Crosstalk between histone modifications in *Saccharomyces cerevisiae*

Doctorate of Philosophy, University of Oxford

Françoise S. Howe

Merton College, Trinity Term 2012

Approximate word count: 48, 000

Crosstalk between histone modifications in *Saccharomyces cerevisiae*

A thesis submitted to the University of Oxford for the degree of Doctor of Philosophy
Françoise S. Howe, Merton College, Trinity Term 2012.

Abstract

The N-terminal tails of histone proteins protrude from the nucleosome core and are extensively post-translationally modified. These modifications are proposed to affect many DNA-based processes such as transcription, DNA replication and repair. Post-translational modifications on histone tails do not act independently but are subject to crosstalk. One example of crosstalk is on histone H3 between lysine 14 (H3K14) and trimethylated lysine 4 (H3K4me3), a modification found at the 5' end of most active or poised genes. In this work, Western blots and chromatin immunoprecipitation (ChIP) experiments show that different amino acid substitutions at histone H3 position 14 cause varying degrees of H3K4me3 loss, indicating that H3K14 is not essential for H3K4me3 but acts as a modulator of H3K4me3 levels. A neighbouring residue, H3P16 is also important for H3K4me3 and may operate in concert with H3K14 to control H3K4me3. These crosstalk pathways have gene-specific effects and the levels of H3K4me3 are influenced to different extents on genes that fall into functionally distinct classes. A model is proposed to explain how H3K14/H3P16 may exert these varying effects on H3K4me3 at individual genes. In addition to its ability to regulate H3K4me3, H3K14 also influences the levels of two modifications on H3K18, acetylation and monomethylation. A ChIP-sequencing experiment has shown that H3K18me1, a previously uncharacterised modification in *S. cerevisiae*, is widely distributed throughout the genome and correlates strongly with histone H3 levels. The potential for a functional acetyl/methyl switch at H3K18 is explored. Together, these data indicate that, with gene-specific effects, crosstalk between histone modifications may be even more complex than originally thought.

Acknowledgements

There are a number of people I would like to acknowledge for their contributions towards this thesis. The most important thanks must go to Jane Mellor, my supervisor, who has been very supportive throughout the last four years. David Clynes produced the dot blot in Figure 40 and also began the early work on this project. Anitha Nair constructed the strains and performed the Western blots in Figure 30, and is invaluable for the smooth running of the lab. Michael Youdell collected the yeast doubling time data in Figure 16. Between them, these people taught me everything I needed to know when I first arrived in the lab four years ago and made me feel immediately welcome. Ana Serra Barros produced the Northern blot in Figure 47 and Matthew Sale performed the Western blots in Figure 24 and pioneered the early work on the H3P16 project. Ivan Boubriak carried out the SPR experiments in Figure 22. Special thanks must go to Struan Murray for the initial analysis of the ChIP-seq data and other genome-wide analyses featured in this thesis.

Collaborators outside the Mellor lab also need to be acknowledged. The DNA sequencing for the ChIP-seq experiment was performed by the Wellcome Trust Centre for Human Genetics, Oxford and I would particularly like to thank Lorna Gregory and her team. The mass spectrometry in this project was performed in collaboration with Ben Thomas at the Proteomics Facility, Dunn School of Pathology. The synthetic histone H3 mono-, di- and trimethylated on lysine 18, calf thymus histones, recombinant H3 and the *RPH1* plasmids were kindly provided by Rob Klose. I would also like to thank Rob for his advice on antibody design, His_6-tagged protein purification and the *in vitro* demethylase assays. Mahesh Chandrasekharan (Utah) provided the *JHD2* plasmids. Hugh Patterton shared unpublished data on the unmodified and modified histone H3 tail molecular dynamics simulations, which helped add weight to the H3P16 story.

Lastly I would like to thank all those who have given me moral support throughout these last four years, especially my family. Additionally, the people in the Mellor lab have managed to make coming into work a truly enjoyable experience, for which I am very grateful.

Abbreviations

Abbreviation	Term in full
ac	acetylation
ALL	acute lymphoblastic leukaemia
AML	acute myeloid leukaemia
ATP	adenosine triphosphate
bp	base-pair
BSA	bovine serum albumin
ChIP	chromatin immunoprecipitation
ChIP-qPCR	chromatin immunoprecipitation-quantitative real-time PCR
ChIP-seq	chromatin immunoprecipitation-sequencing
COMPASS	complex of proteins associated with Set1
CPF	cleavage and polyadenylation factor
CSM	complete synthetic media
CTD	C-terminal domain
CUT	cryptic unstable transcript
DHFR	dihydrofolate reductase
dO_2	dissolved oxygen
EDTA	ethylene diamine tetra-acetic acid
ESR	environmental stress response
FKBP	FK506-binding protein
GO	gene ontology
H2Bub	H2B monoubiquitination
HA	haemagglutinin
HAR	H2A repression domain
HAT	histone acetyltransferase
HDAC	histone deacetylase
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
kDa	kilo-Dalton
KMT	lysine methyltransferase
me	methylation
me1	monomethylation
me2	dimethylation
me2a	asymmetric dimethylation
me2s	symmetric dimethylation
me3	trimethylation
MAPK	mitogen-activated protein kinase
MNase	micrococcal nuclease
mRNA	messenger RNA
NET-seq	native elongating transcript sequencing
ncRNA	non-coding RNA
NP-40	nonidet P-40
nt	nucleotide
OD	optical density
OX	oxidative phase of the yeast metabolic cycle
PHD	plant homeodomain
PIC	pre-initiation complex
PKA	protein kinase A
PPIase	peptidyl prolyl cis-trans isomerase
PRMT	protein arginine methyltransferase
qPCR	quantitative PCR

Abbreviation	Term in full
R/B	reductive building phase of the yeast metabolic cycle
R/C	reductive charging phase of the yeast metabolic cycle
RNA FISH	RNA fluorescence in situ hybridisation
RPG	ribosomal protein gene
RRM	RNA recognition motif
SAGA	Spt-Ada-Gcn5 acetyltransferase
SAP	sense-antisense pair
SEC	super elongation complex
SLIK	SAGA-like
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
SUMO	small ubiquitin-related modifier
SUT	stable unannotated transcript
TORC	target of rapamycin complex
ub	ubiquitination
WT	wildtype
XUT	Xrn1-senstive unstable transcript
YMC	yeast metabolic cycle
YPD	yeast extract – bactopeptone – glucose
YPG	yeast extract – bactopeptone – galactose

In this thesis, 'yeast' will be used interchangeably with *S. cerevisiae*.

Table of Contents

Table of Figures

Chapter 1.

Introduction

1. Introduction

1.1. Gene expression is a highly regulated process

Gene expression, the synthesis of proteins using a DNA-based code, is a fundamental cellular process in all organisms. Misregulation of this process in higher eukaryotes can result in developmental defects and diseases such as cancer, demonstrating its importance. During gene expression, a gene is transcribed to produce a messenger RNA (mRNA), which is then exported from the nucleus and translated by the ribosomes to produce a polypeptide. This polypeptide is then folded and may be subject to post-translational modifications. At each of these stages, the process of gene expression can be controlled.

There are many circumstances that require either small- or large-scale changes in gene expression. In yeast these include temporal changes in gene expression during the cell cycle (Cho et al. 1998; Spellman et al. 1998; Granovskaia et al. 2010) or the yeast metabolic cycle (Tu, BP et al. 2005, 2007) and changes in response to alterations in the environment. For example, roughly 10 % of yeast genes are commonly induced or repressed in response to a variety of environmental stresses, allowing the cells to adapt to the change in environment and increase their chance of survival (Gasch et al. 2000; Causton et al. 2001). At the other end of the scale, changes in nutrient conditions can result in the induction of a small subset of genes specifically required for that condition, such as the induction of the *GAL* genes in response to the absence of glucose and the presence of galactose.

1.2. Transcription

Transcription, as the first step in gene expression, is a key point of control. DNA is transcribed by three RNA polymerases in eukaryotes, RNA polymerase I-III. Protein-coding genes are transcribed by RNA polymerase II and gene transcription can be broken down into three stages: initiation,

elongation and termination (reviewed by Hahn and Young 2011). During transcription initiation, the subunits in the pre-initiation complex (PIC) are recruited to a gene promoter by gene-specific transcription factors. Once correctly assembled, the transcription machinery can begin to transcribe the gene. The transformation from an initiating polymerase to a processive polymerase signifies the start of transcription elongation. Transcription termination is accompanied by processing of the RNA 3' end to produce an mRNA molecule.

This transcription cycle is in part controlled by phosphorylation of the largest RNA polymerase II subunit, Rpb1 on its C-terminal domain (CTD) (reviewed by Buratowski 2009). This domain consists of multiple copies of the heptapeptide repeat, YSPTSPS. The number of repeats varies between organisms and in yeast, the copy number is 26 (Allison et al. 1985). RNA polymerase II is unphosphorylated when recruited to a promoter. At initiation, the fifth serine in the CTD repeat is phosphorylated by the kinase Kin28 within TFIIH (Feaver et al. 1994). Serine 7 is also phosphorylated by TFIIH-associated Kin28 and has a 5' localisation (Akhtar et al. 2009; Kim, M et al. 2009). Levels of serine 5 phosphorylation drop after transcription of the first few hundred base-pairs (Komarnitsky 2000) and serine 2 is phosphorylated by Ctk1 (Cho et al. 2001) and Bur1 (Liu et al. 2009; Qiu et al. 2009) during transcription elongation. Recently, phosphorylation of tyrosine, the first residue in the heptapeptide repeat has been identified and this phosphorylation peaks prior to the 3' end of the gene (Mayer et al. 2012). Dephosphorylation of the tail by the phosphatases Rtr1 (serine 5) (Mosley et al. 2009), Ssu72 (serines 5 and 7) (Krishnamurthy et al. 2004; Bataille et al. 2012; Zhang et al. 2012) and Fcp1 (serine 2) (Cho et al. 2001) enables recycling of the polymerase for another round of transcription. This cycle of phosphorylation can influence transcription by the recruitment of phospho-specific binding factors such as the mRNA capping machinery (Komarnitsky 2000; Rodriguez et al. 2000; Schroeder 2000) and 3' end-processing factors (Licatalosi et al. 2002; Ahn et al. 2004) to the RNA polymerase CTD phosphorylated on serines 5 and 2 respectively. Tyrosine phosphorylation suppresses the binding of the termination

factors, Nrd1, Pcf11 and Rtt103 to phosphorylated serine 2 and promotes the binding of the transcription elongation factor, Spt6, thus repressing transcription termination until levels of tyrosine phosphorylation drop just before the 3' end of the gene (Mayer et al. 2012). The function of serine 7 phosphorylation is unknown (Akhtar et al. 2009; Kim, M et al. 2009).

The traditional view of transcription, the process by which RNA polymerase reads the DNA sequence flanked by a defined gene promoter and terminator to produce a functional RNA molecule that is then translated into proteins in the cytoplasm or that fulfils structural or catalytic roles, has been challenged in recent years. Studies performed with strains defective in components of the RNA degradation machinery, such as Rrp6, a catalytic component of the nuclear exosome, have indicated that the genome is subject to widespread pervasive transcription, much of which produces 'non-functional' or non-coding RNAs (ncRNAs) that are rapidly degraded (Wyers et al. 2005; David et al. 2006; Steinmetz et al. 2006). Importantly, the production of these transcripts in wildtype yeast has been validated by strand-specific mapping of nascent transcripts using native elongating transcript sequencing (NET-seq) (Churchman and Weissman 2011). These ncRNAs are produced as a result of inter- and intragenic transcription, the latter can occur in the same direction as coding gene transcription (sense) or in the opposite, antisense direction. There are several classes of ncRNAs, defined by their stability and sensitivity to certain ribonucleases. Cryptic unstable transcripts (CUTs) can only be detected in the absence of functional RNA degradation machinery (Neil et al. 2009; Xu et al. 2009). These transcripts are capped, short, with an average length of 200-800 nt and are terminated by an Nrd1-dependent pathway (Arigo et al. 2006b; Thiebaut et al. 2006) that directs the RNAs to be polyadenylated by the TRAMP complex and degraded by the exosome (LaCava et al. 2005; Vanácová et al. 2005; Vasiljeva and Buratowski 2006; Wyers et al. 2005). In contrast, stable unannotated transcripts (SUTs) can be detected at low levels in wildtype yeast (Xu et al. 2009), perhaps due to their termination by a different pathway (Marquardt et al. 2011). A third class, Xrn1-sensitive unstable

transcripts (XUTs) are sensitive to the cytoplasmic 5' to 3' exonuclease Xrn1 (also called Kem1) (Van Dijk et al. 2011). Whilst the majority of unannotated transcripts initiate at 5' ends of coding genes, reading in a sense or antisense direction, a substantial fraction originate from the 3' ends of genes and these are mostly transcribed in the antisense direction (Xu et al. 2009). These latter, overlapping transcripts give rise to sense-antisense pairs (SAPs) at a particular gene, which are often anti-regulated (Neil et al. 2009; Xu et al. 2009). The function for ncRNAs, or their transcription, is still unclear. Both repressive and activating roles have been described (reviewed by Berretta and Morillon 2009), although work in this laboratory may suggest that antisense transcription sets up an environment at gene promoters that gives scope for the regulation of gene transcription in either a positive or negative manner (S. Murray, unpub.).

Transcription can be regulated by a number of factors and processes. Basal transcription can occur *in vitro* in the absence of transcription factors but only on a non-nucleosomal DNA template (Juan et al. 1993). This demonstrates that a major obstacle to transcription is the fact that DNA is packaged into nucleosomes and gaining access to the underlying DNA is a highly regulated process.

1.3. Nucleosome structure

The nucleosome, the basic unit of chromatin, consists of 146 base-pairs of DNA wrapped 1.65 times around a core of histone proteins. The wrapping of DNA around basic histone proteins neutralises the negative charge of the DNA backbone phosphate groups and allows compaction of the DNA inside the nucleus. Nucleosomes are then further arranged to form higher order chromatin structures of various compactions: less compacted chromatin is termed euchromatin whereas highly compact chromatin is termed heterochromatin. A consequence of this packaging is that it allows the accessibility of the DNA to DNA-processing factors to be regulated. In this way, the nucleosome and higher order chromatin structure can influence many DNA-based processes

including transcription, DNA replication and DNA repair. The positions of nucleosomes along the DNA are in part determined by the underlying DNA sequence but also by a number of *trans*-acting factors such as chromatin remodellers, RNA polymerase and transcription factors (reviewed in Jansen and Verstrepen 2011).

Luger *et al.* published the X-ray crystal structure of the nucleosome at 2.8 Å resolution (Luger et al. 1997a), showing the atomic detail of the packing of two histone H3 and H4 homodimers with two histone H2A-H2B heterodimers to form the histone octamer (Figure 1). The DNA backbone interacts with the underlying histone proteins at many sites, causing it to have an irregular path over the octamer surface. Whilst the overall conformation of the nucleosome is conserved from single-celled eukaryotes up to metazoa, there are subtle differences in the amino acid sequence and therefore the structure of the histone octamer. One example of this is the interaction between H2A-H2B dimers. In the 2.8 Å X-ray structure determined using recombinantly-expressed histones from *Xenopus laevis*, the buried surface area between the H2A-H2B dimers is 150 $Å^2$ (Luger et al. 1997a). Because of sequence differences that alter some interactions, the buried surface area between the *S. cerevisiae* H2A-H2B dimers is reduced to 90 $Å^2$ (C L White et al. 2001). These and other weaker interactions may facilitate the disassembly of the nucleosome, for example to allow the passage of RNA polymerase II, and may reflect the more open, active nature of the *S. cerevisiae* genome when compared to higher eukaryotes.

Histone proteins have flexible N- and C-terminal tails that do not contribute to the stability of the nucleosome core particle or the rotational positioning of the DNA on the nucleosome (Luger et al. 1997b) but instead may be involved in internucleosomal contacts in the assembly of nucleosomes into higher order chromatin fibres (Luger and Richmond 1998). These tails protrude from the nucleosome core through channels formed by aligned DNA minor grooves (Luger et al. 1997a) and are thus accessible for extensive post-translational modifications.

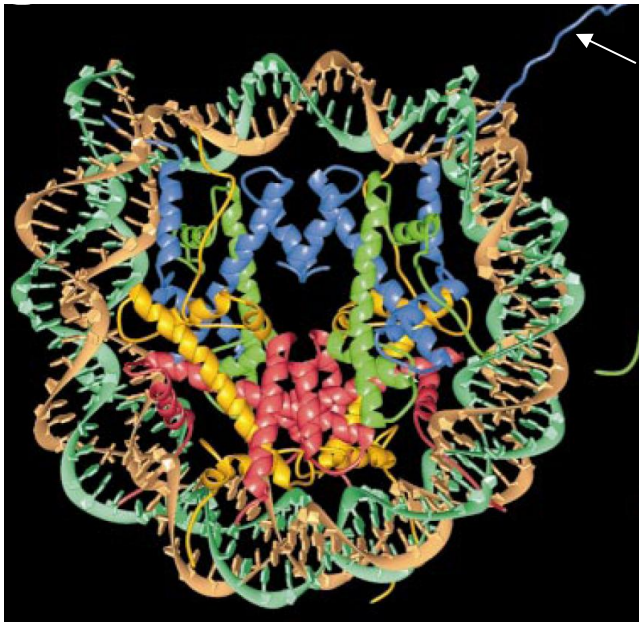


Figure 1. The X-ray crystal structure of the nucleosome to 2.8 Å resolution (Luger et al. 1997a)**.**

DNA (green and orange/turquoise) is wrapped around a core of histone proteins comprising two copies of H3 (blue), H4 (green), H2A (yellow) and H2B (red). One of the histone H3 tails in the nucleosome is indicated.

1.4. Histone tail structure

Whilst the histone tails are too disordered to appear in the X-ray crystal structure of the nucleosome (Luger et al. 1997a), molecular dynamics simulations predict that these tails are not simply irregular chains but do contain local structures (H. Patterton, pers. comm.). Circular dichroism experiments demonstrate that approximately 50 % of histone H3 and H4 N-terminal tail residues are in an α-helical conformation when bound to DNA in the nucleosome (Banères et al. 1997). Furthermore, predicted changes in backbone dihedral angles by modification of the histone tail may cause the adoption of tail conformations that are no longer permissive for the binding of histone modifying enzymes (Sanli et al. 2011). Thus it should not be assumed that the histone tails are completely unstructured but instead the effects of local structures should be considered.

1.5. The histone code evolves into a histone-based language

As mentioned, the tails of the histone proteins protrude from the core of the nucleosome and are accessible for extensive post-translational modifications. These include acetylation, methylation, phosphorylation, ubiquitination and sumoylation, modifications that have the potential to directly influence histone-DNA interactions or indirectly affect chromatin structure via interacting effector proteins. When the plethora of possible histone modifications first became apparent, the histone code hypothesis was put forward (Strahl and Allis 2000; Turner 2000; Jenuwein and Allis 2001). This proposed that different modifications had distinct functions and that, in combination, they could be read as part of a histone code to contribute to the formation of an active or silent chromatin environment. However, this view now seems a little too simplistic with the discovery that the same modifications can have different functional outputs that are dependent on the genomic context. Instead, it has been suggested that histone modifications should be treated more as the words in a complex chromatin language (Berger 2007; Gardner et al. 2011).

1.6. Readers of the histone language

A number of domains have been described that specifically recognise the different varieties of histone modifications. These include bromodomains, recognising acetylated lysines and Royal family proteins (chromodomains, Tudor domains) and plant homeodomain (PHD) fingers, both recognising methylated lysine residues (reviewed in Taverna et al. 2007). These domains have binding pockets that can recognise the modification and, in the case of methylation, distinguish between the states, in addition to surface channels or residues that enable recognition of the surrounding protein sequence. The proteins containing these domains are often found in complexes with potential chromatin-modifying activity.

1.7. Histone post-translational modifications

The proposed functions, deposition and removal of the histone modifications are discussed below and depicted in Figure 2. This discussion focuses predominantly on the modifications to histone H3 found in *S. cerevisiae* connected to transcriptional regulation, although, when relevant, examples from higher eukaryotes are mentioned.

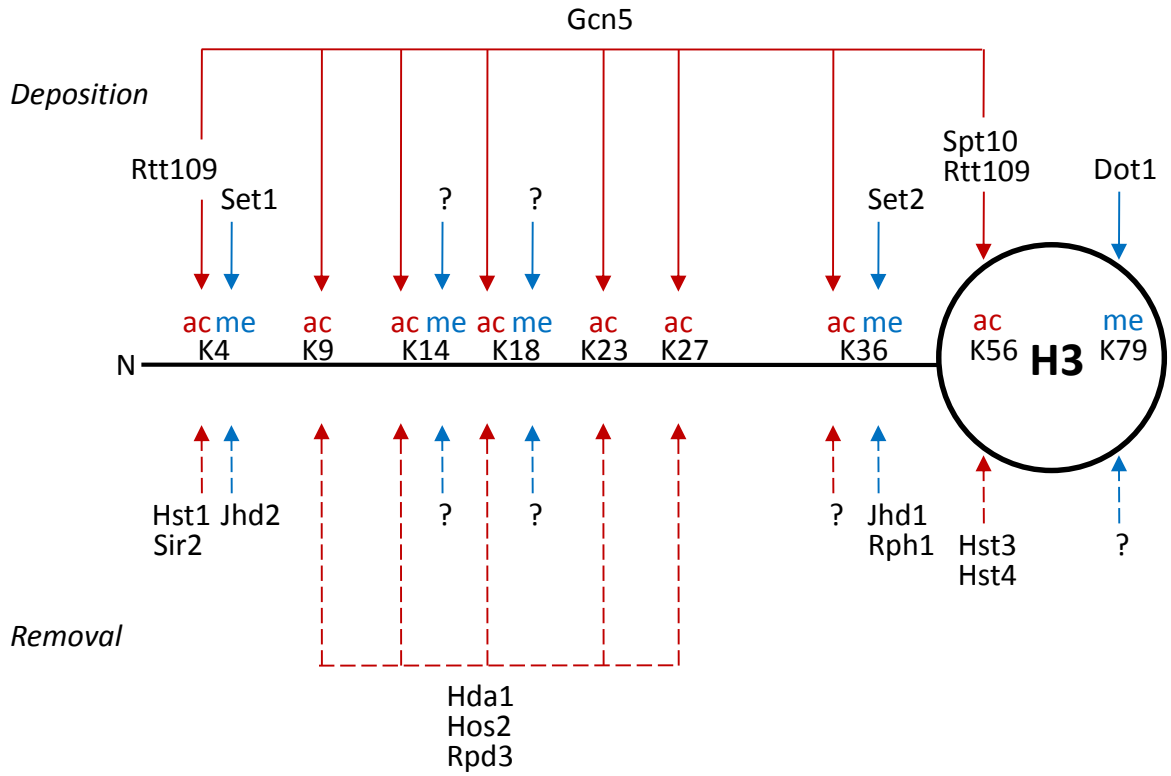


Figure 2. Deposition and removal of histone H3 modifications.

The enzymes required for deposition of acetylation (solid red lines) and methylation (solid blue lines) are shown at the top of the diagram. The enzymes required for removal of acetylation (dotted red lines) and methylation (dotted blue lines) are shown at the bottom of the diagram. Substrate preferences for each enzyme have not been shown for clarity.

1.8. Histone acetylation

Histone acetylation is found on many different lysine residues within all four histones in *S. cerevisiae* (although only acetylation of histones H3 and H4 is discussed below). Histone acetyltransferases (HATs) catalyse the formation of acetyl-lysine whereas histone deacetylases (HDACs) are responsible for the removal of the acetyl group.

1.8.1. Histone acetyltransferases

HATs transfer the acetyl group from acetyl-coA to the ε-amino group of the substrate lysine residue. There are a number of HATs in *S. cerevisiae* including Gcn5, Elp3, Sas3 and Esa1 that have numerous and overlapping target residues.

The discovery of Gcn5 HAT activity linked the previously described transcriptional co-activator function of this protein with histone acetylation (Brownell et al. 1996), although It should be noted that Gcn5 also acetylates a number of non-histone proteins, such as Rsc4 (Charles et al. 2011). Gcn5 acetylates free histones H3 and H4 primarily on H3K14, H4K8 and H4K16 *in vitro* (Kuo et al. 1996) but requires associated proteins for acetylation of nucleosomes *in vivo* (Grant et al. 1997; Ruiz-García et al. 1997). Gcn5 has been found in three complexes with differently directed HAT activities: ADA, Spt-Ada-Gcn5 acetyltransferase (SAGA) and SAGA-like (SLIK). ADA, the smallest Gcn5-containing complex at 0.8 MDa, consists of Ada2, Ada3 and Ahc1 in addition to Gcn5 (Grant et al. 1997; Eberharter et al. 1999). Within ADA, Gcn5 acetylates primarily H3K14 followed by H3K18 (Grant et al. 1999). In contrast, within SAGA, Gcn5 still preferentially targets H3K14 and H3K18 but can also acetylate H3K9 and H3K23 (Grant et al. 1999). SAGA is much larger than ADA (1.8 MDa) and consists of 20 subunits that can be split into five modules: the architecture module, required for SAGA integrity; the TBP-binding module; the HAT module containing Gcn5, together with Ada2 and Ada3 to moderate substrate specificity; the ubiquitin protease module; and the activator-binding/recruitment module (reviewed in Koutelou et al. 2010). SAGA controls the transcription of ~10% genes in *S. cerevisiae* (Lee, TI et al. 2000) and evidence points towards the function of SAGA as a transcriptional co-activator (Utley et al. 1998; Vignali et al. 2000; Bhaumik and Green 2001; Larschan and Winston 2001). The third Gcn5-containing complex, SLIK, is related to SAGA but has a truncated version of Spt7 and the retrograde response protein Rtg2 in place of Spt8 (Pray-Grant et al. 2002; Sterner et al. 2002; Wu and Winston 2002). SAGA and SLIK have partially overlapping roles in transcriptional regulation

but, in addition, SLIK functions in the retrograde response pathway (Pray-Grant et al. 2002) (see section 1.8.4).

As mentioned, approximately 10% of genes in *S. cerevisiae* are regulated by SAGA; the remainder are regulated by the general transcription factor, TFIID (Huisinga and Pugh 2004). Both of these complexes possess TBP-binding and histone acetyltransferase activities (Brownell et al. 1996; Mizzen et al. 1996; Grant et al. 1997). Interestingly, SAGA-regulated genes are mainly stress-induced and their expression is highly dynamic due to the recruitment of a large number of different chromatin regulators (Venters et al. 2011). In contrast, TFIID-regulated genes, comprising the majority of housekeeping genes, are less regulated and have a more constant expression pattern. Thus, these two mechanisms of transcriptional regulation balance the need for an inducible stress response and the steady output of housekeeping genes.

Elp3 is a subunit of the RNA polymerase II elongator complex and can acetylate all four histones *in vitro* (Wittschieben et al. 1999). However, in the context of intact elongator, Elp3 acetylates H3 and H4 predominantly on H3K14 and H4K8 (Winkler et al. 2002). Whilst deletion of *ELP3* only has mild effects on histone acetylation, deletion of both *ELP3* and *GCN5* results in severe H3 hypoacetylation and a decrease in expression of several genes, indicating that these HATs function synergistically (Kristjuhan et al. 2002).

Sas3 was identified as a component of the NuA3 HAT complex and is required both for HAT activity and complex integrity (John et al. 2000). NuA3 primarily acetylates H3K14 but shows some activity towards H3K23 (Howe et al. 2001; Martin et al. 2006a). Deletion of *SAS3* in combination with *GCN5* is synthetic lethal, demonstrating the essential overlapping functions of these two proteins (Howe et al. 2001). Another component of the NuA3 complex, Yng1, is required for the

chromatin recruitment and activity of NuA3 *in vivo* (Howe et al. 2002; Martin et al. 2006b; Taverna et al. 2006).

