Govin J, Khochbin S. Histone variants and sensing of chromatin functional states. *Nucleus* 2014;4:438–42.
- [38] Tropberger P, Schneider R. Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat Struct Mol Biol* 2013;20:657–61.
- [39] Di Cerbo V, Mohn F, Ryan D, Montellier E, Kacem S, Tropberger P, et al. Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. *eLIFE* 2014 [in press].
- [40] Caron C, Pivot-Pajot C, van Grunsven LA, Col E, Lestrat C, Rousseaux S, et al. Cdy1: a new transcriptional co-repressor. *EMBO Rep* 2003;4:877–82.
- [41] Hazzouri M, Pivot-Pajot C, Faure AK, Usson Y, Pelletier R, Sele B, et al. Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur J Cell Biol* 2000;79:950–60.
- [42] Lahn BT, Tang ZL, Zhou J, Barndt RJ, Parvinen M, Allis CD, et al. Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. *Proc Natl Acad Sci U S A* 2002;99:8707–12.
- [43] Lu LY, Wu J, Ye L, Gavrilina GB, Saunders TL, Yu X. RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Dev Cell* 2010;18:371–84.
- [44] Sin HS, Barski A, Zhang F, Kartashov AV, Nussenzweig A, Chen J, et al. RNF8 regulates active epigenetic modifications and escape gene activation from inactive sex chromosomes in post-meiotic spermatids. *Genes Dev* 2012;26:2737–48.
- [45] Leduc F, Maquennehan V, Nkoma GB, Boissonneault G. DNA damage response during chromatin remodeling in elongating spermatids of mice. *Biol Reprod* 2008;78:324–32.
- [46] Baarends WM, Hoogerbrugge JW, Roest HP, Ooms M, Vreeburg J, Hoeijmakers JH, et al. Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. *Dev Biol* 1999;207:322–33.
- [47] Meyer-Ficca ML, Scherthan H, Burkle A, Meyer RG. Poly(ADP-ribosylation) during chromatin remodeling steps in rat spermiogenesis. *Chromosoma* 2005;114:67–74.
- [48] Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 2011;146:1016–28.
- [49] Meyer-Ficca ML, Ihara M, Lonchar JD, Meistrich ML, Austin CA, Min W, et al. Poly(ADP-ribose) metabolism is essential for proper nucleoprotein exchange during mouse spermiogenesis. *Biol Reprod* 2011;84:218–28.
- [50] Chen YS, Qiu XB. Transcription-coupled replacement of histones: degradation or recycling? *J Genet Genomics* 2012;39:575–80.
- [51] Cook AJ, Gurard-Levin ZA, Vassias I, Almouzni G. A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3-H4 in the histone supply chain. *Mol Cell* 2011;44:918–27.
- [52] Gunjan A, Paik J, Verreault A. The emergence of regulated histone proteolysis. *Curr Opin Genet Dev* 2006;16:112–8.
- [53] Toyama BH, Savas JN, Park SK, Harris MS, Ingolia NT, Yates JR, et al. Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell* 2013;154:971–82.
- [54] Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, et al. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat Struct Mol Biol* 2013;20:868–75.
- [55] van der Heijden GW, Derijck AA, Ramos L, Giele M, van der Vlag J, de Boer P. Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. *Dev Biol* 2006;298:458–69.