Esa1 is a subunit of the NuA4 HAT complex that acetylates the four conserved lysines on histone H4 (Allard et al. 1999) but can also acetylate H3K14 *in vitro* (Smith et al. 1998; Clarke et al. 1999). Interestingly, Esa1 is the only essential HAT in *S. cerevisiae* (Smith et al. 1998) perhaps due to its role in cell cycle progression (Clarke et al. 1999).

1.8.2. Histone deacetylases

There are several HDACs in *S. cerevisiae* including Rpd3, Hda1, Hos2, Hst1 and Sir2. These are often found in multi-subunit complexes and have overlapping, relatively low specificity for target lysines.

Deletion of *RPD3* results in increased acetylation of histones H3 and H4 *in vivo* (Rundlett et al. 1996), implicating Rpd3 in histone deacetylation. Rpd3 is a subunit of two complexes: a complex of 0.6 MDa, termed Rpd3S (Rundlett et al. 1996) and a larger complex of 1.2 MDa, termed Rpd3L (Kasten et al. 1997). Rpd3L contains at least 12 subunits including the transcriptional co-repressor Sin3 (Kasten et al. 1997) and gene-specific regulators such as Ash1 and Ume6 (Carrozza et al. 2005a). Hda1 has sequence similarity to Rpd3 but is found in a different complex, HDA (Rundlett et al. 1996) and targets an overlapping but distinct subset of genes (Robyr et al. 2002). As with Rpd3, deletion of *HDA1* causes increased histone H3 and H4 acetylation (Rundlett et al. 1996).

Hos2 was identified as an HDAC due to its sequence similarity with Hda1 and Rpd3 (Rundlett et al. 1996) and is found in the Set3 complex together with another HDAC, Hst1 (Pijnappel et al. 2001). The Set3 complex has HDAC activity *in vitro* and is involved in the repression of sporulation genes during meiosis (Pijnappel et al. 2001). Contrary to this, Set3 and the H3/H4-specific *in vivo*

deacetylase activity of Hos2 have been shown to be required for proper *GAL1* activation and there is a genome-wide correlation between Hos2 enrichment and transcript levels (Wang et al. 2002). This indicates that the effects of Hos2-mediated deacetylation are context-dependent.

Sir2 has long been characterised as a protein required for heterochromatin formation at the mating type loci, telomeres and rDNA locus (reviewed in Millar and Grunstein 2006). The discovery of Sir2 NAD-dependent HDAC activity linked these effects to a requirement for histone deacetylation, especially of H4K16 (Imai et al. 2000). Recently, Sir2 was also shown to deacetylate H3K4 in heterochromatin (Guillemette et al. 2011). Interestingly, deletion of *SIR2* results in a shortened replicative lifespan, perhaps relating nutritional status (in the availability of NAD) to ageing (Imai et al. 2000).

1.8.3. The effects of histone acetylation

Acetylation neutralises the positive charge of the lysine residue and may therefore have a direct effect on chromatin compaction by weakening the histone tail-DNA and internucleosomal interactions. Indeed, histone tail removal or acetylation, albeit to a lesser extent, both result in increased accessibility of nucleosomal DNA (Anderson et al. 2001) and greater instability of the nucleosome (Widlund et al. 2000). This increased exposure of nucleosomal DNA enables binding by DNA-processing factors such as the transcription factors Gal4 (Vettese-Dadey et al. 1996) and TFIIA (Lee et al. 1993) or nucleosome remodellers to previously occluded sites. In agreement with this model, acetylation enhances transcription of nucleosomal arrays by RNA polymerase III *in vitro* (Tse et al. 1998). Alternatively, acetylation can recruit effector proteins via acetyl-binding bromodomains. For example, Rsc4, a component of the RSC chromatin remodelling complex, is recruited to Gcn5-deposited H3K14ac via its tandem bromodomain (Kasten et al. 2004).

Histone acetylation generally has a positive effect on transcription: acetylation of H3K4 (Guillemette et al. 2011), H3K9, H3K14, H3K18, H4K5 and H4K12 maps to gene promoters (Liu et al. 2005; Pokholok et al. 2005) and there is a positive correlation between levels of H3 acetylation and transcriptional activity (Kurdistani et al. 2004; Pokholok et al. 2005; Guillemette et al. 2011) (Figure 3A). Exceptions to this are H4K8ac and H4K16ac, found in the 3' coding regions of genes (Liu et al. 2005). The levels of these modifications negatively correlate with transcription, perhaps explaining the need for Hos2-mediated deacetylation for active transcription (Wang et al. 2002).

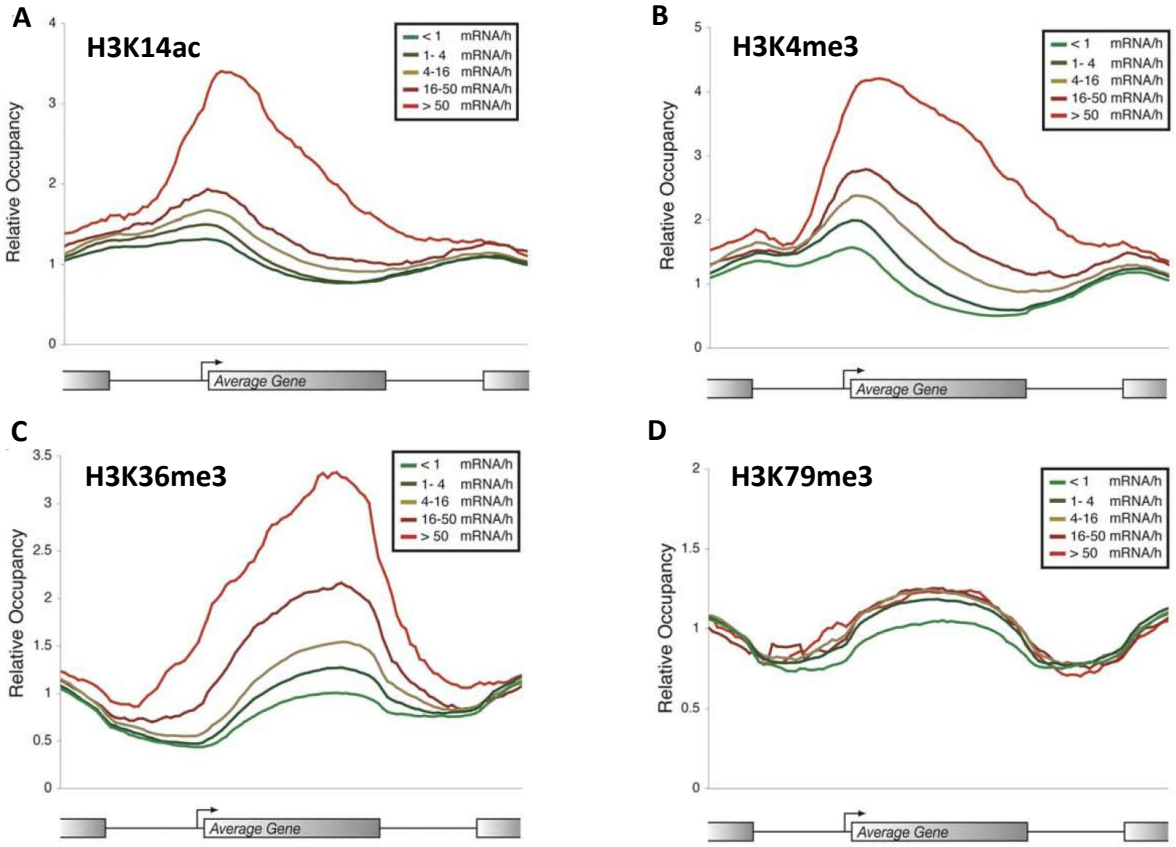


Figure 3. Genic distributions of histone acetylation and methylation.

Distribution profiles of **(A)** H3K14ac **(B)** H3K4me3 **(C)** H3K36me3 and **(D)** H3K79me3 relative to H3 across averaged genes on chromosome XII grouped according to rate of transcript production. Plots are adapted from Pokholok et al. (2005).

Histone acetylation can also have transcription-independent effects, such as the requirement of H3K56ac for replication-coupled nucleosome assembly (Li et al. 2008; Burgess et al. 2010) and H4K16ac to prevent the spreading of heterochromatin domains (reviewed in Millar and Grunstein 2006).

1.8.4. Signalling to H3K14

Acetylation of H3K14 is a downstream event for a number of signalling pathways, especially those responding to the cellular environment. The first suggestion that these pathways to H3K14 might exist was the discovery that the retrograde response protein, Rtg2, is a component of SLIK, the SAGA-like complex that predominantly acetylates H3K14 in addition to other H3 lysines (Pray-Grant et al. 2002). The retrograde response is an interorganelle communication that signals the functional status of the mitochondria to alter the expression of a subset of nuclear genes (Figure 4). Rtg2 is one of three components of the retrograde response, the others being Rtg1 and Rtg3 (Liao and Butow 1993). These latter components are helix-loop-helix leucine zipper transcription factors that heterodimerise to bind target gene promoters (Jia et al. 1997). Rtg2 transduces mitochondrial quality signals and promotes the nuclear import of the Rtg1-Rtg3 heterodimer by binding to and inactivating the negative regulator Mks1 and coordinating the partial dephosphorylation of Rtg3 (Komeili et al. 2000; Sekito et al. 2000; Dilova et al. 2002; Liu et al. 2003). The finding that Rtg2 is present in SLIK raised the question of whether SLIK is also involved in retrograde signalling. Further investigation showed that SLIK (but not SAGA) did indeed regulate the expression of the retrograde response target gene *CIT2* and that Rtg2 association with the *CIT2* promoter correlated with gene activity (Pray-Grant et al. 2002). This regulation potentially occurs through the acetylation of H3K14.

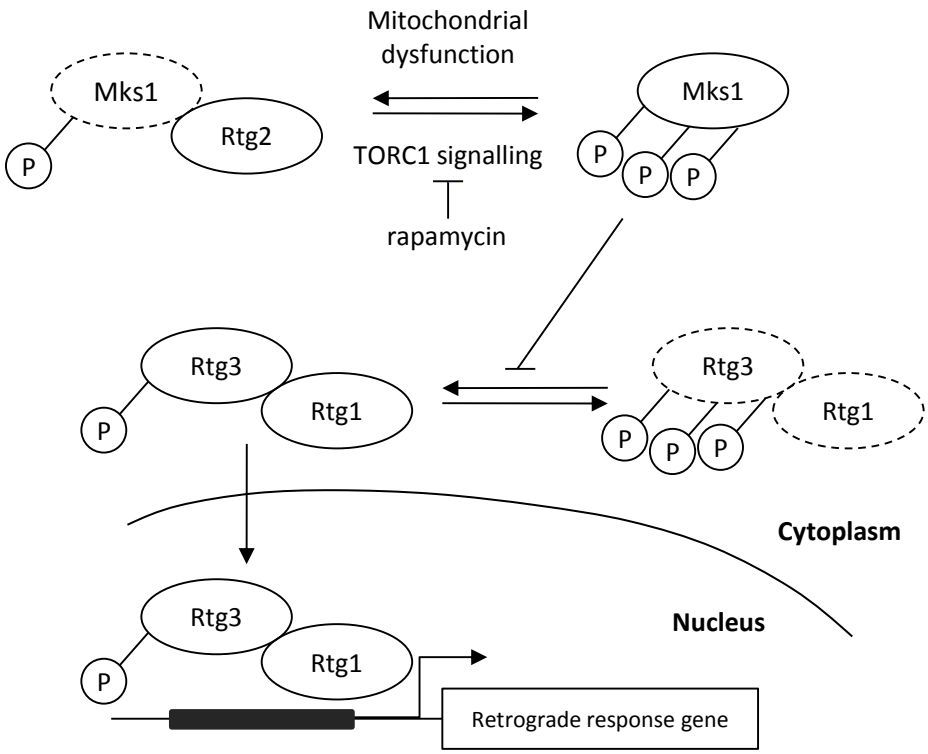


Figure 4. The retrograde response pathway in yeast.

The components of the retrograde response pathway. P represents phosphorylation, solid and dashed outlines indicate active and inactive proteins respectively. See text for explanation.

Other studies have shown that additional signalling pathways feed into the retrograde response pathway. This was demonstrated by the induction of Rtg1-Rtg3 nuclear localisation and target gene transcription upon treatment of yeast with rapamycin (Komeili et al. 2000). Rapamycin binds and inhibits target of rapamycin (Tor) proteins 1 and 2, components of the TORC1 complex. This complex is involved in nutrient signalling and metabolic homeostasis (comprehensively reveiwed by Loewith and Hall 2011). Although it is still unclear exactly how TORC1 affects retrograde signalling, a direct interaction between TORC1 and Mks1 has been observed (Breitkreutz et al. 2010). Potentially this interaction is involved in the inhibition of the Mks1-Rtg2 association, therefore preventing the inactivation of Mks1 by Rtg2 and subsequent retrograde signalling (Liu et al. 2003). Alternatively, TORC1 signalling may alter metabolite levels, thus indirectly influencing retrograde signalling.

TORC1 also interacts with the effector phosphatase, Tap42 (reviewed by Loewith and Hall 2011). Together, TORC1/Tap42 inhibit Snf1, the kinase contributing both to H3S10ph *in vivo* (Lo et al. 2001) and Gcn5 phosphorylation *in vitro* (Liu, Y et al. 2005). This phosphorylation of Gcn5 may influence its ability to acetylate H3K14: *in vitro,* phosphorylated Gcn5 exhibits slightly lower activity towards H3 and H4 than the unphosphorylated form although it is not known how this phosphorylated form of Gcn5, if in existence *in vivo*, would behave within its native complexes (Liu, Y et al. 2005). These examples illustrate the potential of H3K14 to be regulated by intracellular signalling pathways, especially those involved in the sensing of nutritional status.

1.9. Histone methylation

Histones can be methylated on lysine or arginine residues. Lysine residues are methylated on the ε-nitrogen and can be mono- (Kme1), di- (Kme2) or trimethylated (Kme3). In contrast, arginine can only be mono- (Rme1) or dimethylated, but the dimethylated form can exist in two conformations: symmetric (Rme2s) or asymmetric (Rme2a) methylation. Separate roles have been attributed to the different methylation states of a single residue, as discussed below.

1.10. Histone lysine methylation

The enzymes responsible for the addition of the methyl group(s) to histone lysine residues are called lysine methyltransferases (KMTs) and contain a catalytic SET domain, using S-adenosyl-L-methionine as the methyl donor. (An exception to this is Dot1, responsible for histone H3 lysine 79 (H3K79) methylation, which is a non-SET KMT (Feng et al. 2002).) The SET domain is a 130-140 residue motif found in several chromatin-associated proteins (Stassen et al. 1995; Jenuwein et al. 1998). There are twelve SET-domain containing proteins in *S. cerevisiae* (Wlodarski et al. 2011). The characterised yeast KMTs had been shown to be highly specific towards a particular residue for example, Set1 exclusively methylates histone H3 lysine 4 (H3K4) (Briggs et al. 2001; Roguev et al. 2001; Nagy et al. 2002; Krogan et al. 2002) whereas Set2 targets histone H3 lysine 36 (H3K36)

(Strahl et al. 2002) and it was thought that the same enzyme was capable of catalysing all three methylation states on a particular residue. However, Set5 has recently been identified as the methyltransferase for three residues on histone H4 (H4K5, H4K8 and H4K12) (Green et al. 2012) and only monomethylation has been detected on these residues.

1.10.1. H3K4 methylation

1.10.1.1 H3K4 methyltransferase

As mentioned, Set1 methylates histone H3 on lysine 4 (Briggs et al. 2001; Roguev et al. 2001; Nagy et al. 2002; Krogan et al. 2002; Santos-Rosa et al. 2002), although it should be noted that Set1 can also methylate non-histone proteins such as the kinetochore protein, Dam1 (Zhang, K et al. 2005). Set1 is classified as an exo-methylase since adding a Flag tag to the H3 N-terminus ablates H3K4 methylation (Takahashi et al. 2011). Whereas most yeast SET domain-containing KMTs can act independently, Set1 requires a complex of associated proteins for full activity (Roguev et al. 2001). This complex is termed COMPASS (complex of proteins associated with Set1) (Miller et al. 2001) (Figure 5). In addition to Set1, COMPASS consists of 7 subunits: Swd1, Swd3, Swd2, Sdc1, Bre2, Spp1 and Shg1 (Roguev et al. 2001), which differentially affect the ability of the complex to mono-, di- and trimethylate H3K4. Swd1 and Swd3 interact strongly to form a heterodimer (Murton et al. 2010) and are required for COMPASS stability (Dehé et al. 2006). Therefore deletion of either *SWD1* or *SWD3* results in complete loss of H3K4 methylation (Morillon et al. 2005; Schneider et al. 2005). Swd2 is an essential protein, maybe due to the fact that it is also found in the cleavage and polyadenylation factor (CPF) complex, a complex required for proper RNA polymerase II transcription termination (Cheng et al. 2004; Dichtl et al. 2004). A role has been described for Swd2 in the *trans*-tail crosstalk between H2BK123 monoubiquitination and H3K4me3 (discussed in section 1.16.2). Deletion of *SDC1* or *BRE2* causes loss of detectable H3K4me3 and a severe reduction in the levels of H3K4me2 (Morillon et al. 2005; Schneider et al. 2005; Dehé et al. 2006). Sdc1 and Bre2 form a heterodimer via the Dpy-30 domain of Sdc1 and

the SDI domain of Bre2 and deletion of either protein results in the inability of the other protein to interact with Set1 (South et al. 2010), perhaps explaining their similar effects on H3K4 methylation. 3D cryo-electron microscopy and immunolabelling-2D electron microscopy studies depict a structure for COMPASS, with the Swd1-Swd3 and Bre2-Sdc1 heterodimers on opposite sides of Set1 (Takahashi et al. 2011).

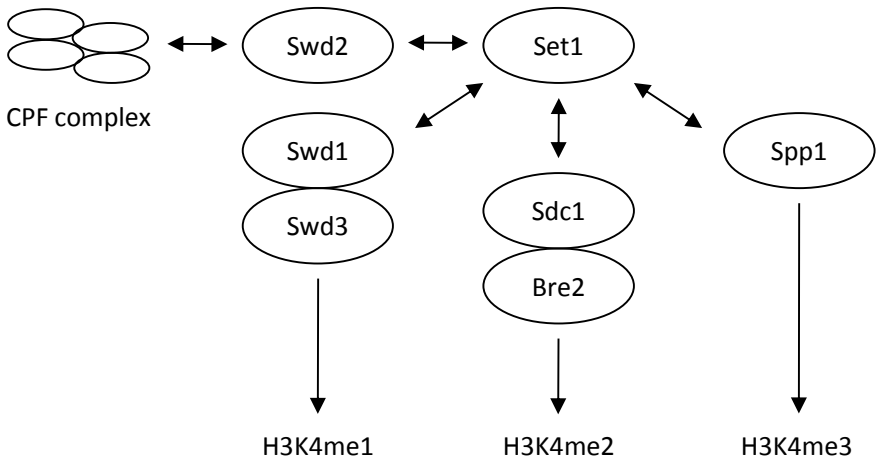


Figure 5. The COMPASS complex in yeast.

The subunits associated with the Set1 methyltransferase and their effect on the H3K4 methylation state deposited by the complex. As indicated, the essential protein Swd2 is also present in the cleavage and polyadenylation factor (CPF) complex.

Deletion of *SPP1* mainly affects H3K4me3, with a small reduction in H3K4me2 (Morillon et al. 2005; Schneider et al. 2005). Spp1 directly interacts with the Set1 N-SET domain (Takahashi et al. 2011) and this has been suggested to influence the catalytic activity of Set1. Takahashi et al. (2009) proposed the Phe/Tyr switch model for Set1 based on the observations that in higher eukaryotes, KMTs catalysing monomethylation have a tyrosine in the catalytic site whereas KMTs depositing di- and trimethylation have a phenylalanine/similarly hydrophobic residue in the corresponding position (Xiao, B et al. 2003; Zhang et al. 2003). In yeast Set1, there is a tyrosine in this position, predicting that the isolated enzyme would only be able to monomethylate H3K4 and that the COMPASS subunits Bre2-Sdc1 and Spp1 are required to influence the Set1 catalytic activity to allow H3K4me2/3. Evidence in support of this model is that mutation of the conserved

tyrosine to phenylalanine results in higher H3K4me3 and suppresses the loss of Spp1 on H3K4me3 (Takahashi et al. 2009). However, *in vitro* studies by the same group demonstrate that Spp1 is not required for trimethylation of H3K4 on free histone H3, arguing against a strict requirement of Spp1 for Set1-induced H3K4me3 (Takahashi et al. 2011). Spp1 has a PHD finger that recognises H3K4me2 and H3K4me3 (Shi et al. 2007), perhaps suggesting that Spp1, through its PHD domain, may bind H3K4me2 and stabilise COMPASS on chromatin to bring about H3K4me3 (Kirmizis et al. 2007). Another possibility is that, by binding H3K4me2/3 on one histone tail, Spp1 can recruit COMPASS to methylate the other histone H3 tail in the same nucleosome. Additionally, it was speculated that the binding of Spp1 to H3K4me2/3 may protect these modifications from demethylation (Kirmizis et al. 2007).

The machinery for H3K4 methylation is well conserved from yeast to humans. Humans have six complexes that catalyse H3K4 methylation: SET1A/B and MLL1-4. These complexes are not redundant, as demonstrated by the embryonic lethality caused by deletion of individual *MLL* genes (Yu et al. 1995; Glaser et al. 2006). The two human SET1 complexes are most similar to yeast COMPASS and consist of hSet1A or B, RBBP5 (Swd1 homologue), WDR5 (Swd3 homologue), Wdr82 (Swd2 homologue), hDPY30 (Sdc1 homologue), Ash2 (Bre2 homologue) and CXXC1 (Spp1 homologue) (reviewed by Shilatifard 2008). The MLL complexes are COMPASS-like and consist of similar core components, although the structural and functional organisation of these subunits differs from the yeast COMPASS (Dou et al. 2006; Steward et al. 2006). Additionally, the MLL complexes contain specific subunits, for example the MLL1/2 complexes contain the tumour suppressor gene *menin* (Hughes et al. 2004; Yokoyama et al. 2004).

1.10.1.2 H3K4 demethylase

A conditional mutant of Set1 demonstrated the rapid turnover of H3K4me2/3 *in vivo* (Seward et al. 2007). H3K4 methylation is removed by the demethylase Jhd2 (Ingvarsdottir et al. 2007; Liang

et al. 2007; Seward et al. 2007; Tu, S et al. 2007), which is capable of removing all three H3K4 methylation states *in vivo* (Huang et al. 2010). Interestingly, Jhd2 may further repress H3K4 methylation by reducing the levels of Set1 on chromatin, an effect that requires Jhd2 demethylase activity (Ingvarsdottir et al. 2007). Jhd2 acts as a monomeric protein and does not have any stably associated protein partners (Liang et al. 2007). Protein levels of Jhd2 are tightly controlled *in vivo* by Not4-induced polyubiquitination followed by proteasomal degradation (Laribee et al. 2007; Mersman et al. 2009). Studies have shown that the Jhd2 PHD domain is not required for demethylase activity *in vitro* (Ingvarsdottir et al. 2007) but is important for chromatin association and demethylase activity *in vivo* (Huang et al. 2010). An attractive hypothesis is that the Jhd2 PHD domain interacts with H3K4 methylation or another methylated residue on the H3 tail before H3K4 demethylation. However, neither deletion of *SPP1* (thus decreasing H3K4me3) nor deletion of the first 28 residues of H3 has an effect on Jhd2 chromatin association (Huang et al. 2010) and the Jhd2 PHD domain does not interact with mono-, di- or trimethylated K4 peptides *in vitro* (Shi et al. 2007).

1.10.1.3 The effects of H3K4 methylation

H3K4 methylation has been mapped to the promoters and coding regions of actively transcribed genes (Santos-Rosa et al. 2002; Bernstein et al. 2002) (Figure 6). At these genes there is a peak of H3K4me3 at the promoter and 5' coding region that correlates positively with transcript levels (Figure 3B). Dimethylated H3K4 peaks further into the gene than H3K4me3 and monomethylated H3K4 is highest towards the 3' end, especially on long genes (Liu et al. 2005; Pokholok et al. 2005) (Figure 6). These different distributions suggest that the three H3K4 methylation states may have distinct functions. Further evidence for the link between H3K4 methylation and gene activity is the fact that COMPASS is recruited to chromatin via the Paf1 complex and the Ser5-phosphorylated RNA polymerase II C-terminal domain (Krogan et al. 2003a; Ng et al. 2003a). To support a role in gene activation, H3K4me3 has been shown to recruit a number of gene-

regulatory factors, which have characterised functions in gene activation. These include Isw1, an ATP-dependent chromatin remodeller (Santos-Rosa et al. 2003), Yng1, a subunit of the NuA3 HAT complex leading to acetylation of H3K14 (Martin et al. 2006b; Taverna et al. 2006) and Sgf29 within the SAGA complex, promoting H3K9/K14/K18 acetylation in yeast and humans (Vermeulen et al. 2010; Bian et al. 2011). Additionally, in humans, H3K4me3 recruits the chromatin remodellers Chd1 (in SAGA/SLIK) (Sims et al. 2005) and BPTF (in NURF) (Wysocka et al. 2006), and the TFIID subunit TAF3 (Vermeulen et al. 2007). Further evidence for a positive function of H3K4 methylation in gene transcription is that, in vertebrates, H3K4me1 localises to gene enhancers (Heintzman et al. 2007; Wang et al. 2008).

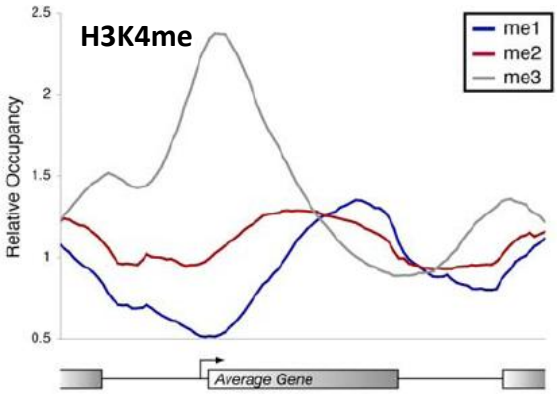


Figure 6. Average gene distributions of the three states of H3K4 methylation.

Distribution profiles of monomethylated (me1), dimethylated (me2) and trimethylated (me3) H3K4 relative to H3 across averaged genes on chromosome XII. Plot adapted from Pokholok et al. (2005).

In contrast to the above evidence, there are a number of reports suggesting a repressive role for H3K4 methylation. Set1 was initially identified as a protein required for rDNA and telomeric silencing (Briggs et al. 2001; Bryk et al. 2002; Krogan et al. 2002). The requirement for Set1 was refined by Fingerman et al. (2005), who used a mutant of Set1 that did not support H3K4me3 to show that H3K4me3, but not H3K4me1/2, was required for proper silencing at the telomeric, rDNA and *HML/HMR* loci. Furthermore, the levels of H3K4me3 increase on the ribosomal protein genes during their repression in response to diamide stress (Weiner et al. 2012). Deletion of *SET1* results in increased expression of *PHO5, PHO84* and *GAL10* (Carvin and Kladde 2004). Further

work at the *PHO5* locus demonstrated that H3K4me3 (and to some extent, H3K4me2) interacts with the PHD domains of the Rpd3L HDAC complex subunits Pho23 and Cti6, thus recruiting the HDAC complex to the *PHO5* promoter. The subsequent histone H3 deacetylation results in a higher nucleosome occupancy, which is repressive to transcription initiation (Wang et al. 2011). In human cells, the ING2 PHD domain binds H3K4me3 upon DNA damage, causing recruitment of the Sin3-HDAC complex to silence genes involved in proliferation (Shi et al. 2006).

H3K4me2 has also been proposed to have repressive functions (Pinskaya and Morillon 2009). H3K4me2 recruits the Set3 histone deacetylase complex and functions to prevent the spreading of histone acetylation from the promoter region into the body of the transcribed gene (Kim and Buratowski 2009). Several studies dissecting the functions of H3K4 methylation have made use of the inducible *GAL* locus. Reb1-dependent long non-coding RNAs have been proposed to trigger H3K4me2/3 deposition at this locus (Houseley et al. 2008). This H3K4me2/3 delays the recruitment of RNA polymerase II and TBP to the *GAL1* promoter upon gene induction via the recruitment of Rpd3S (Pinskaya et al. 2009). After gene inactivation, H3K4me2/3 persists at the *GAL* locus for around 5 hours and was suggested to act as a molecular mark of recent transcription (Ng et al. 2003a). Building on this, Zhou and Zhou (2011) showed that deletion of *SET1* (but not *SPP1* and therefore this effect is more due to loss of H3K4me1/2 than H3K4me3) results in a faster re-activation rate upon a second exposure to galactose. H3K4 methylation is also present at the promoters of inositol- and fatty acid-responsive genes after recent transcription, inhibiting their re-induction possibly through a pathway involving Isw1-mediated inhibition of transcription elongation (Zhou and Zhou 2011).

Recently, a new role for H3K4me3 has been described in the promotion of efficient Nrd1-dependent termination of RNA polymerase II transcription (Terzi et al. 2011). Nrd1, in combination with its complex members Nab3 and Sen1, binds to specific sites in the nascent RNA

(Carroll et al. 2004, 2007) and mediates the termination of small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), some short mRNAs and CUTs (Steinmetz et al. 2001; Arigo et al. 2006b; Kim et al. 2006; Thiebaut et al. 2006). Nrd1-dependent termination directs these transcripts to the exosome where they are either trimmed at their 3' ends, in the case of sn/snoRNAs, or degraded (Vasiljeva and Buratowski 2006). One function of Nrd1-mediated termination is to promote the attenuation of certain protein-coding genes, thus repressing their expression. Genes regulated in this manner include *NRD1* itself, allowing the autoregulation of Nrd1 protein levels (Steinmetz et al. 2001; Arigo et al. 2006a), and a number of stress-regulated genes such as *FKS2* (Kim and Levin 2011). Terzi et al. (2011) showed that, as expected, mutation of *NRD1* results in slow growth and defective transcription termination. However these phenotypes were exacerbated by both simultaneous deletion of *SET1* and mutations in *SET1* that specifically disrupt H3K4me3, indicating that H3K4 methylation, in particular H3K4me3, contributes to Nrd1-dependent transcription termination. The authors observed reduced Nrd1 recruitment in the absence of H3K4 methylation. Furthermore, H3K4me3, through its ability to recruit both the HDAC Rpd3L via Pho23 and Cti6 (Wang et al. 2011) and the HAT complex NuA3 via Yng1 (Martin et al. 2006b; Taverna et al. 2006), is important for controlling the level of histone acetylation, which influences the kinetics of early elongation by RNA polymerase II and therefore the efficiency of Nrd1-mediated termination.

Another link between H3K4 methylation and ncRNAs is that H3K4me3 may act to antagonise the repressive activity of XUTs that are transcribed in an antisense direction through open reading frames and the retrotransposon *TY1* (Van Dijk et al. 2011). However, the same group previously published a function for H3K4me2 and/or H3K4me3 in ncRNA-mediated *silencing* of *TY1* (Berretta et al. 2008). These apparent disparities between the effects of H3K4 methylation on *TY1* transcription may be a result of the different states of H3K4 methylation removed by the mutations used. In the study by van Dijk et al., H3K4me3 was specifically repressed whereas the

study by Berretta et al. used a mutant strain with reductions in both H3K4me2 and H3K4me3. Therefore H3K4me2 may promote whereas H3K4me3 may antagonise ncRNA-mediated silencing of the *TY1* retrotransposon. This would agree with the repressive functions of H3K4me2 in chromatin as described above.

In spite of this collective evidence, it is difficult to infer any overarching causality in the relationship between H3K4 methylation and transcription. Indeed, genome-wide microarray studies indicate that neither deletion of *SET1* nor *SPP1* causes major changes in transcript levels: only 55 genes change in expression more than 1.7-fold upon deletion of *SET1* during steady-state exponential growth (Lenstra et al. 2011). It is possible that, instead of being actively involved in the regulation of transcription, H3K4 methylation may be deposited as a result of transcription, possibly to regulate transcript stability. This hypothesis remains consistent with the correlation between H3K4 methylation and gene activity. Alternatively, H3K4 methylation may be more important in regulating dynamic changes in gene transcription than steady-state transcription (Weiner et al. 2012). The function and regulation of H3K4 methylation in chromatin is therefore still a topic of intense research.

1.10.2. H3K36 methylation

H3K36 methylation is catalysed by Set2 (Strahl et al. 2002) and removed by the demethylases Jhd1 and Rph1 (Kim and Buratowski 2007; Tu, S et al. 2007). Set2 is recruited to the Ser2-phosphorylated RNA polymerase II CTD to deposit H3K36me2/3 in the body of active genes (Krogan et al. 2003b; Li et al. 2003; Schaft 2003; Xiao, T et al. 2003; Kizer et al. 2005) (Figure 3C) (although H3K36me2, in contrast to H3K36me3, does not require the interaction of Set2 with phosphorylated RNA polymerase II (Youdell et al. 2008) and does not correlate with transcriptional frequency (Batta et al. 2011)). The histone deacetylase complex Rpd3S recognises H3K36me2 via the chromodomain of Eaf3 and brings about histone deacetylation in the wake of

gene transcription to prevent the initiation of spurious intragenic transcripts (Carrozza et al. 2005b; Joshi and Struhl 2005; Keogh et al. 2005). H3K36 methylation also has an Rpd3-independent role in the prevention of ectopic silencing (Tompa and Madhani 2007). H3K36 methylation provides a good example of the nuances of the histone language: H3K36 methylation is localised in the bodies of active genes and yet has a 'repressive' function in the inhibition of cryptic transcription.

1.10.3. H3K79 methylation

Dot1 was identified in a screen for genes involved in telomeric silencing (Singer et al. 1998) and was subsequently found to be responsible for H3K79 methylation (Ng et al. 2002; van Leeuwen et al. 2002). The human homologue, hDOT1L, has a similar protein structure and catalytic activity towards H3K79 as the yeast enzyme (Feng et al. 2002). As mentioned, Dot1 is different from other KMTs in that it does not contain the catalytic SET domain (Feng et al. 2002). Instead the catalytic domain of the yeast and human Dot1 is more similar to that of the arginine methyltransferases (Min et al. 2003; Sawada et al. 2004). Unlike the other lysine methyltransferases, Dot1 is a non-processive enzyme, resulting in the inability to controllably generate specific H3K79 methylation states (Frederiks et al. 2008). This may explain the overlapping gene distributions of the three H3K79 methylation states (Shahbazian et al. 2005) and the functional redundancy of these states in chromatin (Frederiks et al. 2008). An additional difference is that H3K79 is found on an exposed surface of the nucleosome body (rather than the histone tails) that does not contact other histones or DNA (Luger et al. 1997a). It has been estimated that approximately 90 % of H3K79 is methylated *in vivo* (Van Leeuwen et al. 2002) and there is little correlation between H3K79me3 and gene activity (Pokholok et al. 2005) (Figure 3D). Instead, this methylation is proposed to restrict the silencing proteins Sir2 and Sir3 to telomeric regions. In the absence of H3K79 methylation (in H3 K79A point mutation or *dot1Δ* strains), the Sir proteins become delocalised over the genome, thus titrating out the silencing potential at the telomeres (Van Leeuwen et al.

2002; Ng et al. 2002). A demethylase acting against H3K79 has yet to be identified (Katan-Khaykovich and Struhl 2005; Liang et al. 2007; Tu, S et al. 2007) and may not exist. This would be supported by the fact that H3K79 methylation levels are proportional to histone residence time in chromatin (De Vos et al. 2011).

1.10.4. H3K18 methylation

H3K18 monomethylation (H3K18me1) was discovered in a mass spectrometry study on histones purified from yeast, mouse and humans (Garcia et al. 2007). The only subsequent mention of H3K18me1 in the literature was in a study measuring the dynamics of histone modifications and correlating these with the proposed function of the modification *i.e.* whether the modification was deemed 'activating' or 'silencing' (Zee et al. 2010). This study concluded that H3K18me1 had a half-maximal time slower than other monomethyl marks with activating roles and similar to H3 turnover, suggesting that H3K18me1 has a silencing function. They further hypothesised that there may be a binary switch at H3K18, with H3K18me1 acting to antagonise the activating H3K18ac in the regulation of gene expression. Crosstalk of this type has been described in yeast at H3K36, as discussed in section 1.16.1 (Morris et al. 2007).

1.11. Histone arginine methylation

Histone arginine methylation is catalysed by protein arginine methyltransferases (PRMTs). There are three classes of arginine methyltransferases that catalyse: Rme1 and Rme2a (type I), Rme1 and Rme2s (type II) or only Rme1 (type III) (reviewed in Di Lorenzo and Bedford 2011). All three of these classes utilise S-adenosyl-L-methionine to add methyl group(s) to the terminal guanidino nitrogens of the arginine side chain. Three PRMTs have been identified in *S. cerevisiae* (Rmt1, Rmt2, Hsl7), although these appear to have non-histone substrates (Gary et al. 1996; Henry and Silver 1996; Niewmierzycka and Clarke 1999; Lee, J-H et al. 2000; Chern et al. 2002).

1.11.1. H3R2 methylation

H3R2 methylation was discovered in *S. cerevisiae* in 2007 (Kirmizis et al. 2007). A ChIP-chip experiment mapped H3R2me2a to heterochromatic regions (silent mating type loci, telomeres and rDNA repeats), inactive euchromatic genes and the 3' ends of moderately transcribed genes. Substitution of H3R2 with alanine resulted in loss of heterochromatic silencing, suggesting a repressive function for a modification on this residue (Kirmizis et al. 2007). In contrast, H3R2me1 maps to the coding regions of 85 % of yeast genes and correlates with active transcription (Kirmizis et al. 2009).

The mammalian arginine methyltransferase PRMT6 is predominantly responsible for H3R2me2a (rather than H3R2me1 (Hyllus et al. 2007; Lakowski and Frankel 2008)) in human cell lines (Guccione et al. 2007; Iberg et al. 2008). The PRMT for H3R2 in *S. cerevisiae* has yet to be identified, despite screening 35 known and candidate methyltransferases in yeast (Kirmizis et al. 2009). Although JMJD6 was initially identified as a mammalian arginine demethylase (Chang et al. 2007), this finding was questioned by subsequent reports suggesting that JMJD6 was instead a lysine hydroxylase (Webby et al. 2009; Hong et al. 2010; Mantri et al. 2010). No other enzyme capable of removing arginine methylation has yet been identified in yeast or mammals (Di Lorenzo and Bedford 2011).

1.12. Histone serine/threonine phosphorylation

Phosphorylation of histone tails on serine and threonine is catalysed by a number of kinases and requires ATP as a phosphate donor. This adds a negative charge to the tail, which may directly influence tail-DNA interactions, causing decondensation of the chromatin fibre. However, the phosphorylation of H3S10, H3T11 and H3S28 has been shown to play a role in chromosome condensation during mitosis and meiosis (reviewed in Banerjee and Chakravarti 2011). Additionally, H3 phosphorylation is involved in the induction of the immediate-early genes *c-fos*

and *c-jun* in metazoa (Banerjee and Chakravarti 2011). H3S10ph may also function in transcriptional regulation in *S. cerevisiae*: Snf1-induced H3S10ph works in concert with Gcn5-catalysed H3K14ac to enhance the expression of *INO1* (Lo et al. 2001). Possible removal of histone phosphorylation by phosphatases has not been well characterised in *S. cerevisiae*.

1.13. Histone proline isomerisation

Because peptide bonds have partial double bond character, rotation around the peptide bond is restricted. This causes the peptide bond to adopt one of two conformations: *cis*, when the dihedral angle, ω, is 0° or *trans,* when ω is 180° (Ramachandran and Sasisekharan 1968). The *trans* peptide bond is energetically much more favourable than the *cis* bond due to the steric repulsion between neighbouring C_α atoms in the *cis* conformation, hence the *cis* peptide bond is a rarity in proteins (Ramachandran and Mitra 1976). This was illustrated by the analysis of 571 non-redundant X-ray-solved protein structures, in which only 0.03 % of peptide bonds are in the *cis* conformation (Weiss et al. 1998). An exception to this is the peptide bond formed between a proline and its neighbouring residue. The *cis* conformation of the peptidyl-prolyl imide bond, while still less energetically favourable than the *trans* conformation, occurs more frequently than for non-proline peptide bonds. Indeed, when studying the same 571 protein structures, the incidence of *cis* peptide bonds involving proline is 5.2 % (Weiss et al. 1998).

Conversion between the *cis* and *trans* conformations of the peptidyl-prolyl bond can occur, but because of the high activation energy (~ 90 kJ mol^{-1} for oligopeptides (Schmid 1995)), this process is intrinsically very slow. Enzymes exist that can lower the activation energy and so catalyse *cis-trans* prolyl isomerisation (Figure 7). There are 13 known peptidyl-prolyl *cis-trans* isomerases (PPIases) in *S. cerevisiae,* which fall into three distinct classes: eight cyclophilins (Cpr1-8), four FK506-binding proteins (FKBPs, Fpr1-4) and a parvulin (Ess1) (reviewed in Lu et al. 2007). Of note, Cpr1, Cpr6 and Cpr7 interact with the HDAC complexes Set3 and Rpd3L, although their exact roles

within these complexes are unknown (Duina et al. 1996; Arévalo-Rodríguez et al. 2000; Pijnappel et al. 2001), and Fpr3 and Fpr4 interact with the histone H3 and H4 tails (Nelson et al. 2006). The H3 tail contains three proline residues at positions 16, 30 and 38. Of these, only H3P38 has a known PPIase, Fpr4 (Nelson et al. 2006).

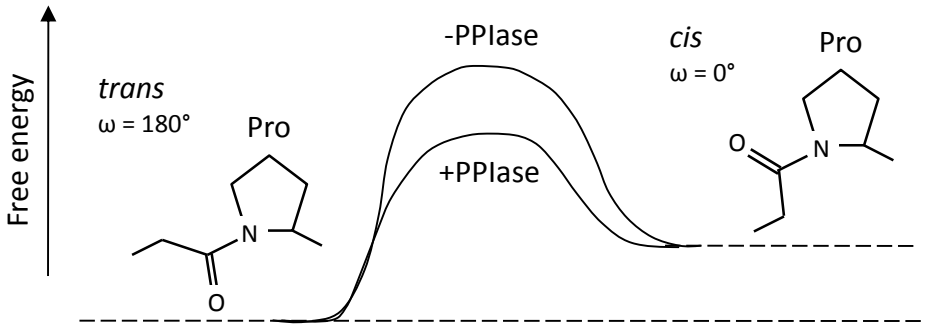


Figure 7. Proline isomerisation.

A schematic showing the free energy profile for uncatalysed and peptidyl-prolyl isomerase (PPIase)-catalysed *trans-cis* peptidyl-prolyl isomerisation.

Whilst histone proline isomerisation may not be classified as a true histone modification, it has the potential to dramatically alter the local structure of the histone tail and thus the recognition by effector proteins or enzymes depositing/removing adjacent modifications. Indeed H3P16 and H3P38 have been implicated in the recognition of the H3 tail by Gcn5 in *Tetrahymena* (Rojas et al. 1999) and crosstalk with H3K36 methylation (Nelson et al. 2006, discussed in section 1.16.3) respectively.

1.14. Histone H2B ubiquitination

Ubiquitin is a 76 amino acid protein that is added to target proteins either singly (monoubiquitination) or multiply to form ubiquitin chains (polyubiquitination). Polyubiquitination targets the substrate protein for proteasomal degradation whereas monoubiquitination is used to covalently mark the target protein. Proteins are ubiquitinated by a three-step reaction. Firstly, a ubiquitin molecule is activated by an E1 ubiquitin-activating enzyme in an ATP-dependent

reaction. The activated ubiquitin is then conjugated via a thiol-ester bond to a cysteine residue in an E2 ubiquitin-conjugating enzyme before transfer to the target lysine on the substrate protein by an E3 ubiquitin ligase.

1.14.1. Histone H2B ubiquitin conjugation

Histones H2A and H2B can be monoubiquitinated, although only H2B monoubiquitination (H2Bub) occurs in *S. cerevisiae*. H2B ubiquitination has been shown to occur similarly in humans (Kim, J et al. 2009). Ubiquitination is deposited on H2B lysine 123 by the E2 ubiquitin-conjugating enzyme Rad6 (Robzyk et al. 2000). Whilst several E3 ligases have been shown to cooperate with Rad6 *in vivo,* only Bre1 has been shown to influence H2B monoubiquitination (Hwang et al. 2003). Bre1 confers substrate specificity on Rad6 towards H2BK123 (Kim and Roeder 2009) and is required for the association of Rad6 with chromatin (Wood et al. 2003a; Kao et al. 2004).

H2Bub is found at actively transcribed genes but is absent from silent chromatin (Kao et al. 2004). Transcriptional activators, such as Gal4, recruit Bre1 (and subsequently Rad6) to promoters but this localisation of Rad6/Bre1 is not sufficient for H2B monoubiquitination. In addition, Rad6/Bre1 requires the Paf1 complex for catalytic activity but not recruitment to promoters (Ng et al. 2003b; Wood et al. 2003b). The Bur1/2 cyclin-dependent kinase complex phosphorylates Rad6, which controls the Rad6/Bre1-Paf1 complex interaction (Wood et al. 2005; Laribee et al. 2005). The Paf1 complex promotes the interaction between Rad6/Bre1 and the serine 5-phosphorlyated RNA polymerase II C-terminal domain (via Bre1 (Kim and Roeder 2009)) and subsequently allows the entry of Rad6 into the coding region of a gene (Xiao et al. 2005).

1.14.2. The effect of histone H2B ubiquitination

Yeast strains lacking H2Bub show genetic interactions with genes affecting transcription elongation such as *CTK1* and *SET2* (Xiao et al. 2005), supporting a role for H2Bub in the regulation

of transcription elongation. H2Bub aids the histone chaperone complex FACT in allowing transcription through nucleosomes (Pavri et al. 2006). Additionally, H2Bub cooperates with the FACT subunit Spt16 in promoting the reassembly of nucleosomes following transcription (Fleming et al. 2008). These studies have been corroborated by a micrococcal nuclease chromatin immunoprecipitation-sequencing (MNase ChIP-seq) experiment showing that nucleosome formation, mediated by H2Bub, influences the process of transcription negatively and positively at promoters and gene bodies respectively (Batta et al. 2011). H2Bub stabilises the nucleosome by an unknown mechanism (Chandrasekharan et al. 2009), which might partially account for the repressive effect of promoter-localised H2Bub on transcription. Contrary to the effect on nucleosomes, H2B monoubiquitination has been shown to lower the compaction of higher order chromatin structure, as measured by sedimentation velocity experiments, MNase sensitivity and FRET between neighbouring nucleosomes in an array (Fierz et al. 2011). This is specific for ubiquitination and not merely a function of its bulk since a similarly sized protein, Hub1, only had weak effects when conjugated to the site of ubiquitination.

H2B deubiquitination is also needed for gene activation and is catalysed by Ubp8 (Henry et al. 2003; Daniel et al. 2004) and Ubp10 (Emre et al. 2005; Gardner et al. 2005). Interestingly Ubp8 and Ubp10 appear to deubiquitinate different pools of H2B, correlating with the enrichment of H3K4me3 or H3K79me3 on nucleosomes respectively (Schulze et al. 2011). Ubp10 interacts with Sir2 and is enriched at silent loci (Emre et al. 2005; Gardner et al. 2005). In contrast, Ubp8 is a component of the SAGA complex and is dependent on the association with SAGA, mediated by Sgf11 (Shukla et al. 2006), for activity. This association of a ubiquitin protease with an activating transcription complex suggests that H2B ubiquitination is a dynamic modification and that its removal shortly after the initiation of transcription is important for gene activation (Henry et al. 2003; Daniel et al. 2004). Additionally, Ubp8 in a SAGA-related complex is required for optimal recruitment of Ctk1 to phosphorylate the RNA polymerase II CTD on serine 2 and promote Set2

recruitment to chromatin (Wyce et al. 2007), perhaps indicating that H2Bub acts as a checkpoint for RNA polymerase II progression during early transcription elongation.

1.15. Histone sumoylation

Small ubiquitin-related modifiers (SUMOs) are ~11 kDa proteins with similar structure to ubiquitin. They are added to proteins by a similar E1/E2/E3 pathway to ubiquitination (reviewed by Johnson 2004). Sumoylation has been detected at numerous sites on all four core histones in *S. cerevisiae* (Nathan et al. 2006). Whilst this modification is found throughout the genome, there is a slight peak at telomere-proximal regions and levels of sumoylation decline upon induction of the *GAL1* gene. The authors hypothesised that sumoylation is repressive and that its effect is due to an inhibition of acetylation and/or ubiquitination, possibly by direct competition for the substrate lysine (Nathan et al. 2006).

1.16. Crosstalk between histone modifications

Crosstalk can be defined as the effect of one residue on the modification of a nearby or distant residue. There are several possible mechanisms of crosstalk, as reviewed by Bannister and Kouzarides (2011). Firstly, the same residue may be targeted for more than one type of modification that cannot occur simultaneously such as lysines, which can be acetylated, methylated or ubiquitinated. Therefore there will be competitive antagonism between these modifications at the same site. Secondly, dependence may exist between one modification and a second modification. An example of this is the requirement of H2B monoubiquitination for H3K4 and H3K79 di- and trimethylation (discussed below). Thirdly, a modification of one site may ablate the binding of an effector protein to, or alter the catalytic activity of an enzyme against, a neighbouring site. Finally, there may be cooperation between modifications to recruit effector proteins. Each of these types of crosstalk is discussed in turn below and illustrated in Figure 8.

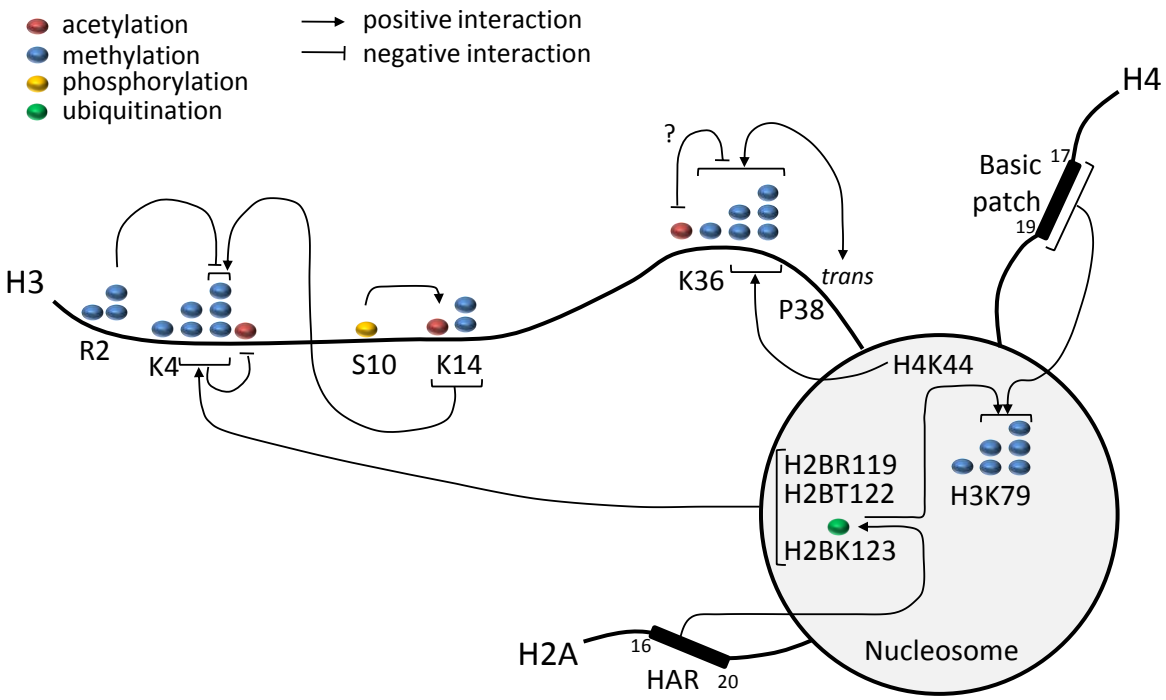


Figure 8. Histone crosstalk.

A schematic showing examples of crosstalk between histone modifications in *S. cerevisiae*. For details, please refer to the text.

1.16.1. Same residue crosstalk

The same residue cannot be modified by two different modifications simultaneously. Therefore the presence of one modification on a residue will antagonise the deposition of the other, until the first modification has been removed. If those two modifications have different functions in chromatin, it is possible that a binary switch can exist at that residue. The best-documented example of this type of crosstalk is that between H3K9 methylation and H3K9ac in mammals. These modifications map to different genomic locations, with H3K9ac primarily in euchromatin and H3K9me3 localised in heterochromatin (Wang et al. 2008). A switch in modifications at H3K9 also occurs temporally at the dihydrofolate reductase (DHFR) promoter during the cell cycle (Nicolas et al. 2003). The authors propose a model whereby two retinoblastoma-related proteins and the transcriptional co-repressor E2F mediate firstly deacetylation followed by methylation of H3K9, and this switch correlates with changes in DHFR expression.

A possible example of a binary switch in *S. cerevisiae* occurs at H3K36. With the discovery of H3K36ac by mass spectrometry and the genome-wide anti-correlation of this modification with H3K36me2, it was speculated that these two modifications may have opposite effects on chromatin and transcription (Morris et al. 2007). However, the precise function of H3K36ac is still unknown and there is no mechanistic link proposed for the putative acetyl/methyl interplay at H3K36.

It should be noted, however, that the ability of one site to be the substrate for two separate modifications does not automatically lead to functional competitive antagonism between those modifications. This was rigorously investigated during the recent characterisation of H3K4ac and its crosstalk with H3K4 methylation (Guillemette et al. 2011). Whilst deletion of *SET1* results in a global increase in H3K4ac, this effect is concentrated at gene promoters (*i.e.* not all sites at which H3K4 methylation previously occupied) and is not reciprocal (the *gcn5Δ rtt109Δ* and *hst1Δ sir2Δ* strains did not have increased or decreased H3K4 methylation respectively despite alterations in global H3K4ac levels). Indeed, since the same increase in H3K4ac is observed in the *SET1* and *BRE2* deletion strains (lacking all H3K4 methylation or only H3K4me2/3 respectively), it can be concluded that H3K4me2/3 has more effect on H3K4ac than H3K4me1, confirming that the abundance of H3K4ac cannot reflect a simple competition for substrate. Instead a mechanism was proposed in which H3K4me2/3, mapping just downstream of H3K4ac at promoters, acts to prevent the spread of H3K4ac into the coding region.

1.16.2. Interdependence of histone modifications at different residues

To investigate the interdependence of histone modifications on the histone H3 tail, Nakanishi et al. (2008) created an alanine-scanning point mutation library in which every residue in H3 was individually mutated to alanine. They identified the requirement of a lysine at H3 position 14 for proper H3K4me3 but not H3K4me1/2. A link between these residues had previously been made

since the deletion of *GCN5*, a HAT that acetylates H3K14 (among other residues), causes decreased H3K4me3 both at *ARG1* (Govind et al. 2007) and in bulk histones by mass spectrometry and Western blot (Jiang et al. 2007). However, the mechanism for this crosstalk is unknown. There is also evidence that the H3K14-H3K4 crosstalk may be reciprocal since the deletion of *SET1* resulted in the loss of H3K14ac, H3K18ac and H3K23ac, as measured by mass spectrometry (Jiang et al. 2007). Yng1, a PHD domain-containing protein in the NuA3 HAT complex, interacts specifically with H3K4me3 over H3K4me1/2 to recruit NuA3 to chromatin (Martin et al. 2006b; Taverna et al. 2006). Disruption of the Yng1-H3K4me3 interaction resulted in altered H3K14ac and transcription at a subset of genes. Furthermore, H3K4me2/3 can recruit the SAGA complex to promoters via the tandem Tudor domains in the SAGA subunit Sgf29 and this interaction is required for optimal acetylation of H3K9, H3K14 and H3K18 (Vermeulen et al. 2010; Bian et al. 2011).

An example of *trans*-tail crosstalk involving the dependence of one modification on the presence of another is the requirement of H2B ubiquitination for H3K4me2/3 (Dover et al. 2002; Sun and Allis 2002). Interestingly, H2Bub does not appear to be required for H3K4me1 (Dehé et al. 2005; Shahbazian et al. 2005). The crosstalk between H2Bub and H3K4me2/3 is not reciprocal since mutation of H3K4 does not perturb levels of H2B ubiquitination (Sun and Allis 2002). COMPASS is able to trimethylate H3K4 in the absence of H2Bub and chromatin *in vitro* (Wood et al. 2007), suggesting that the requirement for H2Bub stems from the increased complexity of the chromatin environment *in vivo*. It is unlikely that H2Bub exerts its effects simply as a result of the bulkiness of the ubiquitin moiety since substitution with the similarly-sized SUMO protein cannot mimic the effect of H2Bub on H3K4 methylation (Chandrasekharan et al. 2009). Instead, H2Bub may act as a bridge to recruit regulatory factors. There are two conflicting hypotheses on the mechanism of crosstalk between H2Bub and H3K4 methylation. The Shilatifard lab propose that the integrity of COMPASS is regulated by H2Bub through the recruitment of Swd2 to chromatin (Lee et al. 2007)

since COMPASS purified from a strain in which H2BK123 is mutated to arginine (and consequently does not have H2B ubiquitination) is deficient in Swd2. In contrast, Vitaliano-Prunier *et al.* claim that H2Bub leads to ubiquitination of Swd2 by Rad6/Bre1, which then allows the recruitment/stabilisation of Spp1 to/within COMPASS; H2Bub does not affect the Swd2-Set1 interaction (Vitaliano-Prunier et al. 2008). Two subsequent studies from the Shilatifard group lend support to the mechanism proposed by Lee *et al*. In human cells, partial knock-down of *BRE1* correlates with a partial decrease in the level of the mammalian Swd2 homologue in the SET1A/B complexes, Wdr82, on chromatin (Wu et al. 2008) and in yeast, deletion of *RAD6*, resulting in a loss of H2Bub, does not affect Spp1 association with *GAL1* or *PMA1* (Takahashi et al. 2009).

The complexity of this crosstalk was increased by the characterisation of the previously-identified H2A repression (HAR) domain (Parra and Wyrick 2007), close to H2BK123 and required for proper H2Bub and H3K4 methylation (Zheng et al. 2010). This effect is dependent on H2Bub since deletion of both *UBP8* and *UBP10* in the H2A tail-deletion strain, thus restoring H2Bub, nearly restores H3K4me3 to wildtype levels. The HAR may provide a docking site for the ubiquitination machinery, although this region is partially occluded by nucleosomal DNA and thus might only be accessible following transcription. Whilst deletion of the H2A tail does not affect the chromatin recruitment of Paf1, Rad6/Bre1 or Bre2, there is a three-fold reduction in the level of Swd2 associated with Bre2 (COMPASS), in agreement with Lee *et al.*

Recently, a ubiquitin-independent effect of H2B on H3K4 methylation has been described. Chandrasekharan *et al.* discovered that a helix in the C-terminal tail of H2B is required for H3K4me3 (Chandrasekharan et al. 2010). Mutation of H2BR119 or H2BT122 (on the same face of the helix) to alanine or aspartate respectively causes an increase in H2Bub but a decrease in H3K4me3, confirming the ubiquitin-independence of this effect. Furthermore, in the double H2BR119A/T122D mutant, there is decreased binding of recombinantly-expressed Spp1 to H2B *in*

vitro, and decreased *in vivo* chromatin association of several COMPASS subunits, including Set1, Swd1 and Swd2. The authors hypothesise that, rather than a direct role of H2Bub in COMPASS recruitment, ubiquitination of H2B may instead indirectly influence H3K4 methylation by increasing the stability of the nucleosome (Chandrasekharan et al. 2009) and so preserving the docking site for COMPASS in the H2B C-terminal helix.

Similar to the crosstalk with H3K4 methylation, H2B ubiquitination is required for H3K79me2/3 (Briggs 2002, Ng 2002) but not H3K79me1 (Shahbazian et al. 2005). Because H3K79 is positioned on the same face of the nucleosome as H2BK123, it is possible to imagine how H2B ubiquitination might directly influence H3K79 methylation by Dot1. Indeed, a lysine-rich region on Dot1 has been shown to interact directly with ubiquitin (Oh et al. 2010), although Dot1 associates with chromatin independently of H2Bub (Shahbazian et al. 2005). The presence of ubiquitin may therefore stabilise Dot1 on the nucleosome or induce a conformational change in Dot1 to allow processive H3K79 methylation. The involvement of Swd2 in the H2Bub-H3K79 crosstalk has also been proposed since Swd2 interacts physically with Dot1, and *swd2*^{ts} mutants decrease H3K79me3 (Lee et al. 2007). Interestingly, the H2B C-terminal helix does not seem to have a role in the regulation of H3K79 methylation, uncoupling the regulation of H3K4 and H3K79 methylation marks (Chandrasekharan et al. 2010).

Other examples of *trans*-tail regulation include the requirement for a basic residue at H4 position 44 and several residues on histones H3 and H2A for Set2-induced H3K36me2/3 (Du et al. 2008; Du and Briggs 2010; Hainer and Martens 2011; Endo et al. 2012) and a basic patch on histone H4 for Dot1-induced H3K79me2/3 (Fingerman et al. 2007).

1.16.3. Modification of enzyme/protein effector binding

The H3R2me2a genomic distribution anti-correlates with that of H3K4me3, being present in heterochromatin, inactive euchromatic genes and the 3' ends of moderately transcribed genes (Kirmizis et al. 2007). Upon further investigation, it was found that H3R2me2a or mutation of H3R2 to alanine inhibits the binding of the Spp1 PHD domain to a peptide methylated on H3K4. Moreover, a ChIP experiment demonstrated that Spp1 is absent from regions with H3R2me2a, even if H3K4me2 is present. The authors propose a model in which H3R2me2a prevents H3K4me3 deposition (but not H3K4me1/2) until removal of H3R2me2a (by an unknown mechanism) allows Spp1 binding, stabilisation of COMPASS on chromatin and subsequent H3K4me3 (Kirmizis et al. 2007). In human cells, a similar mechanism for the inhibitory relationship between H3R2me2a and H3K4me3 has also been described (Hyllus et al. 2007; Guccione et al. 2007) and a reduction in H3R2me2a was shown to alter the expression of the *HOXA5* and cyclin D1 genes regulated by the H3K4 methylation-bound effector proteins WDR5/MLL and ING2 respectively (Iberg et al. 2008). Interestingly, there is no negative crosstalk between H3R2me1 and H3K4 methylation, which have genomic co-localisation (Kirmizis et al. 2009).

A second example of crosstalk in this class is the regulation of Set2 activity towards H3K36 by the isomerisation state of the neighbouring proline residue at H3P38 (Nelson et al. 2006). H3P38 is necessary for proper H3K36me3 and a catalytic mutant of the proline isomerase Fpr4, presumably increasing the proportion of H3P38*trans,* causes an increase in H3K36 methylation *in vivo*. The authors propose that Set2 preferentially recognises the H3 tail when H3P38 is in the *trans* conformation and speculate that the H3K36 demethylase Jhd1 may bind H3P38*cis*. This regulation may be reciprocal since no Fpr4 activity was detected towards a K36-methylated peptide in comparison with an unmethylated H3 tail peptide.

1.16.4. Cooperative recruitment of effector proteins

The use of a dual-specificity antibody demonstrated that Snf1-catalysed H3S10ph and Gcn5-catalysed H3K14ac occur on the same histone tail and are required for proper *INO1* expression (Lo et al. 2001). A peptide affinity assay identified the *S. cerevisiae* 14-3-3 proteins Bmh1/2 as H3S10ph/K14ac-interacting proteins and importantly, 14-3-3 binding to histone H3 is dependent on both modifications (Walter et al. 2008). A similar interaction is found in mammals and an siRNA-mediated knockdown of 14-3-3 proteins abolishes expression of the HDAC1 gene, indicating their importance in transcriptional regulation (Winter et al. 2008). 14-3-3 proteins act as adaptor proteins, most commonly in signal transduction pathways, and interact with a large number of proteins to influence their conformations, protein interactions, enzymatic activities and subcellular localisation (as reviewed in van Heusden 2009).

1.17. Histones and disease

Due to their contribution to the regulation of transcription and therefore gene expression, misregulation of histone protein modification is associated with several diseases.

The involvement of H3K4 methylation in gene regulation is highlighted by the effect of aberrant H3K4me levels in certain cancers. SMYD3, an H3K4-specific KMT, is frequently upregulated in colorectal and hepatocellular carcinomas (Hamamoto et al. 2004). An siRNA-mediated knock-down of SYMD3 represses cell growth whereas overexpression of SMYD3 in cancer cell lines causes increased cell growth and the promotion of transformation. This latter effect of SMYD3 requires the methyltransferase activity. Indeed, high levels of H3K4me3 have been associated with poorer prognosis of patients with hepatocellular carcinomas (He et al. 2012). SMYD3 interacts with the RNA polymerase II complex subunit HELZ RNA helicase to cause transcriptional activation of a number of genes including oncogenes, homeobox genes and cell cycle-regulated genes (Hamamoto et al. 2004).

Another line of evidence for the connection between H3K4me3 misregulation and cancer is the potential tumour suppressor role of the H3K4me3-interacting proteins ING1-5, which have been shown to cooperate with p53 to bring about growth arrest, apoptosis and cellular senescence (reviewed in Shi and Gozani 2005). Decreased levels of ING1, ING3 and ING4 are associated with several cancers including melanoma, glioma, breast and gastric cancer (Garkavtsev et al. 2004; Shi and Gozani 2005). Furthermore, mutations in the ING PHD domains, which abolish binding to H3K4me3, are found in tumours (Shi and Gozani 2005), indicating that the ING H3K4me3-binding activity is important. These apparently opposite effects of H3K4me3 in cancer emphasise the significance of the context of a modification on its downstream consequences.

Fusion proteins formed as a result of chromosomal translocations involving the H3K4 methyltransferase MLL gene occur in roughly 80 % of infants with acute myeloid leukaemia (AML) or acute lymphoblastic leukaemia (ALL) (Shilatifard 2006). Over 60 fusion partners have been identified for MLL, and these are mostly nuclear (Meyer et al. 2009). The MLL chimeras have the greatest effect on the expression of the *HOX* genes and this *HOX* gene misregulation is required for leukaemogenesis (Yu et al. 1995; Ferrando et al. 2003). However, the MLL fusion proteins do not affect levels of histone methylation, indicating that the MLL chimeras must function differently to affect gene expression in these leukaemias. Several MLL translocation partners are associated with the super elongation complex (SEC) and the SEC is required for *HOX* gene overexpression by these MLL fusion proteins (Lin et al. 2010). This indicates that the misregulation of transcription elongation, for example by aberrant recruitment of the elongation machinery to MLL target genes, may be involved in leukaemogenesis. There are other mechanisms by which MLL chimeras may cause leukaemias such as the MLL-AF4, -AF9, -AF10 and -ENL fusions, which have been proposed to promote leukaemogenesis as a result of their interaction with hDOT1L, leading to aberrant H3K79 methylation and persistent activation of the *HOX9* proto-oncogenes

(Okada et al. 2005; Zhang et al. 2006; Bitoun et al. 2007; Mueller et al. 2007; Krivtsov et al. 2008; Mueller et al. 2009).

Misregulation of other histone-modifying enzymes can also be oncogenic. HDAC1-3 and 6 are overexpressed in prostate, gastric, breast and cervical cancer (reviewed in Bolden et al. 2006) and encouragingly, HDAC inhibitors are currently used in the treatment of certain cancers in the clinic (reviewed in Ahmad et al. 2012). Oncogenic translocations involving other histone-modifying enzymes include the MLL-CBP, MLL-p300 (both HATs) and NUP98-NSD1 (H3K36-specific KMT) fusions that promote haematopoietic transformation and AML (Iyer et al. 2004; Wang et al. 2007).

Diseases apart from cancer are also associated with defects in histone modifications. NSD1 haploinsufficiency causes Sotos syndrome, a neurological disorder associated with an increased risk of hepatocellular carcinoma, leukaemia and neuroblastoma (reviewed in Faravelli 2005). JARID1C, a JmjC-type H3K4 demethylase is frequently mutated in a manner that abolishes demethylase activity in patients with X-linked mental retardation (Iwase et al. 2007). These examples demonstrate the importance of the study of histone modifications and their regulation to increase the understanding of human disease. Hopefully the increased knowledge will aid in the development of drug intervention, providing more examples like the use of HDAC inhibitors in the treatment of certain cancers.

1.18. Aims

The work in this thesis aims to increase the understanding of the regulation of H3K4me3. The control of this modification by two residues in the histone H3 tail, H3K14 and H3P16, will be investigated using gene-specific and genome-wide approaches, and a mechanism for this crosstalk will be proposed. Additionally, a newly discovered modification in *S. cerevisiae*, H3K18me1, will be characterised and its involvement in the regulation of H3K4me3 will be explored.

Chapter 2.

Materials and Methods

2. Materials and Methods

2.1. Yeast transformation

All the strains created in this study were made using the homologous recombination-based method described in Longtine et al. (1998). For gene deletion, PCR was performed using Phusion DNA polymerase (New England Biolabs) on plasmid templates encoding an appropriate selectable marker and PCR primers containing 40 nucleotides of sequence homologous to the regions flanking the desired site of disruption in the yeast genome. This PCR product was then transformed into yeast using a LiAc/heat shock transformation method. The transformed yeast were plated onto the appropriate selection plate and positive transformants were confirmed by PCR. The protein tagging and *pGAL1*-driven gene expression strains were created using plasmids containing the desired construct adjacent to the selectable marker.

2.2. Yeast strains

All strains used in this study are detailed in Table 1.

Strain	Parent	Genotype	Origin
WTH3	-	*MAT**α** his3Δ200 leu2Δ0 lys2Δ0 trp1Δ63 ura3Δ0 met15Δ0 can1::MFA1pr-HIS3 hht1-hhf1::NatMX4 hht2-hhf2::[HHTS-HHFS]-URA3*	Open Biosystems
R2A	WTH3	*hht2-hhf2::[HHTS R2A-HHFS]-URA3*	Open Biosystems
K4A	WTH3	*hht2-hhf2::[HHTS K4A-HHFS]-URA3*	Open Biosystems
K9A	WTH3	*hht2-hhf2::[HHTS K9A-HHFS]-URA3*	Open Biosystems
S10A	WTH3	*hht2-hhf2::[HHTS S10A-HHFS]-URA3*	Open Biosystems
K14A	WTH3	*hht2-hhf2::[HHTS K14A-HHFS]-URA3*	Open Biosystems
K14R	WTH3	*hht2-hhf2::[HHTS K14R-HHFS]-URA3*	Open Biosystems
K14Q	WTH3	*hht2-hhf2::[HHTS K14Q-HHFS]-URA3*	Open Biosystems
K18A	WTH3	*hht2-hhf2::[HHTS K18A-HHFS]-URA3*	Open Biosystems
K18R	WTH3	*hht2-hhf2::[HHTS K18R-HHFS]-URA3*	Open Biosystems
K18Q	WTH3	*hht2-hhf2::[HHTS K18Q-HHFS]-URA3*	Open Biosystems
K23A	WTH3	*hht2-hhf2::[HHTS K23A-HHFS]-URA3*	Open Biosystems
K27A	WTH3	*hht2-hhf2::[HHTS K27A-HHFS]-URA3*	Open Biosystems
K36A	WTH3	*hht2-hhf2::[HHTS K36A-HHFS]-URA3*	Open Biosystems
K56A	WTH3	*hht2-hhf2::[HHTS K56A-HHFS]-URA3*	Open Biosystems
K79A	WTH3	*hht2-hhf2::[HHTS K79A-HHFS]-URA3*	Open Biosystems
NSY429	-	*MAT**α** ura3-52 leu2-3,112 trp1-289 his3Δ1* (*hht1-hhf1*) (*hht2-hhf2*) plus pNS329 (*CEN TRP1 HHF1-HHT1*)	Sabet et al. (2003)
H3 Δ1-28	NSY429	plus pNS329 (*hht2 Δ1-28*)	Sabet et al. (2003)

Strain	Parent	Genotype	Origin
WT HHT	K14A	*hhts K14A-hhfs::HHTS-HHFS-KanMX6*	This study
K14A HHT	WTH3	*hhts-hhfs::HHTS K14A-HHFS-KanMX6*	This study
K14R HHT	WTH3	*hhts-hhfs::HHTS K14R-HHFS-KanMX6*	This study
K14Q HHT	WTH3	*hhts-hhfs::HHTS K14Q-HHFS-KanMX6*	This study
P16V HHT	WTH3	*hhts-hhfs::HHTS P16V-HHFS-KanMX6*	This study
K18A HHT	WTH3	*hhts-hhfs::HHTS K18A-HHFS-KanMX6*	This study
K18L HHT	WTH3	*hhts-hhfs::HHTS K18L-HHFS-KanMX6*	This study
K14A P16V	WTH3	*hhts-hhfs::HHTS K14A P16V-HHFS-KanMX6*	This study
K14R P16V	WTH3	*hhts-hhfs::HHTS K14R P16V-HHFS-KanMX6*	This study
K14Q P16V	WTH3	*hhts-hhfs::HHTS K14Q P16V-HHFS-KanMX6*	This study
K14A K18A	WTH3	*hhts-hhfs::HHTS K14A K18A-HHFS-KanMX6*	This study
K14R K18A	WTH3	*hhts-hhfs::HHTS K14R K18A-HHFS-KanMX6*	This study
K14Q K18A	WTH3	*hhts-hhfs::HHTS K14Q K18A-HHFS-KanMX6*	This study
P16V K18A	WTH3	*hhts-hhfs::HHTS P16V K18A-HHFS-KanMX6*	This study
WTP16	WZY42	*MAT**a** ura3-52, lys2-801, ade2-101, trp1Δ63, his3Δ200, leu2Δ1, hht1-hhf1::LEU2, hht2-hhf2::HIS3, YCp50 [HHT2-HHF2]*	Nelson et al. (2006)
P16V	WTP16	*plus YCp50 [HHT2 P16V-HHF2]*	Nelson et al. (2006)
WTH3 *bre1Δ*	WTH3	*bre1::TRP1*	This study
K14A *bre1Δ*	K14A	*bre1::TRP1*	This study
K14R *bre1Δ*	K14R	*bre1::TRP1*	This study
K14Q *bre1Δ*	K14Q	*bre1::TRP1*	This study
WTH3 *bre1Δ jhd2Δ*	WTH3 *bre1Δ*	*jhd2::LEU2*	This study
K14A *bre1Δ jhd2Δ*	K14A *bre1Δ*	*jhd2::LEU2*	This study
K14R *bre1Δ jhd2Δ*	K14R *bre1Δ*	*jhd2::LEU2*	This study
K14Q *bre1Δ jhd2Δ*	K14Q *bre1Δ*	*jhd2::LEU2*	This study
WTH3 *bre2Δ*	WTH3	*bre2::TRP1*	This study
WTH3 *jhd2Δ*	WTH3	*jhd2::KanMX6*	This study
K14A *jhd2Δ*	K14A	*jhd2::KanMX6*	This study
K14R *jhd2Δ*	K14R	*jhd2::KanMX6*	This study
K14Q *jhd2Δ*	K14Q	*jhd2::KanMX6*	This study
P16V *jhd2Δ*	P16V HHT	*jhd2::LEU2*	This study
WTH3 *pGAL1-3HA-JHD2*	WTH3	*KanMX6-pGAL1-3HA-JHD2*	This study
K14A *pGAL1-3HA-JHD2*	K14A	*KanMX6-pGAL1-3HA-JHD2*	This study
K14R *pGAL1-3HA-JHD2*	K14R	*KanMX6-pGAL1-3HA-JHD2*	This study
K14Q *pGAL1-3HA-JHD2*	K14Q	*KanMX6-pGAL1-3HA-JHD2*	This study
WTH3 Jhd2-HA	WTH3	*JHD2-3HA-KanMX6*	This study
K14A Jhd2-HA	K14A	*JHD2-3HA-KanMX6*	This study
K14R Jhd2-HA	K14R	*JHD2-3HA-KanMX6*	This study
K14Q Jhd2-HA	K14Q	*JHD2-3HA-KanMX6*	This study
WTH3 *JHD2-9Myc*	WTH3 *jhd2Δ*	*plus pFH17 (JHD2-9Myc, TRP1)*	This study
WTH3 *JHD2 H261A-9Myc*	WTH3 *jhd2Δ*	*plus pFH20 (JHD2 H261A-9Myc, TRP1)*	This study
WTH3 *rad6Δ*	WTH3	*rad6::KanMX6*	This study
K14A *rad6Δ*	K14A	*rad6::KanMX6*	This study
K14R *rad6Δ*	K14R	*rad6::KanMX6*	This study
K14Q *rad6Δ*	K14Q	*rad6::KanMX6*	This study
WTH3 *rad6Δ jhd2Δ*	WTH3 *rad6Δ*	*jhd2::TRP1*	This study
K14A *rad6Δ jhd2Δ*	K14A *rad6Δ*	*jhd2::TRP1*	This study
K14R *rad6Δ jhd2Δ*	K14R *rad6Δ*	*jhd2::TRP1*	This study
K14Q *rad6Δ jhd2Δ*	K14Q *rad6Δ*	*jhd2::TRP1*	This study
WTH3 *spp1Δ*	WTH3	*spp1::TRP1*	This study
K14A *spp1Δ*	K14A	*spp1::TRP1*	This study
K14R *spp1Δ*	K14R	*spp1::TRP1*	This study
K14Q *spp1Δ*	K14Q	*spp1::TRP1*	This study
P16V *spp1Δ*	P16V HHT	*spp1::TRP1*	This study
K18A *spp1Δ*	K18A	*spp1::TRP1*	This study
K18R *spp1Δ*	K18R	*spp1::TRP1*	This study
K18Q *spp1Δ*	K18Q	*spp1::TRP1*	This study
K18L *spp1Δ*	K18L HHT	*spp1::TRP1*	This study
WTH3 *pGAL1-3HA-SPP1*	WTH3	*KanMX6-pGAL1-3HA-SPP1*	This study

Strain	Parent	Genotype	Origin
K14A *pGAL1-3HA-SPP1*	K14A	*KanMX6-pGAL1-3HA-SPP1*	This study
K14R *pGAL1-3HA-SPP1*	K14R	*KanMX6-pGAL1-3HA-SPP1*	This study
K14Q *pGAL1-3HA-SPP1*	K14Q	*KanMX6-pGAL1-3HA-SPP1*	This study
WTH3 *spp1Δ jhd2Δ*	WTH3 *jhd2Δ*	*spp1::TRP1 jhd2::KanMX6*	This study
WTH3 *pGAL1-3HA-SPP1 jhd2Δ*	*WTH3 pGAL1-3HA-SPP1*	*jhd2::TRP1*	This study
WTH3 *pGAL1-3HA-JHD2 spp1Δ*	WTH3 *pGAL1-3HA-JHD2*	*spp1::TRP1*	This study
WTH3 *pGAL1-3HA-JHD2 pGAL1-3HA-SPP1*	WTH3 *pGAL1-3HA-JHD2*	*TRP1-pGAL1-3HA-SPP1*	This study
WTH3 Spp1-HA	WTH3	*SPP1-3HA-KanMX6*	This study
K14A Spp1-HA	K14A	*SPP1-3HA-KanMX6*	This study
K14R Spp1-HA	K14R	*SPP1-3HA-KanMX6*	This study
K14Q Spp1-HA	K14Q	*SPP1-3HA-KanMX6*	This study
P16V Spp1-HA	P16V HHT	*SPP1-3HA-TRP1*	This study
WTP16 Spp1-HA	WTP16	*SPP1-3HA-KanMX6*	This study
P16V Spp1-HA	P16V	*SPP1-3HA-KanMX6*	This study
WTH3 Swd1-HA	WTH3	*SWD1-3HA-KanMX6*	This study
K14A Swd1-HA	K14A	*SWD1-3HA-KanMX6*	This study
K14R Swd1-HA	K14R	*SWD1-3HA-KanMX6*	This study
K14Q Swd1-HA	K14Q	*SWD1-3HA-KanMX6*	This study
P16V Swd1-HA	P16V HHT	*SWD1-3HA-TRP1*	This study
BY4741	-	*MAT**a** his3Δ leu2Δ met15Δ ura3Δ*	Euroscarf
abp140Δ	BY4741	*abp140::KanMX4*	Open Biosystems
aml1Δ	BY4741	*aml1::KanMX4*	Open Biosystems
bdh2Δ	BY4741	*bdh2::KanMX4*	Open Biosystems
bio1Δ	BY4741	*bio1::KanMX4*	Open Biosystems
bre2Δ	BY4741	*bre2::KanMX4*	Open Biosystems
coq3Δ	BY4741	*coq3::KanMX4*	Open Biosystems
coq5Δ	BY4741	*coq5::KanMX4*	Open Biosystems
crg1Δ	BY4741	*crg1::KanMX4*	Open Biosystems
ctm1Δ	BY4741	*ctm1::KanMX4*	Open Biosystems
dot1Δ	BY4741	*dot1::KanMX4*	Open Biosystems
dph5Δ	BY4741	*dph5::KanMX4*	Open Biosystems
ecm5Δ	BY4741	*ecm5::KanMX4*	Open Biosystems
ecm31Δ	BY4741	*ecm31::KanMX4*	Open Biosystems
elp3Δ	BY4741	*elp3::KanMX4*	Open Biosystems
erg6Δ	BY4741	*erg6::KanMX4*	Open Biosystems
gis1Δ	BY4741	*gis1::KanMX4*	Open Biosystems
hmt1Δ	BY4741	*hmt1::KanMX4*	Open Biosystems
hsl7Δ	BY4741	*hsl7::KanMX4*	Open Biosystems
irc15Δ	BY4741	*irc15::KanMX4*	Open Biosystems
jhd1Δ	BY4741	*jhd1::KanMX4*	Open Biosystems
jhd2Δ	BY4741	*jhd2::KanMX4*	Open Biosystems
kar4Δ	BY4741	*kar4::KanMX4*	Open Biosystems
lip5Δ	BY4741	*lip5::KanMX4*	Open Biosystems
met1Δ	BY4741	*met1::KanMX4*	Open Biosystems
mht1Δ	BY4741	*mht1::KanMX4*	Open Biosystems
mtf1Δ	BY4741	*mtf1::KanMX4*	Open Biosystems
mtq1Δ	BY4741	*mtq1::KanMX4*	Open Biosystems
mtq2Δ	BY4741	*mtq2::KanMX4*	Open Biosystems
nnt1Δ	BY4741	*nnt1::KanMX4*	Open Biosystems
oms1Δ	BY4741	*oms1::KanMX4*	Open Biosystems
ppm1Δ	BY4741	*ppm1::KanMX4*	Open Biosystems
rkm1Δ	BY4741	*rkm1::KanMX4*	Open Biosystems
rkm2Δ	BY4741	*rkm2::KanMX4*	Open Biosystems
rkm3Δ	BY4741	*rkm3::KanMX4*	Open Biosystems
rkm4Δ	BY4741	*rkm4::KanMX4*	Open Biosystems
rmt1Δ	BY4741	*rmt1::KanMX4*	Open Biosystems
rph1Δ	BY4741	*rph1::KanMX4*	Open Biosystems

Strain	Parent	Genotype	Origin
rpl42aΔ	BY4741	*rpl42a::KanMX4*	Open Biosystems
rrp8Δ	BY4741	*rrp8::KanMX4*	Open Biosystems
rsm22Δ	BY4741	*rsm22::KanMX4*	Open Biosystems
sam4Δ	BY4741	*sam4::KanMX4*	Open Biosystems
sdc1Δ	BY4741	*sdc1::KanMX4*	Open Biosystems
see1Δ	BY4741	*see1::KanMX4*	Open Biosystems
set1Δ	BY4741	*set1::KanMX4*	Open Biosystems
set2Δ	BY4741	*set2::KanMX4*	Open Biosystems
set3Δ	BY4741	*set3::KanMX4*	Open Biosystems
set4Δ	BY4741	*set4::KanMX4*	Open Biosystems
set5Δ	BY4741	*set5::KanMX4*	Open Biosystems
set6Δ	BY4741	*set6::KanMX4*	Open Biosystems
set7Δ	BY4741	*set7::KanMX4*	Open Biosystems
slm3Δ	BY4741	*slm3::KanMX4*	Open Biosystems
spp1Δ	BY4741	*spp1::KanMX4*	Open Biosystems
ste14Δ	BY4741	*ste14::KanMX4*	Open Biosystems
swd1Δ	BY4741	*swd1::KanMX4*	Open Biosystems
swd3Δ	BY4741	*swd3::KanMX4*	Open Biosystems
tae1Δ	BY4741	*tae1::KanMX4*	Open Biosystems
tmt1Δ	BY4741	*tmt1::KanMX4*	Open Biosystems
trm2Δ	BY4741	*trm2::KanMX4*	Open Biosystems
trm4Δ	BY4741	*trm4::KanMX4*	Open Biosystems
trm10Δ	BY4741	*trm10::KanMX4*	Open Biosystems
trm11Δ	BY4741	*trm11::KanMX4*	Open Biosystems
trm12Δ	BY4741	*trm12::KanMX4*	Open Biosystems
ybr141cΔ	BY4741	*ybr141c::KanMX4*	Open Biosystems
ybr225wΔ	BY4741	*ybr225w::KanMX4*	Open Biosystems
ybr271wΔ	BY4741	*ybr271w::KanMX4*	Open Biosystems
ygr283cΔ	BY4741	*ygr283c::KanMX4*	Open Biosystems
yhl039wΔ	BY4741	*yhl039w::KanMX4*	Open Biosystems
yil110wΔ	BY4741	*yil110w::KanMX4*	Open Biosystems
yjr129cΔ	BY4741	*yjr129c::KanMX4*	Open Biosystems
ykl162cΔ	BY4741	*ykl162c::KanMX4*	Open Biosystems
ylr063wΔ	BY4741	*ylr063w::KanMX4*	Open Biosystems
ylr137wΔ	BY4741	*ylr137w::KanMX4*	Open Biosystems
ymr209cΔ	BY4741	*ymr209c::KanMX4*	Open Biosystems
ymr310cΔ	BY4741	*ymr310c::KanMX4*	Open Biosystems
ynl022cΔ	BY4741	*ynl022c::KanMX4*	Open Biosystems
ynl024cΔ	BY4741	*ynl024c::KanMX4*	Open Biosystems
ynl092wΔ	BY4741	*ynl092w::KanMX4*	Open Biosystems
ynr029cΔ	BY4741	*ynr029c::KanMX4*	Open Biosystems
yor021cΔ	BY4741	*yor021c::KanMX4*	Open Biosystems
bre2Δ jhd2Δ	*bre2Δ*	*jhd2::His3MX6*	This study (A. Nair)
sdc1Δ jhd2Δ	*sdc1Δ*	*jhd2::His3MX6*	This study (A. Nair)
set1Δ jhd2Δ	*set1Δ*	*jhd2::His3MX6*	This study (A. Nair)
spp1Δ jhd2Δ	*spp1Δ*	*jhd2::His3MX6*	This study (A. Nair)
swd1Δ jhd2Δ	*swd1Δ*	*jhd2::His3MX6*	This study (A. Nair)
swd3Δ jhd2Δ	*swd3Δ*	*jhd2::His3MX6*	This study (A. Nair)
pGAL1-3HA-RPH1	BY4741	*His3MX6-pGAL1-3HA-RPH1*	This lab
W3031a	-	*MAT**α** leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	
set1Δ	W3031a	*set1::KanMX4*	This lab

Table 1. Genotypes of the *S. cerevisiae* strains used in this study.

2.3. Creation of the HHT-Kan double mutant strains

The *HHTS-HHFS* locus was amplified from the Open Biosystems histone mutant strains (J Dai et al. 2008). These fragments were then digested with EcoRI and SpeI at the endogenous restriction sites flanking the locus. The digested fragments were ligated into EcoRI/SpeI-digested pFA6 (KanMX6) (Longtine et al. 1998). Histone H3 double substitution mutants were created using the single mutants as templates for Quikchange site-directed mutagenesis (Stratagene). The resulting plasmids were confirmed by sequencing before use as PCR templates to replace the genomic *HHTS-HHFS* with the newly mutated locus according to the standard protocol (section 2.1). Transformed strains were sequenced once more before use in experiments.

2.4. DNA extraction and PCR

Cells from 500 µl overnight YPD cultures (approximately 7 x 10^7 cells) were pelleted by centrifugation (13 000 rpm, 1 min) and resuspended in 200 µl DNA extraction buffer (100 mM NaCl, 10 mM Tris-Cl pH 8.0, 1 mM EDTA pH 8.0, 1 % SDS, 2 % Triton X-100) and 200 µl phenol:chloroform:isoamyl alcohol. The cells were broken by vortexing with 1 scoop acid-washed glass beads for 10 min. After centrifugation (13 000 rpm, 10 min, 4°C), the upper aqueous layer was removed and the DNA was precipitated in ammonium acetate/ethanol. The DNA pellet was ethanol-washed and dried before resuspension in 300 µl water. PCR reactions for strain checking were performed using Taq polymerase (Sigma).

2.5. Yeast growth conditions

Unless otherwise stated, yeast cells were grown at 30°C, shaking at 200 rpm to exponential phase in YPD medium (1 % yeast extract (Difco), 1 % bactopeptone supplemented with 2 % glucose). Cells were grown overnight in 5 ml cultures before dilution to 0.2 OD_{600} (0.5 x 10^7 cells/ml) in the appropriate volume of fresh media. Cells were harvested during exponential growth at 0.5 OD_{600} (1.25 x 10^7 cells/ml) by centrifugation (3000 rpm, 5 min). Overexpression of genes placed under

the control of the *GAL1* promoter was achieved by growing cells as above except both overnight and exponential growth was in YP medium supplemented with 2 % galactose (YPG). To induce *MET16*, cells were grown overnight and then to exponential phase in complete synthetic medium (CSM) supplemented with 0.2 % yeast nitrogen base without amino acids (Difco) and 2 % glucose before switching of the media to CSM lacking methionine. The *GAL* locus was induced after first growing cells in YPD to exponential phase and washing in water before transferring to YPG for induction. Analysis of yeast during the diauxic shift was achieved by growing the yeast for 25 h in YPD. The assessment of yeast growth by drop plate assays was performed by first normalising the cell number in overnight yeast cultures before pipetting 1.5 µl of five serial ten-fold dilutions onto YPD-agar plates. Plates were incubated for 2 days at 30°C. Assessment of growth in liquid culture (CSM) was achieved using a Bioscreen spectrophotometer that automatically measures the optical density of cultures at 30°C every 20 min for 24 h.

2.6. Preparation of whole cell extracts

Cells were grown at 30°C in 25 ml of the appropriate media to 1.25 x 10^7 cells/ml. Whole cell extracts were prepared by vortexing the cells with glass beads in 300 µl 8 M urea, 240 µl loading buffer (100 mM Tris-Cl pH 6.8, 20 % glycerol, 4 % SDS, 0.1 % bromophenol blue) and 60 µl 1 M DTT for 3 min followed by boiling for 5 min.

2.7. Western blotting and dot blots

For Western blotting, proteins were separated on 10-15 % SDS-polyacrylamide gels and transferred to nitrocellulose membranes using a semi-dry transfer method. For dot blots, appropriate amounts of peptide were pipette onto nitrocellulose membrane and allowed to dry. Membranes were blocked in 5 % bovine serum albumin (BSA) in TBST (20 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.1 % TWEEN-20). Antibodies were added in 2.5 % BSA/TBST. The primary antibodies used are detailed in Table 2. The HRP-conjugated rabbit, mouse or rat secondary antibodies

(Sigma) were used at 1:4000 dilutions. Interactions were visualised using chemiluminescence (Pierce) and exposure to X-ray film.

Antibody	Catalogue number	Species reactivity	Dilution for Western blot	Amount for ChIP (µl)
H3	Mi 07-690	Rabbit	1:2000	5
H3R2me2a	ab80075	Rabbit	1:1000	-
H3K4me2	Mi 07-030	Rabbit	1:2000	5
H3K4me3	Mi 05-745R	Rabbit	1:2000	5
H3K4me3 (K14R validation only, section 3.2.1)	Diagenode CS-003-100	Rabbit	1:1000	-
H3K9ac	Mi 07-352	Rabbit	1:2000	-
H3S10ph	Mi 05-817	Rabbit	1:2000	-
H3K14ac	Mi 07-353	Rabbit	1:2000	-
H3K18ac	Mi 07-354	Rabbit	1:4000	5
H3K18me1 (original)	-	Rabbit	-	40
H3K18me1 (new)	-	Rabbit	1:1000	50
H3K23ac	Mi 07-355	Rabbit	1:2000	-
H3K27ac	Mi 07-360	Rabbit	1:2000	-
H3K36me2	Mi 07-369	Rabbit	1:2000	-
H3K36me3	ab9050	Rabbit	1:2500	5
H3K79me2	ab3594	Rabbit	1:2000	-
H3K79me3	ab2621	Rabbit	1:2000	-
HA (for Western)	Roche 3F10	Rat	1:500	-
HA (for ChIP)	ab9110	Rabbit	-	5
Myc	Sigma M4439	Mouse	1:1000	-
Actin	Mi MAB1501	Mouse	1:1000	-

Table 2. Details of the antibodies used during this study.

2.8. Production and affinity purification of the first H3K18me1 antibody

Peptide synthesis and antibody production were performed by Covalab UK Ltd. The H3K18me1 antibody was raised against a peptide corresponding to the region of the N-terminal tail of histone H3 surrounding lysine 18 (cGKAPR**K(me)**QLASKAARK). The peptide was conjugated to keyhole limpet haemocyanin (KLH) carrier protein via the N-terminal cysteine and injected into a rabbit. The final bleed was collected after 105 days. Two preparations of antibody were used in this project: the serum from the final bleed (for ChIP) and the serum after affinity-purification against the mono-methylated peptide (for Western blot).

2.9. Production and affinity purification of the second H3K18me1 antibody

Peptide synthesis was performed by GL Biochem (Shanghai) Ltd. The antibody was raised against the following histone H3 peptide: TGGKAPR**K(me)**QLASKAARc. The peptide was conjugated in-house to maleimide-activated KLH (Fisher) via the C-terminal cysteine residue. Antibody production was carried out by the Scottish National Blood Transfusion Service. After injection of the rabbit with the conjugated peptide, four bleeds were collected before the final bleed. The fourth bleed serum was affinity purified against the methylated peptide conjugated to sulfolink resin (Pierce). The elution fractions containing antibody were pooled and concentrated to 0.5 mg/ml.

2.10. Chromatin immunoprecipitation

ChIP was performed as described in Morillon et al. (2005). Briefly, cells grown to OD_{600} 0.5 in 50 ml of appropriate media were fixed with 1 % formaldehyde in 45 ml PBS for 30 min at 22°C followed by addition of 125 mM glycine for 5 min. Cell pellets were collected by centrifugation (3000 rpm, 5 min) before washing twice with 10 ml cold PBS. Cells were resuspended in 500 µl cold FA-150 buffer (10 mM HEPES pH 7.9, 150 mM NaCl, 0.1 % SDS, 0.1 % sodium deoxycholate, 1 % Triton X-100) and broken using 1 ml glass beads on a MagnaLyser (Roche) at 4°C. Sample volume was increased to 2 ml with FA-150 buffer before shearing of the fixed chromatin by sonication using a biorupter (Diagenode, 30 min, 1 min on, 20s off, medium setting). Chromatin was cleared by centrifugation (10 000 rpm, 15 min, 4°C) and incubated with antibody (amounts stated in Table 2) in 1.5 ml siliconised Eppendorf tubes for 15-20 h rotating at 4°C. Bound chromatin was immunoprecipitated for 90 min at 22°C with 50 µl protein A-Sepharose pre-blocked with sonicated salmon sperm DNA. Beads and attached chromatin were pelleted by centrifugation (2600 rpm, 1 min) and washed with TSE-150 buffer (20 mM Tris-Cl pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1 % SDS, 1 % Triton X-100) for 3 min, TSE-500 buffer (20 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM EDTA, 0.1 % SDS, 1 % Triton X-100) for 3 min, LiCl buffer (0.25 M LiCl, 10 mM

Tris-Cl pH 8.0, 1 mM EDTA, 1 % dioxycholate, 1 % NP-40) for 15 min and twice with TE, all at 22°C. After washing, chromatin was eluted from the beads for 30 min at 65°C with elution buffer (0.1 M NaHCO3, 1 % SDS). Addition of 350 mM NaCl and incubation for 3 h at 65°C reversed the cross-links before treatment of samples with RNase A for 1 h at 37°C and proteinase K overnight at 65°C. DNA was purified using a PCR-purification kit (Qiagen) and eluted in 400 µl 1 mM Tris-Cl pH 8.0. Input DNA was diluted accordingly. Real-time quantitative PCR (qPCR) was carried out using a Corbett Rotorgene and Sybr green mix (Bioline). Data ([IP - no antibody control]/input) were expressed as a percentage of the input and normalised to levels of H3 where appropriate. Error bars represent the standard error of three real-time PCR reactions. The primers used are listed in Table 3.

Primer name	Sequence
ACS1 fwd	GCGCAGCAGAAGAAGGAACA
ACS1 rev	CCCGTCCAAGTGTGGAGAAT
ADH1 1 fwd	CGCTCACATTCCTCAAGGTA
ADH1 1 rev	CCGGAGATAGCAACCCAGTG
AGA1 fwd	AAGCAGTTCC CAGCAAGCAA
AGA1 rev	GGATCAGATGCCAAGGCAAT
ATP17 fwd	CGGCACCAAATGCCAAGCGC
ATP17 rev	GGCTTTGTATCTGGCCAATC
CIT2 fwd	CCCATCCA TGCTCAAGAT GT
CIT2 rev	CCCATACGCTCCCTGGAATA
DLD3 fwd	CCCAAATTTCAAAGTTCTCG
DLD3 rev	CCAGTCCTGGTTAAACGAAG
ENO2 fwd	CCCAACCGTCGAAGTCGAAT
ENO2 rev	CCCATCCACTTGGATTTGTC
FMP27 1 fwd	TTTGACCTAGAAGCATCCCT
FMP27 1 rev	CCTCCTCCTCTTCAATGTAA
FMP27 2 fwd	GCCCTTCCTTTTATACTTCC
FMP27 2 rev	GATAAGGGGTCTCTATGACA
FMP27 3 fwd	TTCCCGTTCACTCTTGTTCT
FMP27 3 rev	TGGGAAGTATATTTCCCGTG
FMP27 4 fwd	TCTTGGTGGTGTTGCAGTTG
FMP27 4 rev	GCTGTTTGCTTTTACTGTCC
FMP27 5 fwd	CCGTCAGGCTAAAATCCGTT
FMP27 5 rev	CCTCTCTTGGTAATTCTTTA
FMP27 6 fwd	GCGAAACCAATTTGTTCTCT
FMP27 6 rev	GGATCCATGTCAATAGTTAA
FMP27 7 fwd	GGGTAACAATCTTATGGAAG
FMP27 7 rev	CCTTAATATTGTATGCTCGT
FMP27 8 fwd	GCCAAATAATCGCACTCCCA
FMP27 8 rev	GCGTTGGAGATTCTACCATC
FMP27 9 fwd	GGGATTTTATGATAACAGAG
FMP27 9 rev	GGGCACGAACAACGAGTAAT
FMP27 10 fwd	CCGTTACTGGTATGGTCTAG
FMP27 10 rev	CCTTGTCTGCTTTTTCGTTT

Primer name	Sequence
FMP27 11 fwd	CCCATGGCTGTACACTAGGG
FMP27 11 rev	GGCTGATTTGTTTTTTCTAA
FMP27 12 fwd	GCCCAAACTAGAAAATTCGT
FMP27 12 rev	CGTTGTGAATATTGGATAGA
FMP27 13 fwd	CGGAGCCTGAAGAACTTCGT
FMP27 13 rev	GGTCTGAAACATGGTAACA
GAL1 5′ fwd	CGCTTAT GATGCTAAAC CGG
GAL1 5′ rev	CGCAAAGCATATCAAAATCA
GAL1 3′ fwd	G GGTGGTTGTACTGTTCACT
GAL1 3′ rev	GCTCAGCATCAGTGATCTTA
HYP2 fwd	CCTTCTTTCTAACGATTCTA
HYP2 rev	GCGGAGGAACCAGCGTCAGC
MET16 5′ fwd	GCCAGCAAAGGTATCAACCC
MET16 5′ rev	GCGTTTCCAGCTTGATTAGT
MET16 3′ fwd	CCTTGGATATAGATCCATTG
MET16 3′ rev	GGCTTCATGAATTCCACACT
PAA1 fwd	GCCTCCTCAAGTAGCACGCT
PAA1 rev	CGAAAGCTGATAATTTCTTC
PDC1 fwd	GTTTGCCAGG TGACTTCAAC
PDC1 rev	ACCGAAGGTGGTGATGATAC
PGK1 fwd	GCGTGTCTTC ATCAGAGTTG
PGK1 rev	AGTGAGAAGCCAAGACAACG
RPL10 fwd	GAGCTGTTCC AGACTCCAAG
RPL10 rev	GTTGGCACAGATACGAGCAG
RPL39 fwd	CCCAGTTAGA TGCTGAACTG
RPL39 rev	GCGTAGTGGCGACTTTAGAT
RPL43B fwd	GGAGTTAGTGGAG ATGCGTG
RPL43B rev	GGACCTACAGTGAGATTGCC
RPS15 fwd	ACGACCGATC ATGTCTCAAG
RPS15 rev	TCTTCTAACTCTAGCTGGGG
SEN1 fwd	CCCAATAGCGATGTTCAGCT
SEN1 rev	CCCCTAATAATTTTGCCTCT
U4 fwd	ATCCTTATGCACGGGAAA
U4 rev	AACACAATCTCGGACGAATC

Table 3. Sequences for primers used in real-time PCR.

2.11. ChIP-sequencing

Chromatin was immunoprecipitated essentially as described in section 2.10 except, where necessary, multiple immunoprecipitation reactions were performed and then pooled to obtain sufficient material for sequencing (10 ng). Immunoprecipitated DNA was quantified using Qubit Fluorometric Quantitation (Invitrogen). Sequencing was performed by the Wellcome Trust Centre for Human Genetics, Oxford. DNA was multiplexed during library preparation and subjected to 50 nt paired end single lane sequencing. Initial analysis was performed by S. Murray using MATLAB (MathWorks). Reads were aligned to the *S. cerevisiae* genome using Bowtie with the command

line: Bowtie s_cerevisiae –n 2 reads.fastq > bowtie_output, which allows for up to two mismatches per read. Sample reads were normalised to the input read counts for each strain to account for any sequencing bias. To allow comparison of modifications between mutant and wildtype strains, signals in the mutant strains were first normalised to the wildtype read count at *FMP27* +150 to +250 bp and then the ratio of the modifications (mutant/wildtype) at this region of *FMP27*, as measured by qPCR of the DNA prior to sequencing.

2.12. Preparation of RNA and reverse transcription

Yeast cells were resuspended in 400 µl TES (100 mM Tris-Cl pH 7.5, 100 mM EDTA pH 8.0, 0.5 % SDS) followed by addition of 400 µl phenol:chloroform pH 4.7. After incubation at 65°C for 20 min and -80°C for 30 min, the samples were centrifuged (13 000 rpm, 15 min, 22°C). The upper aqueous phase was transferred to 10 mM sodium acetate/ethanol and incubated at -80°C for 30 min. Centrifugation (13 000 rpm, 15 min, 4°C) pelleted the RNA before resuspension in water. Following DNase-I treatment, reverse transcription was carried out at 50°C for 1 h with random primers using Superscript III (Invitrogen). Real-time PCR was performed as above (section 2.10).

2.13. Purification of histones from S. cerevisiae for mass spectrometry

Histones were prepared essentially as Edmondson et al. 1996. Yeast cells were grown in 300 ml YPD to OD_{600} 0.5-0.8. Cells were harvested (5000 rpm JA-10 rotor, 5 min, 4°C) and washed in sterile water. Cells were resuspended in 0.1 mM Tris-Cl pH 9.4, 10 mM DTT and incubated for 15 min at 30°C with gentle shaking. Following a wash with and then resuspension in 1.2 M sorbitol, 20 mM HEPES pH 7.4, cells were spheroplasted with 2.75 g zymolyase/g yeast cells for 60 min at 30°C with gentle shaking. A series of washes was then performed in: ice-cold 1.2 M sorbitol, 20 mM PIPES pH 6.8, 1 mM $MgCl_2$; ice-cold nuclear isolation buffer (0.25 M sucrose, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, $CaCl_2$, 15 mM MES pH 6.6, 0.8 % Triton X-100, protease inhibitors) twice with incubations on ice for 20 min; buffer A (10 mM Tris-Cl pH 8.0, 0.5 % NP-40, 75 mM NaCl, 30

mM sodium butyrate, protease inhibitors) twice with 15 min incubations on ice; and buffer B (10 mM Tris-Cl pH 8.0, 0.4 M NaCl, 30 mM sodium butyrate, protease inhibitors) twice with 10 min incubation on ice for the first wash. The resulting nuclei were resuspended in cold 0.4 N H_2SO_4 and incubated on ice for 30 min. After centrifugation (13 000 rpm, 10 min, 4°C), the supernatant (containing histones) was added to 5 volumes of acetone. Proteins were precipitated overnight at -20°C. Precipitated protein was collected by centrifugation (13 000 rpm, 10 min, 4°C) and air-dried before resuspension in 10 mM Tris-Cl pH 8.0. Samples were separated using 18 % SDS-polyacrylamide gel electrophoresis before excision of an appropriate molecular weight gel band and submission for mass spectrometry (Central Proteomics Facility, Dunn School of Pathology, Oxford).

2.14. Identification of histone H3 post-translational modifications by MS/MS

All the subsequent steps were performed by the Central Proteomics Facility. Briefly, the excised gel band was digested with trypsin overnight at 37°C. The resulting peptides were then analysed on an Ultimate 3000 nano HPLC system (Dionex, Camberley, UK) run in direct injection mode coupled to a LTQ XL Orbitrap mass spectrometer (Thermo Electron, Hemel Hempstead, UK). Samples were resolved on a 15 cm x 75 µm inner diameter PicoTip analytical column (New Objective, Woburn, MA, USA), which was packed in-house with Reprosil-Pur C18-AQ phase, 3 µm bead (Dr. Maisch, Germany). A 120 min gradient was used to separate the peptides. The mass spectrometer was operated in a “Top 5” data-dependent acquisition mode. Precursor scans were performed in the Orbitrap at a resolving power of 60,000, from which five precursor ions were selected and fragmented in the linear ion trap. Charge state +1 ions were rejected. Data analysis was performed using Mascot v. 2.2 against a custom database containing the NCBI *S. cerevisiae* protein database and the mutant (P16V) histone H3 protein sequence. The search was performed with fixed carbamidomethyl modification and variable oxidation, acetyl(K), methyl(K), dimethyl(K) and trimethyl(K) modifications. The peptide mass tolerance (± 20 ppm), the fragment mass

tolerance (± 0.5 Da) and the maximum number of missed cleavages (2) were defined in the Mascot search parameters.

2.15. Jhd2 peptide pull-down assay

Yeast cells were grown to OD_{600} 0.6-0.8 in 200 ml CSM-Trp (for plasmid retention). The cell pellet was resuspended in 1 ml binding buffer (40 mM HEPES-OH pH 7.4, 175 mM NaCl, 0.1 % TWEEN-20, 10 % glycerol, protease inhibitors) and cells were broken using acid-washed glass beads on a MagnaLyser (Roche). Cell extract was incubated with 100 µl pre-blocked peptide-conjugated sulfolink beads (Pierce) for 2 h rotating at 4°C. The beads were washed with binding buffer before boiling in protein loading buffer for 5 min. Extracts were analysed by Western blotting to detect the Myc-tagged Jhd2 protein.

2.16. Expression and purification of recombinant His_6-tagged Rph1 from E. coli

E. coli BL21(DE3) cells were transformed by heat shock with the plasmids carrying the wildtype or catalytically mutant (H235A) *RPH1* gene fused to the sequence encoding a His_6 tag (pET28 *RPH1*, pET28 *RPH1* H235A, R. Klose). Overnight pre-cultures were grown in 5 ml LB (1 % tryptone, 0.5 % yeast extract, 1 % NaCl) plus appropriate antibiotic before dilution in 500 ml LB plus antibiotic. Expression of the recombinant proteins was induced for 3 h by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD_{600} 0.5. Cells were harvested by centrifugation (6000 g, 20 min, 4°C) before resuspension in 12.5 ml lysis buffer (20 mM Tris-Cl pH 8.0, 500 mM NaCl, 0.1 % NP-40, protease inhibitors). Sonication at 60 % amplitude for six 30s on/30s off cycles lysed the cells. Extracts were cleared by centrifugation (12 000 rpm JA-20 rotor, 20 min, 4°C) before incubation with Ni-NTA resin rotating for 2 h at 4°C. The beads were washed with 20-40 column volumes of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole). Bound protein was eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole). Fractions with

protein were pooled and dialysed overnight against BC100 buffer (40 mM HEPES-OH pH 7.9, 100 mM KCl, 10 % glycerol, 0.5 mM PMSF, 0.5 mM DTT).

2.17. In vitro demethylase assay

Reactions were set up in demethylase buffer (50 mM HEPES-OH pH 7.9, 0.023 mM $Fe(NH_4)_2(SO_4)_2$, 1 mM α-ketoglutarate, 2 mM ascorbate) with 8 µg recombinantly expressed enzyme (section 2.16). Calf thymus histones (10 µg) or synthetic histone H3 methylated on lysine 18 (2 µg) were used as substrates and the demethylase inhibitor, EDTA, when included, was present at 10 mM. Reactions were incubated at 37°C for 3 h with occasional agitation before boiling in protein loading buffer and analysis by Western blotting (section 2.7).

Chapter 3.

Investigating the crosstalk between H3K14 and H3K4me3

3. Investigating the crosstalk between H3K14 and H3K4me3

3.1. Introduction

It is becoming increasingly clear that histone modifications do not act independently but participate in crosstalk, which can not only control their deposition and removal but also their effects on chromatin-based processes. Therefore, in order to understand more about how a particular modification is regulated and functions, it is important to discover more about how that modification interacts with other nearby or more distant modifications. One modification that attracts a lot of interest is methylation, and especially trimethylation, of lysine 4 on histone H3. H3K4me3, deposited by the Set1 methyltransferase complex (Briggs et al. 2001; Roguev et al. 2001; Nagy et al. 2002; Krogan et al. 2002) and removed by the Jhd2 demethylase (Ingvarsdottir et al. 2007; Liang et al. 2007; Seward et al. 2007; Tu, S et al. 2007), is found at the 5' end of active genes and has a high positive correlation with both sense transcription (S. Murray, unpub.) and transcript levels (Santos-Rosa et al. 2002; Bernstein et al. 2002; Liu et al. 2005; Pokholok et al. 2005). However, it is still unclear what function this modification has in chromatin. Although H3K4me3 has been shown to interact with a variety of protein effectors such as Yng1 (Martin et al. 2006b; Taverna et al. 2006), Sgf29 (Vermeulen et al. 2010; Bian et al. 2011) and Isw1 (Santos-Rosa et al. 2003) in yeast, and Chd1 (Sims et al. 2005), BPTF (Wysocka et al. 2006) and TFIID (Vermeulen et al. 2007) in humans, deletion of *SPP1*, the COMPASS subunit specifically required for the majority of H3K4me3, has little effect on steady-state transcript levels in yeast (Lenstra et al. 2011). This is not to say, however, that H3K4me3 has no function in the regulation of transcript levels: the field is becoming increasingly aware that H3K4 methylation may instead predominantly contribute to regulation during dynamic changes in transcription (Weiner et al. 2012). It is also possible that, instead of being a prerequisite for transcription, H3K4me3 may be deposited as a result of transcription, possibly influencing the stability of the resulting transcript. A mechanism for how this might occur has already been proposed: H3K4me3 can enhance the Nrd1-dependent

termination of short transcripts that are then targeted for degradation (or processing in the case of snoRNAs) by the exosome (Terzi et al. 2011). It is hoped that by understanding more about the regulation of H3K4me3 deposition and removal by crosstalk with other histone residues/modifications, insights can be made into the function of H3K4me3 in chromatin.

3.1.1. Crosstalk involving H3K4me3

As detailed in Chapter 1, there are a number of histone residues that can influence the levels of H3K4me3, including H3R2 (Kirmizis et al. 2007), H3K14 (Nakanishi et al. 2008) and H2BK123 (Dover et al. 2002; Sun and Allis 2002). It was decided to concentrate on the relationship between H3K14 and H3K4me3 since there have been no proposed mechanisms for this crosstalk. Nakanishi et al. (2008) generated a library of histone mutants in which each residue was systematically substituted with alanine. Whole cell extracts were made from each of these strains and probed for levels of H3K4me3 by Western blot to identify any histone residues that are important for optimal H3K4me3. In this screen, the authors identified a number of *cis*- and *trans*-acting residues, including several that had not previously been identified as being required for H3K4me3. Mutation of the residues flanking the methylation site, H3T3 and H3Q5, resulted in undetectable H3K4me2 and H3K4me3, but this could have been due to antibody epitope effects. The only other previously unidentified *cis*-acting residue to affect H3K4me3 was H3K14. This was interesting since the distance between these two residues on the histone H3 tail precludes any mechanism that relies on a direct interference with the binding of the Set1 catalytic site to its substrate residue: the Set1 homologue SET7/9 catalytic site only makes contact with five residues of the histone H3 tail (H3A1-H3Q5) (Xiao, B et al. 2003) and crystal structures of several SET domain-containing proteins reveal that a maximum of nine amino acids from the histone tail are accommodated in the SET domain catalytic site (Dillon et al. 2005). Most of the crosstalk mechanisms described on the histone H3 tail to date rely on a neighbouring residue directly interfering with enzyme/regulatory subunit binding, for example H3R2me2a prevents the binding

of Spp1 for H3K4me3 deposition (Kirmizis et al. 2007), H3S10ph enhances the histone H3-Gcn5 interaction to promote H3K14ac (Lo et al. 2000; Clements et al. 2003) and H3P38 isomerisation influences the ability of Set2 to methylate H3K36 (Nelson et al. 2006). Thus this relatively long-range *cis*-acting effect of H3K14 on H3K4me3 is novel.

Nakanishi et al. investigated the H3K14-H3K4me3 crosstalk further. They found that mutation of H3K14 specifically affected H3K4me3 but not H3K4me2. Furthermore, they could not detect H3K4me3 in the K14R and K14Q substitution strains. From this the authors concluded that this crosstalk was not due to simple charge-based effects and, since H3K14 can be acetylated and the acetylation is lost in all three H3K14 substitution mutants, proposed that H3K14ac is likely to be important for H3K4me3. However, they did not consider the possible functions of the unmodified lysine, which is also lost in the H3K14 point mutants. Furthermore, low levels of dimethylated H3K14 have been detected in yeast by mass spectrometry (Garcia et al. 2007), although no more is known about this putative modification. In favour of the proposed requirement of H3K14 acetylation for H3K4me3 is the co-localisation of these modifications at the 5' ends of genes (Liu et al. 2005; Pokholok et al. 2005) and on the same histone tails in yeast (Jiang et al. 2007) and mammals (Hazzalin and Mahadevan 2005). Other studies may support the Nakanishi et al. hypothesis, including the reduction in H3K4me3 in the absence of Gcn5, the histone acetyltransferase for a number of residues on histone H3, including H3K14 (Jiang et al. 2007; Govind et al. 2007). However, it is possible that the acetylation of a non-K14 histone H3 residue or the slow-growth phenotype of the *gcn5Δ* strain (Zhang et al. 1998) may be responsible for this phenotype.

Since strains lacking all H3K4 methylation or specifically H3K4me3 (*set1Δ* or Set1[829-1080] respectively), have decreased H3K14ac, H3K18ac and H3K23ac, as measured by mass spectrometry (Jiang et al. 2007), there is a possibility that the H3K14-H3K4me3 crosstalk may be

reciprocal. H3K4me3 specifically interacts directly with the PHD finger of Yng1 *in vitro* (Martin et al. 2006b; Taverna et al. 2006) and this interaction recruits and enhances the activity of the histone acetyltransferase Sas3 within the NuA3 complex towards H3K14-containing H3 peptides (Taverna et al. 2006). Furthermore, there is an overlap in the genomic distributions of H3K4me3 and Yng1 *in vivo* and mutation of the Yng1 PHD finger results in decreased H3K14ac and transcription at a subset of NuA3-targeted genes (Taverna et al. 2006). There is also a link between H3K4 methylation and the activity of the SAGA complex in yeast and humans. The tandem Tudor domains of the SAGA subunit Sgf29 specifically interact with peptides containing H3K4me2/3 but not unmodified H3K4 *in vitro* (Vermeulen et al. 2010; Bian et al. 2011). In humans, a ChIP-seq experiment demonstrated that Sgf29 and H3K4me3 co-localise genome-wide (Vermeulen et al. 2010) and in *S. cerevisiae,* deletion of *SGF29* results in decreased SAGA occupancy at promoters genome-wide, reduced H3K9/K14/K18 acetylation and decreased transcription of *GAL1* (Bian et al. 2011). A mutant of Sgf29 with defective H3K4me2/3 binding could not rescue the reduction in histone H3 acetylation caused by *SGF29* deletion, demonstrating the importance of the H3K4me2/3-Sgf29 interaction for SAGA-mediated acetylation of histone H3 (Bian et al. 2011). Thus there are a number of links between H3K14 and H3K4me3.

3.1.2. Studying the functions of histone residues using substitution mutations

H3K14 is acetylated by a number of HATs *in vivo,* including Gcn5 (Grant et al. 1999; Jiang et al. 2007), Sas3 (Howe et al. 2001; Martin et al. 2006a) and Elp3 (Kristjuhan et al. 2002; Winkler et al. 2002). Because these HATs also target other histone residues and combined deletion of some of these genes results in synthetic lethality (Howe et al. 2001), the function of H3K14ac cannot be isolated easily using simple genetics. Instead, a common method for studying the function of a histone modification *in vivo* is to substitute the residue and assess a phenotype. However, a frequent but incorrect assumption is that the phenotype of the mutant is due to the loss of the

modification, rather than the combined loss of both the modified and unmodified forms of the residue. One approach that can aid separation of the latter two possibilities involves using several different substitutions, with some substitutions roughly mimicking different wildtype states. In the study of lysine, and in particular acetylated lysine residues, three substituting residues are frequently employed: alanine, arginine and glutamine. Replacement of lysine with alanine removes both the positive charge of and the potential to modify the residue. In contrast, arginine maintains the positive charge of the lysine and can be used to mimic the unmodified lysine state. Glutamine has often been used to mimic acetyl-lysine due to its neutral charge and similar chemical structure, and can functionally substitute for acetyl-lysine *in vivo*. Two examples of this include the substitution of residues that can be acetylated in p53 with glutamine, which can recapitulate the effect of acetylation-mediated inhibition of p53 ubiquitination by the E3 ubiquitin ligase Mdm2 (Li et al. 2002). Secondly, the repression of genes caused by overexpression of the HDAC Rpd3 can be overcome by mutation of the substrate lysines in either the histone H3 or H4 tails to glutamine (De Nadal et al. 2004; Sabet et al. 2004). Two caveats associated with using glutamine to mimic acetylated lysines in histone tails is that *in vivo*, histone acetylation has a discrete localisation at particular gene regions (Liu et al. 2005; Pokholok et al. 2005) and is also highly dynamic (Hazzalin and Mahadevan 2005) whereas the glutamine substitution will mimic fixed acetyl-lysine and be present on every histone molecule throughout the genome. Therefore, it must be remembered that these mimics cannot substitute exactly for any of the natural states *in vivo* and so caution must be taken when drawing any conclusions based solely on experiments with substitution mutants.

3.1.3. Large scale changes in gene expression in yeast

As mentioned above, H3K4me3 does not appear to regulate steady-state transcription (Lenstra et al. 2011) but may instead function during dynamic changes in gene expression. There are several examples of both large- and small-scale changes in gene transcription in yeast, including periodic

variations during the cell cycle (Cho et al. 1998; Spellman et al. 1998) or yeast metabolic cycle (Tu et al. 2005) and specific or general transcriptional changes in response to a cellular stimulus.

The cycling of different transcripts during the yeast metabolic cycle (YMC) is of particular interest since it links the metabolic status of the cell to transcription. Metabolic cycling has been observed in a number of studies on synchronous and continuously slow-growing yeast cultures (Chance et al. 1964; Beck and von Meyenburg 1968; Ghosh et al. 1971; Mochan and Pye 1973). The maintenance of slow-growth conditions in a chemostat facilitates temporal resolution of the oscillations in the concentration of dissolved oxygen (dO_2), which are used as a measure of the phases of the YMC. These phases (oxidative, reductive/building and reductive/charging), are accompanied by specific transcriptional signatures with certain transcripts peaking during different phases (Figure 9) (Klevecz et al. 2004; Tu et al. 2005). Over half of all *S. cerevisiae* genes (~3552) exhibit cycling during the YMC, a fraction that is much greater than those periodically expressed during the cell division cycle (~500-800 genes (Cho et al. 1998; Spellman et al. 1998)). This phenomenon is not simply a result of the synchronous, carbon-limited growth conditions: cycling of transcripts has also been observed by RNA fluorescence *in situ* hybridisation (RNA FISH) in single cells from an asynchronously-growing yeast population that are slow-growing as a result of phosphate limitation (Silverman et al. 2010; Wyart et al. 2010). The changes in transcription during the YMC have functional cellular consequences since levels of protein reporters correlate with mRNA production, although there is some heterogeneity within the cycling population (Laxman et al. 2010; Romagnoli et al. 2011). Thus metabolic and concomitant transcriptional cycling enables coordination of cellular processes with the metabolic status of the cell. This is not a phenomenon that is unique to *S. cerevisiae*: metabolic and transcriptional cycling have also been observed in the fission yeast *Schizosaccharomyces pombe* and human epithelial and primary fibroblast cells (Slavov et al. 2012).

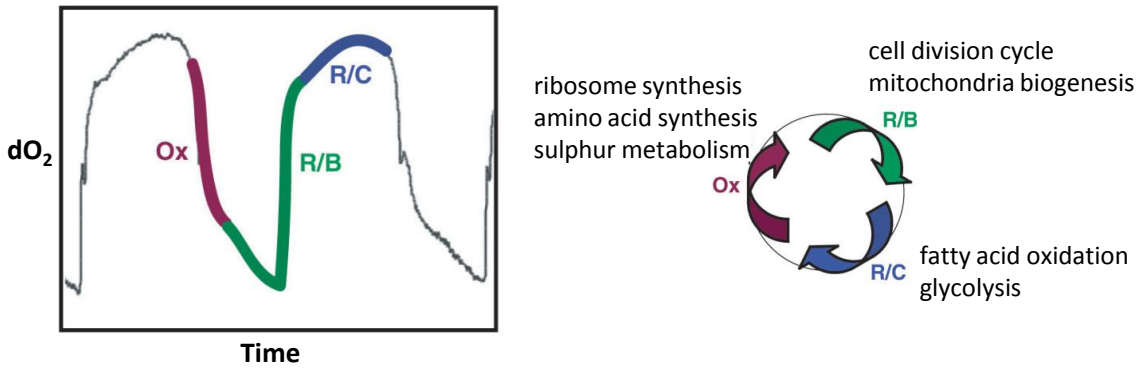


Figure 9. The yeast metabolic cycle.

The periodic phases of the yeast metabolic cycle, as monitored by the oscillations in dissolved oxygen (dO_2) concentration, and the associated cellular and metabolic processes. (Ox, oxidative; R/B, reductive/building; R/C, reductive/charging). Figure adapted from Tu et al. (2005).

The oxidative (OX) phase occurs following the peak in dO_2 levels and represents a burst of mitochondrial respiration to produce ATP. This is accompanied by the expression of genes involved in growth such as the energy-demanding processes of ribosome biogenesis and translation (Tu et al. 2005). After the concentration of dO_2 has dropped, genes expressed during the reductive/building (R/B) phase peak. These include nuclear-encoded mitochondrial genes, histone genes and genes required for DNA replication and cell division. It is thought that cell division might be restricted to this phase of the metabolic cycle to minimise oxidative DNA damage by the reactive oxygen species produced during mitochondrial respiration (Klevecz et al. 2004; Tu et al. 2005; Reinke and Gatfield 2006), although mitotic S phase has been shown to coincide with this period of high oxidative metabolism (Slavov and Botstein 2011). Moreover, the cell division cycle is not required for metabolic cycling (Slavov et al. 2011) and only ~50 % of the cell population enters the cell cycle in each R/B phase (Laxman et al. 2010). Therefore there does not seem to be as tight a link between cell division and the metabolic cycle as originally proposed. Following the reductive/building phase is the reductive/charging (R/C) phase, during which genes encoding proteins involved in non-respiratory modes of metabolism, protein degradation and the starvation/stress response peak in expression. This is proposed to be a quiescent-like state (Allen

et al. 2006; Shi et al. 2010) that enhances the ability of the cell to survive under low glucose conditions and prepares the cell for growth upon improvement of conditions (Laxman et al. 2010). For this reason, the genes expressed during the R/C phase negatively correlate with growth rate (Brauer et al. 2008) and cells in this phase are more resistant to stress (Lu et al. 2009; Slavov et al. 2012).

Of particular relevance to the work presented in this thesis is the differential regulation of genes expressed during the three YMC stages. As mentioned above, the YMC is accompanied by oscillations in various intracellular metabolites, including acetyl-CoA. Acetyl-CoA is synthesised and consumed by a number of metabolic pathways but its levels peak during the oxidative phase of the YMC (Tu, BP et al. 2007). Indeed, entry into oxidative phase can be prematurely induced by the addition of acetate, a precursor of acetyl-CoA, to the culture medium, indicating that acetyl-CoA is an important stimulant for growth (Cai et al. 2011). Interestingly, this peak of acetyl-CoA correlates with a peak in protein acetylation, including three components of the SAGA complex, Spt7, Ada3 and Sgf73. This acetylation is dependent on both Gcn5 and Ada3, indicating that it is mediated by SAGA itself. Gcn5 within SAGA acetylates a number of residues on histone H3 and analysis of histone acetylation levels showed that H3K9, K14, K18, K23, K27, K56 and H4K5, K8 and K12 acetylation are also highly periodic during the YMC (all these residues are acetylated upon entry into the oxidative phase except H3K56 and H4K16). Many of the genes regulated by Gcn5/SAGA are involved in growth, including ribosomal genes (peaking in oxidative phase), and the stress response genes (peaking in reductive/charging phase) (Huisinga and Pugh 2004). However, although SAGA is bound to these genes in their respective phases of the YMC, histone H3 acetylation is only increased at SAGA-bound genes expressed in the oxidative phase. Therefore, genes expressed during the reductive/charging phase are subjected to a different form of regulation by SAGA that is not dependent on histone H3 acetylation. This raises the possibility

that mechanisms of gene regulation other than SAGA-mediated histone acetylation could also be temporally employed.

Like histone acetylation, histone methylation could also be preferentially deposited in different phases of the YMC. The levels of the methyl donor, S-adenosyl-L-methionine (SAM) do not vary periodically during the YMC, however, oscillations in S-adenosylhomocysteine (SAH) have been observed, with a peak level in the R/B phase (Tu, BP et al. 2007). SAH is produced following the loss of the methyl group from SAM during the methylation of substrates, including proteins, and so the oscillations in SAH could reflect cycles of protein methylation. In support of this finding, the majority of yeast methyltransferases are periodically expressed during the YMC, including Set2 and Dot1 in the R/B phase (Tu et al. 2005; Wlodarski et al. 2011). However, although Set1 expression does not oscillate (Tu et al. 2005; Wlodarski et al. 2011) and the global levels of H3K4me3 remain constant during the YMC (Cai et al. 2011), it is possible that H3K4me3 may cycle on individual genes. It would be predicted that the potential peaking of H3K4me3 on individual genes would coincide with their period of maximum expression.

Another large-scale change in gene expression occurs in response to external stimuli. Over 50 % of the genome is involved in responses to various changes in the cellular environment. Of these, ~900 genes behave similarly over almost all environmental stresses tested (Gasch et al. 2000; Causton et al. 2001). Approximately two-thirds of these commonly responsive genes are repressed during the environmental stress response (ESR), including those participating in energy-intensive processes such as RNA metabolism, protein synthesis and molecular biosynthesis. It is thought that repression of genes involved in these processes allows the cell to divert the energy to other processes upregulated during the stress response (Jelinsky and Samson 1999). The genes that are induced during the ESR include those involved in carbohydrate and fatty acid metabolism, removal of reactive oxygen species, DNA damage repair, vacuolar and mitochondrial

functions, autophagy and intracellular signalling. The transcription of these genes is predominantly controlled by the transcription factors Msn2 and Msn4 (Martínez-Pastor et al. 1996; Schmitt and McEntee 1996). Interestingly, the genes that are induced and repressed in response to stress have a bias towards transcriptional regulation by SAGA and TFIID respectively (Huisinga and Pugh 2004), which has been proposed to balance the requirement for a dynamic, rapid induction of stress-inducible genes and the more steady transcriptional output of housekeeping genes. This again highlights the potential for differential transcriptional regulation of genes falling into separate functional classes, although it should be noted that transcript levels can also be regulated by an alteration in the rates of degradation during the stress response (Lindquist 1981; Albig and Decker 2001).

Although the changes in transcription profile during the yeast metabolic cycle and stress responses have been discussed separately, they are in fact strongly overlapping (Machné and Murray 2012; Slavov et al. 2012). Many of the genes expressed during the OX phase of the YMC are repressed during the ESR and many of those expressed during the reductive phases are induced in response to environmental stress. Therefore there are two distinct classes of genes that are oppositely regulated under different growth/metabolic conditions (Gasch et al. 2000; Tu et al. 2005; Brauer et al. 2008; Slavov et al. 2012). In the natural environment, the availability of nutrients is sporadic and yeast require mechanisms both to survive periods of low nutrition by entry into a quiescent-like state but also to rapidly initiate growth following improvements in conditions. Major regulators of the switch between these states are the TOR and protein kinase A (PKA) signalling cascades (reviewed by Loewith and Hall 2011; Cai and Tu 2012), which transduce nutrient signals to either induce the stress response pathways during poor nutrition or promote the ribosome biogenesis and cell growth required for subsequent cell division when nutrients are available (Jorgensen and Tyers 2004). In contrast, standard laboratory growth conditions favour continuous rapid exponential growth. Since the duration of the OX phase increases relative to the

reductive phases with growth rate whilst maintaining the same overall cycle period length (Slavov and Botstein 2011), more cells will be in the OX phase at any one time under fast-growth conditions and therefore, at a population level, OX phase genes are predominantly expressed. Conversely, the short duration of the reductive phases at high growth rates results in a low population-level expression of the genes that peak in this phase. Collectively, the studies described above inform us that, because these classes of genes are expressed under opposite conditions, it is probable that the regulatory mechanisms controlling their transcription are also different.

3.1.4. Stress genes are enriched for antisense transcripts

Stress response genes have both high levels of antisense transcription running into their promoters (Yassour et al. 2010; Xu et al. 2011) and higher levels of RNA polymerase II at their 3' ends relative to their 5' ends (Kim, T et al. 2010) during exponential growth when these genes are repressed. The antisense transcripts at these genes have been proposed to allow a rapid, switch-like induction in response to stress (Xu et al. 2011).

The Nrd1-Nab3-Sen1 complex is responsible for the termination of snoRNAs (Steinmetz et al. 2001), cryptic unstable transcripts (CUTs) (Arigo et al. 2006b; Thiebaut et al. 2006) and some pre-mRNAs including *NRD1*, *IMD2* and *URA2* (Arigo et al. 2006a; Kuehner and Brow 2008; Thiebaut et al. 2008). Transcription termination mediated by Nrd1-Nab3 directs transcripts to be polyadenylated by the TRAMP complex and then degraded by the nuclear exosome (LaCava et al. 2005; Vanácová et al. 2005; Vasiljeva and Buratowski 2006; Wyers et al. 2005). Nrd1 and Nab3 are required for repression of the growth genes *CLN3, TYE7* and *DLD3* upon glucose limitation, acting oppositely to PKA/Ras signalling (Darby et al. 2012). The authors suggest that this Nrd1-dependent repression may operate through transcription attenuation, although they could not rule out promoter occlusion by an upstream CUT or Nrd1-dependent degradation of the full-

length transcript. Similar to this described role of Nrd1 in the repression of genes in response to nutrient depletion, suppression of Nrd1 binding is required for the activation of the stress-inducible gene *FKS2* upon cell wall stress (Kim and Levin 2011). Under non-inducing conditions, the Nrd1-Nab3-Sen1 complex associates with the Paf1 complex on the RNA polymerase II CTD (Sheldon et al. 2005; Gudipati et al. 2008; Vasiljeva et al. 2008) to attenuate *FKS2* transcription by premature termination. Upon cell wall stress, the mitogen-activated protein kinase (MAPK) Mpk1 binds to Paf1 in a way that prevents the recruitment of the Nrd1 complex to *FKS2*. This therefore allows transcription elongation to proceed and the production of the full-length *FKS2* transcript. In their analysis of the Neil et al. (2009) dataset, Kim and Levin note that approximately 10 % of protein coding genes have short sense transcripts across their promoters and suggest that this method of transcription attenuation by Nrd1 may be relatively widespread.

Of relevance to this study, H3K4me3 has been shown to promote Nrd1-dependent termination for a number of snoRNAs and CUTs (Terzi et al. 2011). However, deletion of *SET1* has little effect on the abundance of improperly terminated snoRNAs (Sheldon et al. 2005) or the transcription within the body of genes/downstream of mRNA termination sites, as measured by NET-seq (Terzi et al. 2011). Furthermore, the effects of *SET1* deletion are less severe than those caused by mutation of the termination factors, indicating that H3K4 methylation may increase the efficiency of, rather than be required for, Nrd1-dependent termination.

3.1.5. Aims

The aim of this chapter is to further characterise the crosstalk between H3K14 and H3K4me3 and its potential effects on gene transcription. In particular, it will be investigated whether there are any gene-specific effects of this crosstalk and if so, whether these genes fall into any functionally related classes.

3.2. Results

3.2.1. H3K14 is required for optimal H3K4me3

A reduced version of the alanine scanning mutagenesis screen was performed to assess global levels of H3K4me3 in whole cell extracts from strains in which the histone H3 residues with characterised modifications had been replaced with alanine (Figure 10A). The absence of H3K4me3 in the K4A strain indicates that the antibody raised against H3K4me3 is specific for a modification on H3K4. The lack of detectable H3K4me3 in the R2A strain has been documented previously and is proposed to be due to a crosstalk between H3R2 and H3K4me3 rather than antibody epitope masking (Guccione et al. 2007; Kirmizis et al. 2007). The effect of H3K14 substitution on H3K4me3 observed by Nakanishi et al. was confirmed by the strong reduction in the H3K4me3 signal in the extract from the K14A strain. No other histone H3 substitution tested in this experiment affected the global levels of H3K4me3 as greatly as the three described above.

To investigate the H3K14-H3K4me3 crosstalk further, whole cell extracts were prepared from yeast strains in which H3K14 had been substituted with alanine, arginine and glutamine and analysed by Western blot for global levels of H3K4me2 and H3K4me3 (Figure 10B). (H3K4me1 was not assessed in this project due to the low quality of the commercial antibodies raised against this modification. Nakanishi et al. (2008) did not observe an effect of H3K14 mutation on H3K4me1.) The arginine substitution was used to mimic the positively-charged, unmodified lysine residue whereas glutamine has been used widely as a mimic of the neutral, acetylated lysine (Wang and Hayes 2008). As before and in agreement with Nakanishi et al., the K14A strain had almost undetectable levels of H3K4me3, with only a very slight reduction in H3K4me2. Substitution of H3K14 with arginine or glutamine similarly had very little/no effect on global H3K4me2 levels. However, the effect of H3K14 substitution on H3K4me3 was not uniform in the three substitution strains: both the K14R and K14Q strains had higher levels of H3K4me3 than the K14A strain, with

near-wildtype levels of H3K4me3 in the K14R strain. This result was obtained over 20 times by Western blot (Figure 62, Appendix). The different effects of the three H3K14 substitutions on H3K4me3 were validated by chromatin immunoprecipitation (ChIP) at the 5' end of *FMP27*, a constitutively transcribed gene (Figure 10C) and numerous other genes (data not shown). Together, the Western and ChIP data are contrary to the result published by Nakanishi et al., who failed to detect H3K4me3 in extracts from any of the three H3K14 substitution strains. It is unclear why these experiments failed to reproduce the published results. To ensure that the effects observed were real and not an artefact of particular strain/antibody conditions, H3K4me3 levels were assessed in yeast with the K14R substitution in a range of strain backgrounds obtained from different laboratories including RMY200 (Michael Grunstein), WZY63a (Vincent Géli) (Figure 10D) and the strains used in the Nakanishi et al. study (data not shown). The antibody raised against H3K4me3 (Millipore 05-745R) was specific for this modification by Western blot and ChIP, as demonstrated by the loss of signal in the *spp1Δ* (and *set1Δ*) strains (Figure 10E, F). Additionally, a different H3K4me3 antibody (Diagenode CS-003-100) was used to rule out antibody effects (data not shown). Under all these conditions, a reproducible near-wildtype level of H3K4me3 was observed in the K14R strains. It is possible that the differences between the current results and those of Nakanishi et al. arose from different growth conditions or threshold effects on a Western blot.

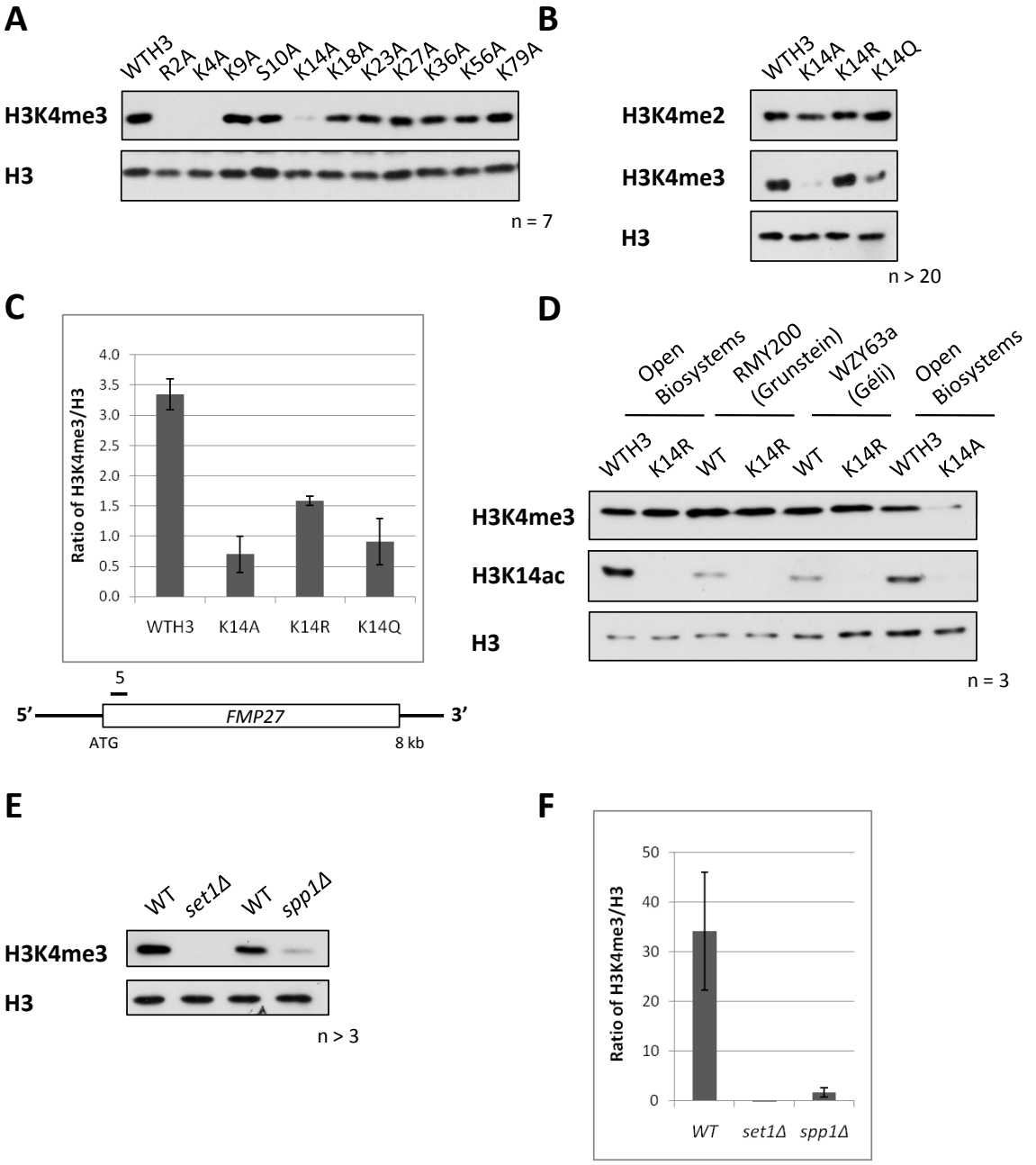


Figure 10. Mutation of H3K14 affects levels of H3K4me3.

(A) A Western blot showing levels of H3K4me3 in whole cell extracts from strains in which individual histone H3 residues had been mutated to alanine. Histone H3 levels are shown as a loading control. **(B)** A Western blot showing levels of H3K4me2 and H3K4me3 in three H3K14 substitution strains. Histone H3 levels are shown as a loading control. The replicates of this experiment are shown in Figure 62 (Appendix) to give an indication of its reproducibility. **(C)** Levels of H3K4me3 assessed by chromatin immunoprecipitation (ChIP)-qPCR in the 5' coding region of FMP27. Signals are normalised to levels of H3. Results shown are representative of at least three experiments. Error bars show standard errors of the real-time PCR reaction. **(D)** A Western blot showing levels of H3K4me3 in wildtype and K14R strains from three different yeast backgrounds. H3K14ac levels are included as a control for the H3K14 mutation. Histone H3 levels are shown as a loading control. **(E)-(F)** A Western blot and ChIP-qPCR showing the levels of H3K4me3 in the wildtype (BY4741), set1Δ and spp1Δ strains. H3K4me3 signals are normalised to levels of H3. qPCR was performed at the same primer position as (C).

The fact that there is a robust level of H3K4me3 in the K14R strain indicates that a lysine at position 14 is not absolutely required for this modification. Additionally, the result in this strain may argue against a requirement of H3K14ac for H3K4me3, as proposed by Nakanishi et al. Experimental evidence for a requirement of H3K14ac for H3K4me3 includes the reduction in global and gene-specific H3K4me3 in the *gcn5Δ* strain (Govind et al. 2007; Jiang et al. 2007). However, Gcn5 does not solely acetylate H3K14 but also a number of other lysine residues on histone H3 (Grant et al. 1999; Jiang et al. 2007). Furthermore, this strain has a significantly reduced growth rate (Zhang et al. 1998). Therefore it is possible that the reduction in H3K4me3 observed in the *gcn5Δ* strain is not a direct consequence of loss of H3K14ac. To test this, *GCN5* was deleted in the K14R strain and levels of H3K4me3 were assessed (Figure 11). Deletion of *GCN5* reduced the global level of H3K4me3 in the presence of a wildtype copy of histone H3, as reported previously (Govind et al. 2007; Jiang et al. 2007). However, deletion of *GCN5* in the K14R strain further reduced the level of H3K4me3 when compared to the wildtype and K14R strains with wildtype *GCN5*. Because there is already no H3K14ac in the K14R strain, the additional reduction in H3K4me3 upon deletion of *GCN5* cannot be due to loss of this modification. Therefore, Gcn5 must be able to modulate H3K4me3 independently of H3K14ac, possibly via acetylation of other H3 lysines, global changes in transcription or indirectly through the reduced growth rate of this strain.

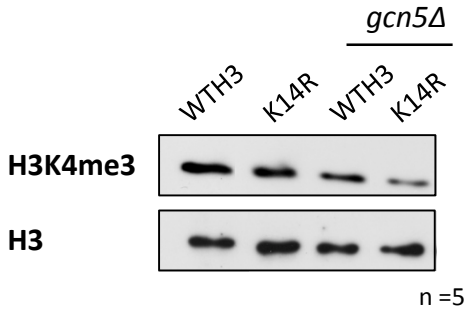


Figure 11. The requirement of *GCN5* for proper H3K4me3 is not fully dependent on H3K14.

A Western blot showing the levels of H3K4me3 in whole cell extracts from the indicated yeast strains. Histone H3 levels are shown as a loading control.

Further evidence that may suggest a lack of requirement of H3K14ac is that the substitution of H3K14 with glutamine, a proposed acetyl-lysine mimic, resulted in a dramatic loss of H3K4me3, implying that fixed wide-spread acetylation of H3K14 may act negatively towards H3K4me3. However, another interpretation of this result is that, in this case, glutamine may not be able to recapitulate the functions of acetylated lysine. For instance, if H3K14ac is required for the binding of a protein effector, the glutamine substitution, whilst neutralising the positive charge of the residue, may not be sufficient for effector binding. Together, the varying levels of H3K4me3 in the three H3K14 substitution strains points towards a more modulatory role rather than a strict requirement for the residue at that position in the H3 tail in the regulation of H3K4me3.

3.2.2. Investigating the H3K4me3-H3K14ac crosstalk

As detailed in the introduction to this chapter, this crosstalk has been proposed to be reciprocal: H3K4me3 has been suggested to promote H3K14ac via the recruitment of HAT complexes (Taverna et al. 2006; Martin et al. 2006b; Jiang et al. 2007; Vermeulen et al. 2010; Bian et al. 2011). Preliminary experiments presented in this work do not confirm these findings. Two strains, R2A and K4A, which lack H3K4me3 and all H3K4 methylation respectively, did not have altered global levels of H3K14ac (Figure 12). This was also the result with the *spp1Δ* strain, which had a near-complete loss of H3K4me3 but no reduction in H3K14ac (D. Clynes, data not shown). H3K4 methylation has also been shown to recruit the histone deacetylase complexes Set3 (Kim and Buratowski 2009) and Rpd3L (Wang et al. 2011). Conceivably, the combined removal of the potential to recruit HATs and HDACs via H3K4 methylation could result in no overall change in global H3K14ac upon substitution of H3K4/inhibition of H3K4me3. It is also possible that the effect of H3K4me3 loss on H3K14ac may only be detected on individual genes rather than at a global level. More experiments are required to validate the findings presented here, especially in light of the collective evidence for a link between these two modifications in the literature.

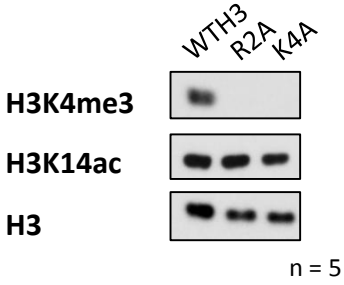


Figure 12. H3K4me3 is not required for H3K14ac.

A Western blot showing the levels of H3K14ac in two histone H3 mutant strains that lack H3K4me3. Histone H3 levels are shown as a loading control.

3.2.3. H3K14 is also required for H3K18ac

In order to discover other modifications that may be regulated by H3K14 or involved in the crosstalk between H3K14 and H3K4me3, the levels of a range of histone H3 modifications were assessed by Western blot in whole cell extracts from the three H3K14 substitution strains (Figure 13A). H3S10ph, H3K27ac and H3K36me2 remain at wildtype levels in the H3K14 mutants. H3K9ac is very slightly decreased equally in the three mutants whereas H3K36me3 is decreased only in the K14Q substitution strain. This latter finding is in accordance with Psathas et al. (2009), who observed decreased H3K36me3 in a strain in which six H3 tail lysines (H3K4, 9, 14, 18, 23, 27) had been mutated to glutamine but not when mutated to arginine. The most striking result however is the lack of detectable H3K18ac in all three H3K14 substitution strains. This is not due to an inability of the antibody raised against H3K18ac to recognise its epitope when H3K14 is mutated (Figure 13B). Mutation of no other histone H3 tail residue with a characterised post-translational modification has as dramatic an effect on H3K18ac as mutation of H3K14 (Figure 13C).

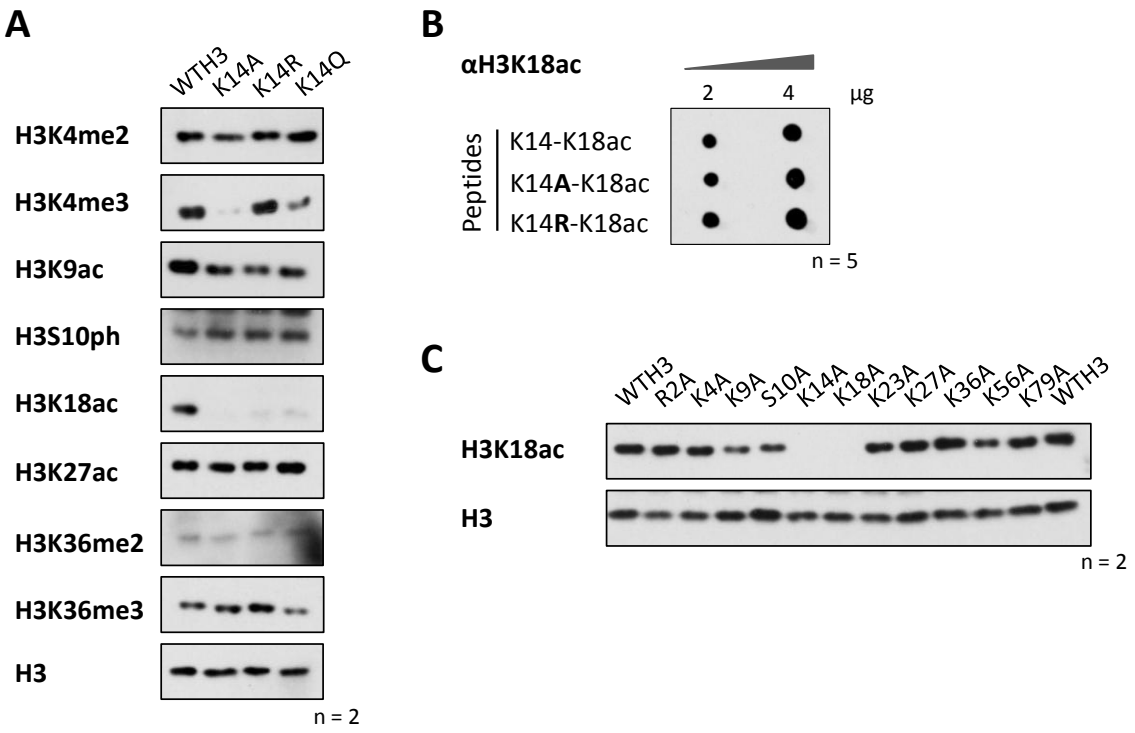


Figure 13. Mutation of H3K14 affects levels of H3K18ac.

(A) Western blots showing the levels of the indicated histone H3 post-translational modifications in the three H3K14 substitution strains. Histone H3 levels are shown as a loading control. **(B)** A dot blot with 2 and 4 µg of acetylated H3 peptides demonstrating that substitution of H3K14 with alanine or arginine does not disrupt recognition of H3K18ac by its antibody. **(C)** A Western blot showing levels of H3K18ac in whole cell extracts from strains in which individual histone H3 residues had been mutated to alanine. Histone H3 levels are shown as a loading control.

3.2.4. Gene-specific effects of H3K14 substitution

The next question to be asked was whether substitution of H3K14 affected H3K4me3 to the same extent at all genes. To this end, a ChIP experiment was performed with the K14A, K14R and K14Q strains and the levels of H3K4me3 at the 5' ends of a number of genes were assessed by real-time quantitative PCR (qPCR). Figure 14 shows three example genes selected for illustrative purposes. As is clearly evident, the effect of each H3K14 substitution varied greatly between individual genes, however, the relative levels of H3K4me3 in the three substitution strains were maintained at each gene. This result therefore suggests that the higher levels of H3K4me3 observed globally in the K14R relative to the K14A/K14Q strains result from a proportional increase in H3K4me3 at each gene in the K14R strain whilst maintaining the relative levels of H3K4me3 between genes.

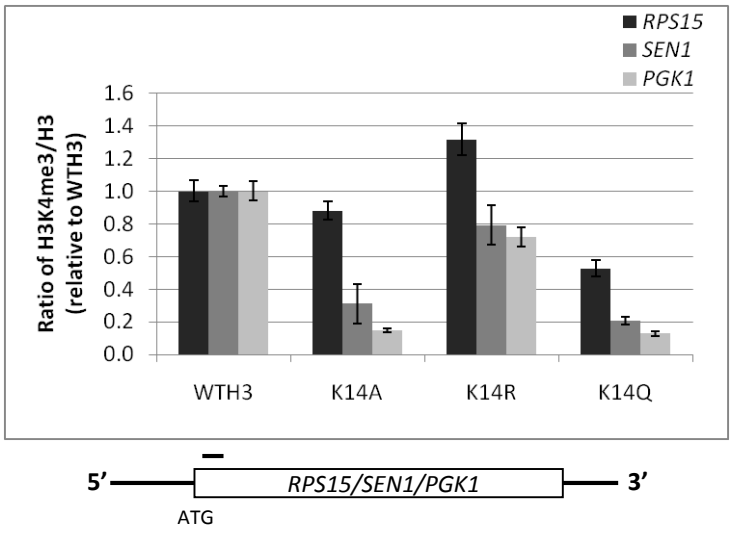


Figure 14. Mutation of H3K14 affects H3K4me3 to different extents at individual genes.

A ChIP-qPCR experiment showing the levels of H3K4me3 at the 5' end of three genes (*RPS15, SEN1* and *PGK1*) in the three H3K14 substitution strains relative to the WTH3 strain. Primer positions are indicated on the locus map. Signals are normalised to levels of H3. Error bars show standard errors of the real-time PCR reaction.

To gain more information about the variations in the effect of the H3K14-H3K4me3 crosstalk between genes, a ChIP-sequencing (ChIP-seq) experiment was performed with the K14A strain. Immunoprecipitated DNA was prepared as described in the Materials and Methods and was sent for high-throughput sequencing at the Wellcome Trust Centre for Human Genetics, Oxford. Immunoprecipitated samples were normalised to the input samples for each strain to remove any sequencing bias. To validate the data, the average WTH3 H3K4me3 gene distribution was compared to the wildtype H3K4me3 distribution obtained in a ChIP-on-chip experiment by Kirmizis et al. (2007). As can be seen from Figure 15A, the two datasets are closely overlapping and have a Spearman rank correlation of 0.83. In order to be able to compare between strains, the K14A sample datasets were first normalised to the read count of the corresponding WTH3 sample dataset and then the ratio of H3K4me3, as measured by qPCR on the immunoprecipitated DNA prior to sequencing, both at *FMP27.* Genes were ranked (from largest to smallest) according to the ratio of the average H3K4me3 over the transcription unit in the K14A strain relative to the WTH3 strain. These ratios were normally distributed with an average of 0.238. The top- and

bottom-ranked 200 genes were used for further analysis (Figure 15B, C) and are listed in Table 25 and Table 26 (Appendix). For the rest of this thesis, the genes with the 200 largest ratios of H3K4me3 in the K14A relative to the wildtype strain will be referred to as the 'top' class of genes whereas the genes with the 200 smallest ratios of H3K4me3 will be referred to as the 'bottom' class of genes.

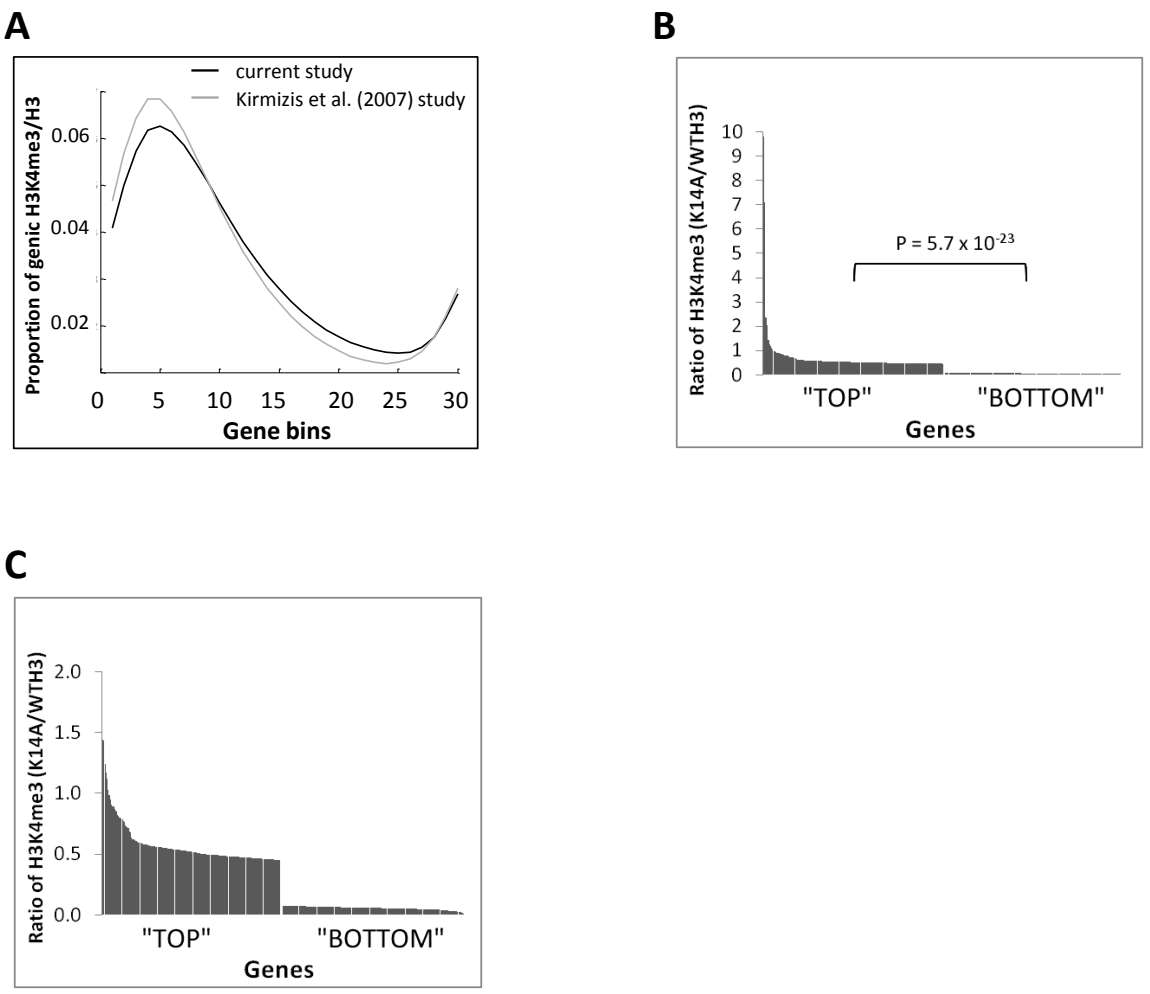


Figure 15. A ChIP-sequencing experiment demonstrates non-uniform effects of H3K14 mutation on H3K4me3.

(A) The average wildtype distribution profile of H3K4me3 normalised to histone H3 levels across all genes in *S. cerevisiae* from the present (black) and Kirmizis et al. (2007) (grey) datasets. Genes were divided into 30 bins and the level of H3K4me3 in each bin is expressed as a proportion of the total H3K4me3 across the gene (S. Murray). **(B)** The ratios of H3K4me3 in the K14A relative to the WTH3 strain at the 'top' and 'bottom' 200 genes after ranking according to the ratio of the average H3K4me3 across the transcription unit (K14A/WTH3) as measured by ChIP-sequencing. Signals are not normalised to levels of H3. P values are derived from student's t-test analysis. **(C)** As (B) after removal of the genes with the four highest H3K4me3 ratios to facilitate visualisation of the ratios of the 'bottom' class of genes.

3.2.5. H3K14 differentially regulates the levels of H3K4me3 on functionally distinct classes of genes

The next aim was to identify any common features between the genes in the two oppositely-defined classes of genes with K14A-affected H3K4me3. Firstly, a gene ontology (GO) analysis was performed. As can be seen from Table 4 and Table 5, the genes in these two classes function in very different biological processes. Genes in the 'top' class, with the highest ratios of H3K4me3 in the K14A relative to the WTH3 strain, are predominantly involved in translation, ribosome biogenesis and rRNA processing. In contrast, genes in the 'bottom' class, with the lowest ratios of H3K4me3 in the K14A relative to the WTH3 strain, are involved in the response to environmental stresses, thiamine and sulphur metabolic processes, and carbohydrate transport and catabolism.

Biological process	**Count**	**%**	**P value**
translation	65	33.9	5.90E-32
ribosome biogenesis	38	19.8	2.90E-18
regulation of translation	22	11.5	3.50E-11
posttranscriptional regulation of gene expression	22	11.5	1.20E-10
regulation of cellular protein metabolic process	22	11.5	2.90E-10
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	15	7.8	6.80E-10
ribosome assembly	14	7.3	6.90E-10
ribosomal subunit assembly	12	6.2	6.50E-09
rRNA processing	20	10.4	2.10E-07
ribonucleoprotein complex assembly	14	7.3	4.90E-07
rRNA export from nucleus	9	4.7	1.50E-06
ncRNA processing	21	10.9	9.00E-06
ribosomal small subunit biogenesis	9	4.7	9.10E-06
ribosomal small subunit assembly	6	3.1	1.10E-05
ncRNA metabolic process	21	10.9	9.20E-05
RNA export from nucleus	10	5.2	9.60E-05
RNA transport	10	5.2	4.00E-04
establishment of RNA localisation	10	5.2	4.00E-04
RNA processing	23	12	5.00E-04
ribosomal large subunit biogenesis	8	4.2	5.30E-04
nuclear export	10	5.2	8.40E-04
ribosomal large subunit assembly	6	3.1	9.20E-04
endonucleolytic cleavage to generate mature 3'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	3	1.6	2.40E-03
cellular macromolecular complex assembly	15	7.8	3.50E-03
macromolecular complex assembly	17	8.9	6.20E-03
endonucleolytic cleavages during rRNA processing	5	2.6	8.40E-03
cellular macromolecular complex subunit organisation	16	8.3	2.60E-02
regulation of translational fidelity	3	1.6	2.80E-02
macromolecular complex subunit organisation	18	9.4	3.10E-02
cleavages during rRNA processing	5	2.6	3.10E-02
endonucleolytic cleavage in *ITS1* to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript	4	2.1	4.10E-02

Table 4. Gene ontology terms associated with the genes with the top 200 H3K4me3 ratios (K14A/WTH3).

Shown are the significantly enriched (p<0.05) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the top 200 ratios of H3K4me3 (K14A/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

Biological process	Count	%	P value
response to temperature stimulus	17	8.5	1.40E-05
thiamine and derivative biosynthetic process	6	3	6.90E-05
thiamine and derivative metabolic process	6	3	1.10E-04
response to abiotic stimulus	20	10.1	1.20E-04
cellular response to heat	14	7	1.30E-04
oxidation reduction	20	10.1	1.70E-04
sulphur compound biosynthetic process	8	4	7.40E-04
sulphur metabolic process	9	4.5	1.40E-03
aromatic compound biosynthetic process	6	3	1.60E-03
water-soluble vitamin biosynthetic process	7	3.5	1.80E-03
response to toxin	6	3	3.00E-03
vitamin metabolic process	7	3.5	3.90E-03
carbohydrate transport	5	2.5	1.30E-02
cellular amide metabolic process	6	3	1.30E-02
polysaccharide catabolic process	3	1.5	1.90E-02
vitamin B6 biosynthetic process	3	1.5	1.90E-02
nitrogen compound biosynthetic process	14	7	1.90E-02
heterocycle catabolic process	4	2	2.40E-02
heterocycle biosynthetic process	6	3	3.40E-02
carbohydrate catabolic process	6	3	3.50E-02
beta-alanine biosynthetic process	2	1	4.30E-02

Table 5. Gene ontology terms associated with the genes with the bottom 200 H3K4me3 ratios (K14A/WTH3).

Shown are the significantly enriched (p<0.05) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the bottom 200 ratios of H3K4me3 (K14A/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

The distinct functions of the genes in the two K14A-affected classes hinted that these genes might be expressed under different cellular conditions. As mentioned in the introduction, the temporal expression of over half of the yeast genome is controlled by the yeast metabolic cycle (Tu et al. 2005). To assess whether the genes in the two classes were regulated by the YMC, the overlaps between the genes in the 'top' and 'bottom' classes and the genes whose transcripts peak in the oxidative (OX), reductive/building (R/B) and reductive/charging (R/C) phases of the YMC were found (Table 6). There is a significant bias for genes with the top 200 ratios of H3K4me3 to be expressed during the oxidative phase and a significant under-enrichment of genes in this class to

be expressed during the reductive phases when compared to the rest of the genome. Conversely, the class with the bottom 200 ratios of H3K4me3 is significantly enriched in genes that are expressed in the reductive/charging phase and significantly depleted for genes expressed in the oxidative phase.

	Expression phase in the yeast metabolic cycle		
	Oxidative	**Reductive/building**	**Reductive/charging**
All genes in the Tu et al. study (n=6575)	1017 (15.5 %)	975 (14.8 %)	1509 (23.0 %)
Genes with top 200 ratios of H3K4me3 (K14A/WTH3)	**48 (24 %)** $P = 3 \times 10^{-4}$	5 (2.5 %) $P < 1 \times 10^{-5}$	19 (9.5 %) $P < 1 \times 10^{-5}$
Genes with bottom 200 ratios of H3K4me3 (K14A/WTH3)	21 (10.5 %) $P = 1.9 \times 10^{-2}$	24 (12.0 %)	**59 (29.5 %)** $P = 4.0 \times 10^{-3}$

Table 6. Overlap between the genes with the top and bottom 200 ratios of H3K4me3 (K14A/WTH3) and the genes temporally expressed during the yeast metabolic cycle.

Yeast metabolic cycle expression data is taken from Tu et al. (2005). The percentages of genes in the 'top' and 'bottom' classes of genes overlapping with each phase are shown in brackets. Bold percentages signify enrichment of genes in that class for a particular feature compared to the rest of the genome. P values were calculated from a simulated probability distribution curve. The absence of P values denotes non-significant results ($P > 5 \times 10^{-2}$).

The connection between the yeast metabolic cycle and the environmental stress response (ESR) has already been documented (Slavov et al. 2012). Based on this link and the opposite effects of H3K14 mutation on the levels of H3K4me3 at the ribosomal and stress-response genes, a prediction would be that the genes in the two K14A-affected classes would also be differentially regulated during the ESR. This was indeed the case, with a significant bias for the genes with the 200 highest ratios of H3K4me3 (K14A/WTH3) to be repressed during the ESR whereas those with the 200 lowest ratios were enriched in genes that are induced during the ESR when compared to the rest of the genome (Table 7). This differential expression of genes in the 'top' and 'bottom' classes during the YMC and ESR is interesting because it implies that the genes in these two classes are unlikely to be transcribed at the same time in any one cell and so are likely to be subjected to different regulatory mechanisms that reflect the metabolic status of the cell.

	Response to environmental stress	
	Induced	**Repressed**
All genes in the Gasch et al. study (n=6152)	278 (4.5 %)	605 (9.8 %)
Genes with top 200 ratios of H3K4me3 (K14A/WTH3)	0 (0.0 %) $P < 1 \times 10^{-5}$	**82 (41 %)** $P < 1 \times 10^{-5}$
Genes with bottom 200 ratios of H3K4me3 (K14A/WTH3)	**20 (10.0 %)** $P = 6.0 \times 10^{-5}$	0 (0.0 %) $P < 1 \times 10^{-5}$

Table 7. Overlap between the genes with the top and bottom 200 ratios of H3K4me3 (K14A/WTH3) and the genes induced and repressed during the common environmental stress response.

Environmental stress response expression data is taken from Gasch et al. (2000). The percentages of genes in the 'top' and 'bottom' classes of genes overlapping with each phase are shown in brackets. Bold percentages signify enrichment of genes in that class for a particular feature compared to the rest of the genome. P values were calculated from a simulated probability distribution curve. The absence of P values denotes non-significant results ($P > 5 \times 10^{-2}$).

As mentioned in the introduction to this chapter, there is a bias towards SAGA-mediated transcriptional regulation of the genes induced during the ESR and TFIID-mediated regulation of those repressed during the ESR (Huisinga and Pugh 2004). To determine if this distinction could be applied to the differential effects of H3K14 on H3K4me3, the genes in the 'top' and 'bottom' classes were stratified according to whether their transcriptional regulation is dominated by SAGA or TFIID (Table 8). There is a significant bias for genes in the 'bottom' class, 10 % of which are upregulated during the ESR, to be predominantly regulated by SAGA, either alone or in concert with TFIID. In contrast, the 'top' class is significantly depleted in genes regulated by SAGA. SAGA-regulated genes tend to have a more dynamic expression pattern and bind a higher variety of gene regulators than TFIID-regulated genes (Huisinga and Pugh 2004; Venters et al. 2011). Thus, the gene ontology terms, the overlap with genes induced and repressed during the ESR and the predominant modes of regulation by SAGA/TFIID of the genes with H3K4me3 differentially affected by mutation of H3K14 suggest that H3K14 may be involved in the stress response. To explore this hypothesis, a preliminary experiment was performed in which the wildtype and H3K14 substitution strains were subjected to a variety of different environmental stresses and assayed for viability by drop plate analysis. However, none of the tested stresses (heat shock, cold

shock, hydrogen peroxide, cycloheximide, caffeine, hydroxyurea and rapamycin) resulted in differential effects between the wildtype and mutant strains. Further experimentation is needed to optimise these assays to ensure the correct severity of stress is applied and the subsequent viability assay is sufficiently sensitive to detect small changes in stress resistance.

	Transcriptional regulation		
	SAGA	**TFIID**	**SAGA and TFIID**
All genes in the Huisinga and Pugh study (n=6226)	577 (9.3 %)	5130 (82.4 %)	165 (2.7 %)
Genes with top 200 ratios of H3K4me3 (K14A/WTH3)	4 (2.0 %) $P < 1 \times 10^{-5}$	135 (67.5 %) $P = 1.2 \times 10^{-3}$	5 (2.5 %)
Genes with bottom 200 ratios of H3K4me3 (K14A/WTH3)	**28 (14.0 %)** $P = 2.1 \times 10^{-3}$	98 (49.0 %) $P < 1 \times 10^{-5}$	**22 (11.0 %)** $P < 1 \times 10^{-5}$

Table 8. Overlap between the genes with the top and bottom 200 ratios of H3K4me3 (K14A/WTH3) and the genes predominantly regulated by SAGA and/or TFIID.

SAGA/TFIID data taken from Huisinga and Pugh (2004). The percentages of genes in the 'top' and 'bottom' classes of genes overlapping with each phase are shown in brackets. Bold percentages signify enrichment of genes in that class for a particular feature compared to the rest of the genome. P values were calculated from a simulated probability distribution curve. The absence of P values denotes non-significant results ($P > 5 \times 10^{-2}$).

3.2.6. The differential changes in H3K4me3 are not a result of a slow growth rate in the H3K14 mutant strains

The field is becoming increasingly aware that many of the changes in transcripts observed in mutant strains may not be a direct consequence of the mutation but instead an indirect effect of the mutation on growth. Indeed, many gene deletions cause slow growth and result in a transcriptional signature that correlates strongly with that observed in the common environmental stress response (Gasch et al. 2000; Regenberg et al. 2006; F. Holstege, pers. comm.). To control for these effects in the H3K14 mutant strains, which could be accompanied by the differential changes in H3K4me3 that were observed at genes oppositely regulated during the environmental stress response, the growth rates of the mutant strains were assessed both by drop plate assays and by growth in liquid culture (Figure 16). Since the growth rates of the H3K14

mutants were not significantly different from the wildtype growth rate, this indirect effect of H3K14 on H3K4me3 can be discounted.

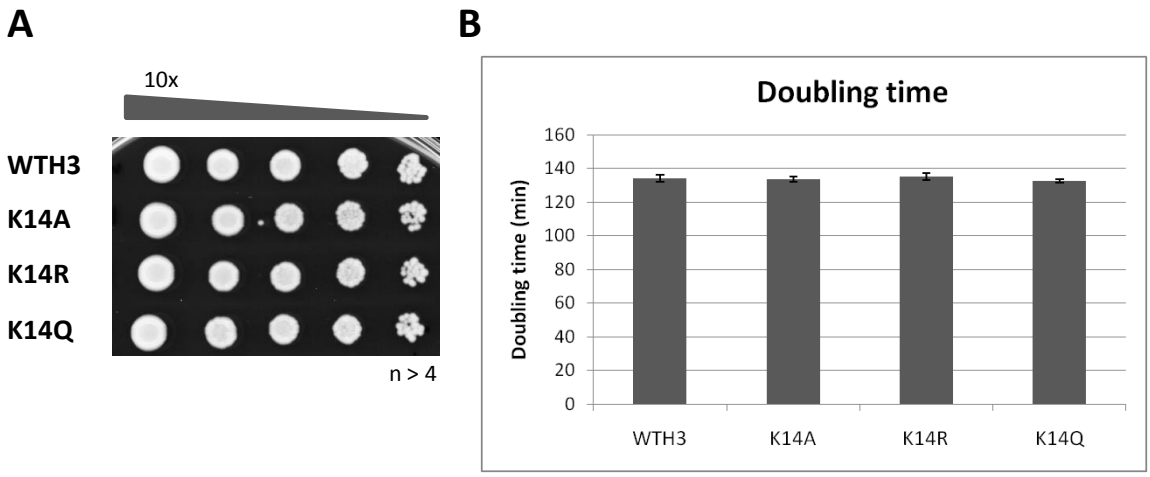


Figure 16. The H3K14 strains have wildtype growth rates.

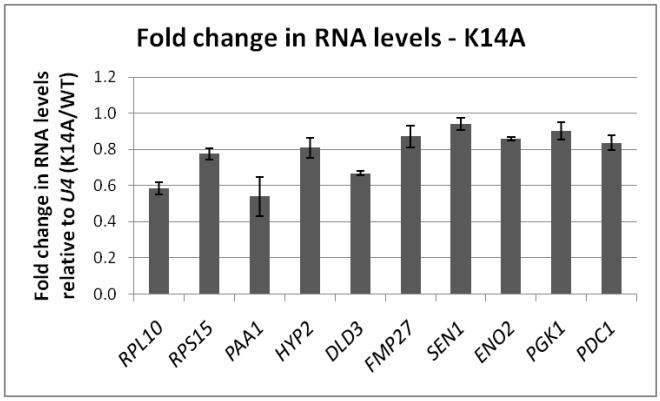
(A) A drop plate assay with the wildtype and H3K14 mutant strains. Strains were plated with ten-fold serial dilutions on YPD-agar plates and incubated for 2 days at 30°C. **(B)** The doubling times of the wildtype and H3K14 mutant strains grown in CSM, as measured by continual monitoring of liquid culture density during a 20 h period (M. Youdell). Error bars display 95 % confidence intervals.

3.2.7. H3K14 and transcript levels

The correlation between H3K4me3 and transcript levels has been established for a decade (Santos-Rosa et al. 2002; Bernstein et al. 2002; Liu et al. 2005; Pokholok et al. 2005), although a definite causality has yet to be confirmed. Therefore it was plausible that H3K14, via its effects on H3K4me3 (and H3K18ac), could influence transcript levels. To explore this possibility, cDNA was prepared from the wildtype and H3K14 mutant strains and the steady-state levels of individual transcripts during exponential growth were assessed and ordered according to the ratio of H3K4me3 in the K14A relative to the wildtype strain (Figure 17). The levels of transcript and H3K4me3 for three example genes are shown in Figure 18 for easier comparison. At all genes analysed, there was less cDNA in the H3K14 mutant strains than the wildtype. Surprisingly, the K14R strain, with near-wildtype levels of H3K4me3, was the H3K14 substitution strain that showed the largest decreases in the levels of all tested transcripts compared to wildtype. This suggests that this H3K14 mutant phenotype is mostly H3K4me3-independent and could instead

be attributed to the lack of H3K14ac and/or undetectable levels of H3K18ac in the absence of H3K14, both modifications that correlate positively with transcript levels (Liu et al. 2005; Pokholok et al. 2005). In agreement with this idea, substitution of H3K14 with the acetyl-lysine mimic glutamine resulted in smaller decreases in transcript levels than for the other H3K14 substitutions, indicating that the loss of H3K14ac may contribute to the reduction in transcript levels in the H3K14 substitution strains. In the K14A and K14Q strains, there was a very slight increase in transcript levels for the tested genes with the lowest ratios of H3K4me3 (K14A/WTH3) but transcript amounts seemed independent to the relative H3K4me3 levels on genes in the K14R strain. This slight anti-correlation between H3K4me3 and transcript levels was unexpected given the genome-wide positive correlation between these features. However, it should be noted that the reverse transcription performed in this experiment was not strand-specific and so this result could instead reflect an increase in the level of antisense transcript rather than the supposed increase in sense transcript at the genes with the lowest ratios of H3K4me3 in the K14A strain. Furthermore, if any reciprocal changes in sense and antisense RNA levels occurred at an individual gene, no net change in total RNA level would be observed. Future experiments employing strand-specific Northern blotting or microarrays could enable specific monitoring of the level of sense transcripts in the H3K14 mutant strains.

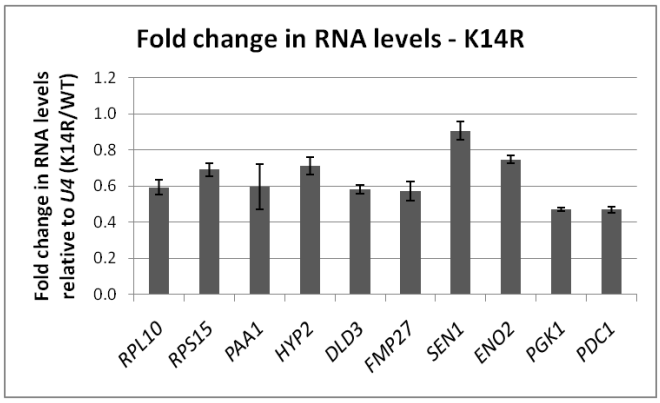
A



D

Gene	Rank of H3K4me3 ratio (K14A/WTH3)
RPL10	65
RPS15	155
PAA1	162
HYP2	280
DLD3	595
FMP27	2749
SEN1	2911
ENO2	4500
PGK1	4855
PDC1	6057

B



C

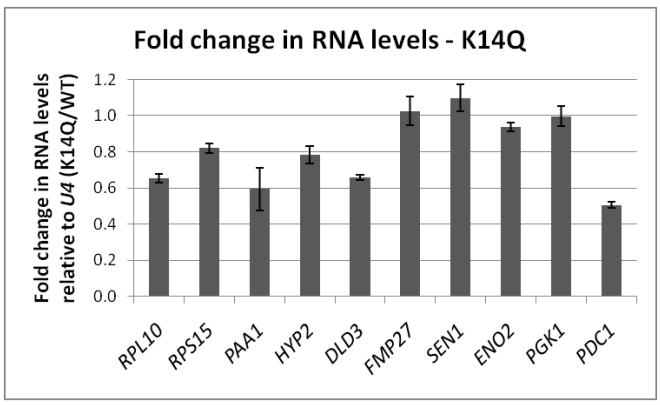


Figure 17. Mutation of H3K14 affects transcript levels to different extents at individual genes.

(A) qPCR after reverse transcription of RNA extracted from WTH3 and K14A yeast strains. K14A signals are normalised to levels of *U4* RNA and expressed as a ratio of the WTH3 normalised RNA level. Primers are in the 5' gene regions and are listed in the Materials and Methods. Error bars show standard errors of the real-time PCR reaction. **(B)** As (A) except with the K14R strain. **(C)** As (A) except with the K14Q strain. **(D)** The rank of the ratio of the average H3K4me3 across the transcription unit (K14A/WTH3) from the ChIP-seq experiment for each gene transcript assessed by reverse transcription-qPCR. Genes were ordered from largest to smallest H3K4me3 ratios.

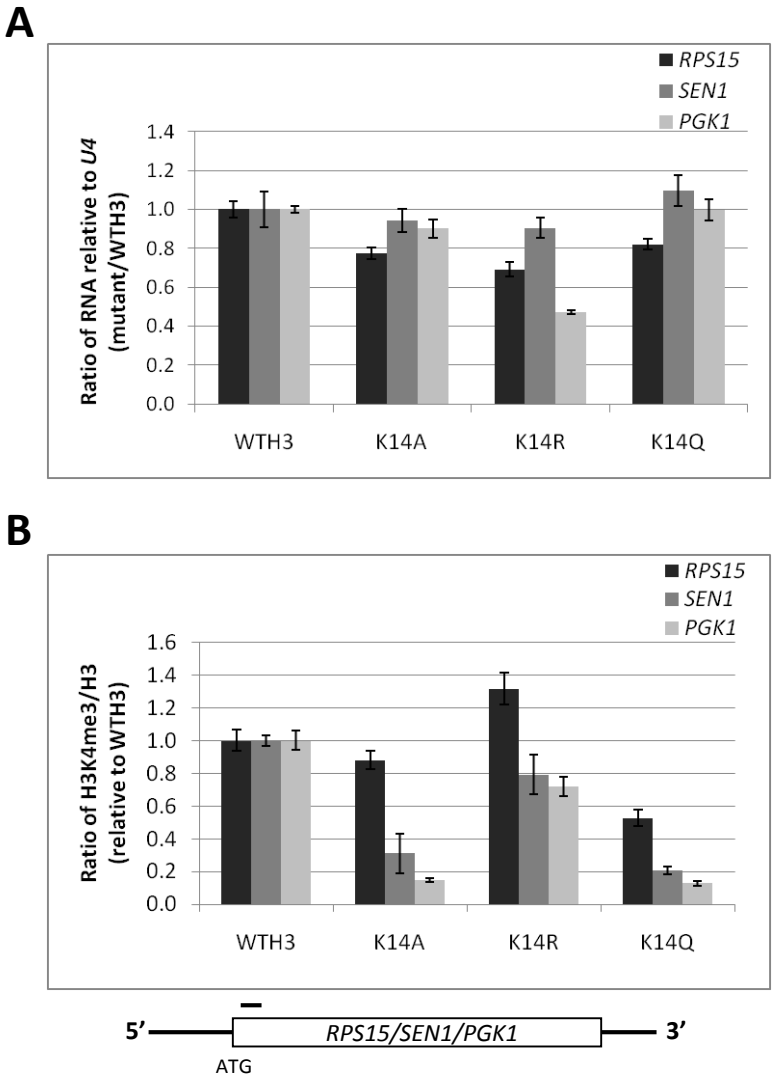


Figure 18. Transcript levels do not correlate with H3K4me3 in the H3K14 substitution strains.

(A) qPCR at the 5' ends of *RPS15, FMP27* and *PGK1* after reverse transcription of RNA extracted from the WTH3 and H3K14 substitution yeast strains. Signals are normalised to levels of *U4* RNA and expressed as a ratio of the WTH3 RNA level. Error bars show standard errors of the real-time PCR reaction. **(B)** ChIP-qPCR of H3K4me3 in the wildtype and H3K14 mutant strains. Signals are normalised to levels of histone H3 and expressed relative to WTH3. Error bars show standard errors of the real-time PCR reaction (reproduced from Figure 14A).

As mentioned, the transcription field is becoming increasingly aware that many chromatin regulators may exert their effects more during dynamic rather than steady-state transcription (Weiner et al. 2012), perhaps partially explaining the very small differences in transcript levels observed between individual genes with different levels of H3K4me3 in the H3K14 mutant strains (Figure 17). Therefore transcript levels in the wildtype and K14A strains were assessed during the diauxic shift, a period that is accompanied by the up- and down-regulation of a large number of

genes including *CIT2* and *RPS15* respectively (Figure 19A, B). (Both the un-normalised and *U4*-normalised data are shown in case the assumption that levels of *U4* RNA do not change during the diauxic shift is not valid.) It is apparent that the K14A mutation affects the kinetics of the changes in transcript levels: *CIT2* is induced faster in the K14A strain and reaches higher absolute levels. In contrast, *RPS15* is repressed more slowly in the K14A strain, although the same absolute levels are reached by 25 h growth. However, these changes in kinetics have little relation to the alterations in H3K4me3 levels at these genes during the diauxic shift (Figure 19C). Thus whilst H3K14 is important for maintaining proper levels of all tested transcripts during steady-state exponential growth and the diauxic shift, these effects do not correlate with the various levels of H3K4me3, neither those in different H3K14 substitution strains nor those found at individual genes. Instead the combined loss of H3K14 integrity and H3K18 acetylation may account for the changes observed.

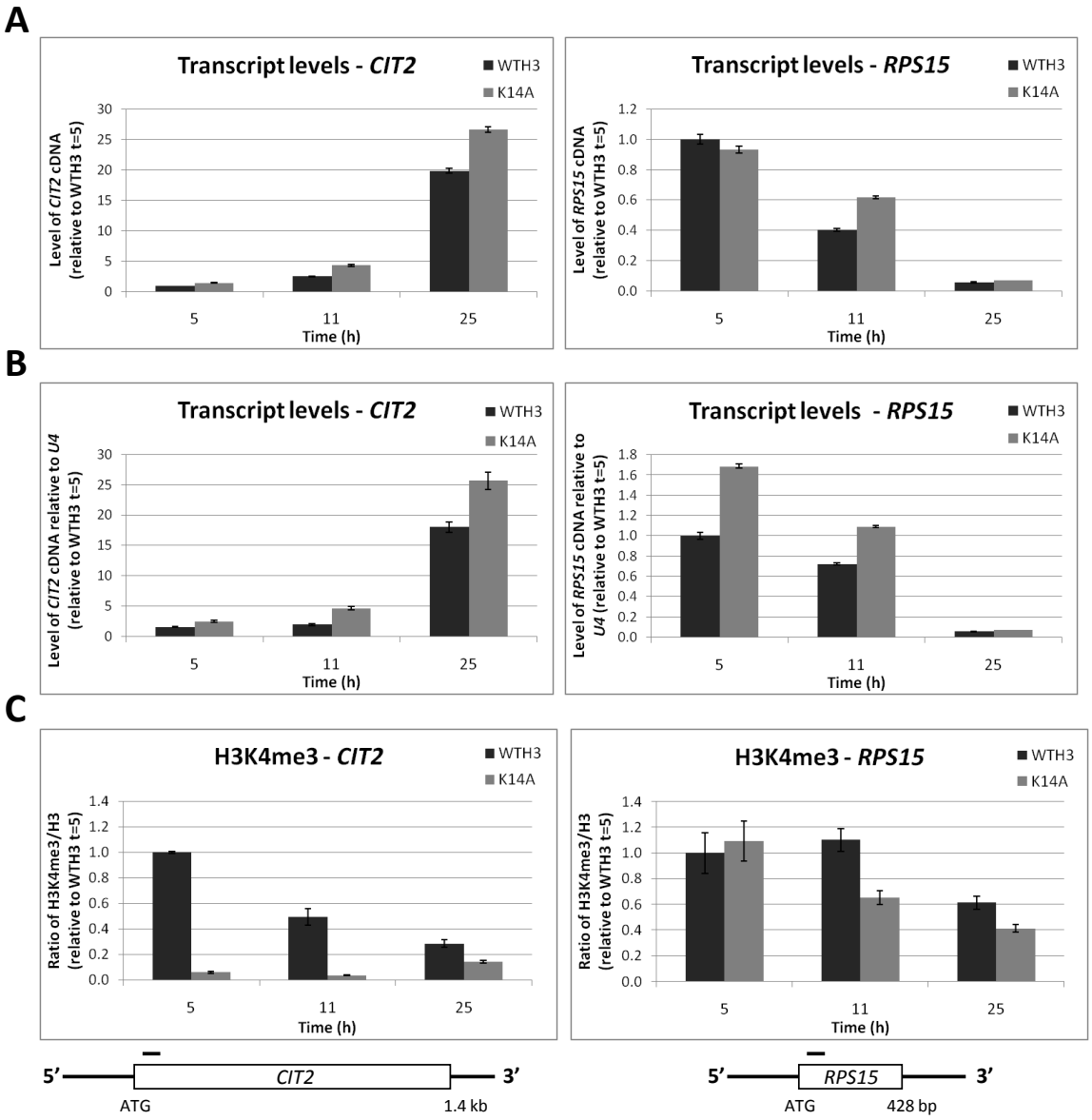


Figure 19. Mutation of H3K14 affects changes in transcript levels during the diauxic shift.

(A) qPCR at the 5' ends of *CIT2* and *RPS15* after reverse transcription of RNA extracted from WTH3 and K14A yeast strains at the times indicated (h). Signals are expressed relative to the WTH3 RNA level at t=5. Primer positions are indicated on the locus maps. Error bars show standard errors of the real-time PCR reaction. **(B)** As (A) except signals are normalised to levels of *U4* RNA. **(C)** A ChIP experiment showing the levels of H3K4me3 in the WTH3 and K14A strains at the times indicated (h). Signals are normalised to levels of histone H3 and are expressed relative to the WTH3 H3K4me3/H3 level at t=5. Error bars show standard errors of the real-time PCR reaction.

3.2.8. Genes with the lowest ratios of H3K4me3 are enriched for antisense transcription in the wildtype strain

In order to dissect any differences in sense or antisense transcription between the genes in the two defined K14A classes, the Churchman and Weissman (2011) dataset was used. Their novel technique involves purifying the RNA associated with RNA polymerase II from a wildtype strain (BY4741), which allows the specific sequencing of nascent transcripts and therefore mapping of active transcription. Firstly, the average level of transcription in the sense direction across the transcription unit for the genes in the 'top' and 'bottom' classes of K14A-affected genes was calculated (Figure 20A). Genes in the 'top' class have a higher level of sense transcription than both the genes in the 'bottom' class and the average of all genes in the genome. This high level of sense transcription is consistent with the enrichment of ribosomal protein genes in the 'top' class of genes, which are genes known to account for approximately 50 % of RNA polymerase II transcription events in the cell (Warner 1999). Next, the average level of antisense transcription in the first 300 bp of the genes was calculated. Since it is known that the majority of antisense transcripts run into the promoter regions of their cognate genes (Xu et al. 2011), this window of antisense transcription was chosen to circumvent any issues relating to the start sites of antisense transcription, which are poorly mapped. In contrast to sense transcription, the level of antisense transcription was much higher for the genes in the 'bottom' class of K14A-affected genes than both those in the 'top' class and the average of all genes in the genome (Figure 20B). Once again, this finding is supported by reports in the literature that stress-response genes, constituting 10 % of the 'bottom' class, are enriched for antisense transcripts when their sense transcription is repressed (Yassour et al. 2010; Xu et al. 2011), as was presumably the case given the standard experimental growth conditions used by Churchman and Weissman (2011).

There is a link between antisense transcription and H3K4me3 since this modification has been shown to enhance Nrd1-dependent termination of non-coding transcripts (Terzi et al. 2011).

Creamer et al. (2011) have produced a transcriptome-wide map of Nrd1 binding sites using an *in vivo* photoactivatable protein-RNA crosslinking immunoprecipitation technique (PAR CLIP). This dataset was used to analyse the distribution of Nrd1 on RNA transcribed over the 5' and 3' ends of the genes in the 'top' and 'bottom' K14A-affected classes in the wildtype strain (Figure 20C, D). The average distribution of Nrd1 for all genes in the genome has been included for comparison. As is clearly visible, Nrd1 is present at much higher levels at both the 5' and 3' ends of the genes in the 'top' class compared to the genes in the 'bottom' class and the genomic average. The low levels of Nrd1 on RNA produced from the 3' ends of genes in the 'bottom' class agrees with the high levels of antisense transcription of these genes: Nrd1 is not present to terminate these RNAs and prevent their transcription into the gene promoter. Interestingly, Nrd1 at the 5' end of the average gene is distributed in two peaks: one at the transcription start site (TSS) and one 250-750 bp upstream. Since the characterised transcripts terminated by Nrd1 include CUTs, which are generally short (200-800 bp), these two peaks of Nrd1 likely represent termination of the sense and antisense CUTs originating from bidirectional gene promoters (Neil et al. 2009; Xu et al. 2009). In contrast to the Nrd1 profile on the average gene, at the genes in the 'top' class, the distribution of Nrd1 is shifted towards a predominant localisation at the TSS, with an almost non-existent upstream peak. In support of a function of Nrd1 at these genes, Nrd1 negatively regulates the levels of the *RPB10* transcript, a gene in the 'top' class, presumably by binding near its 5' end and inducing premature termination (Creamer et al. 2011). Therefore the high level of Nrd1 binding at the 5' ends of the genes in the 'top' class could indicate that Nrd1-induced premature transcription termination may be a common mechanism for the regulation of these genes.

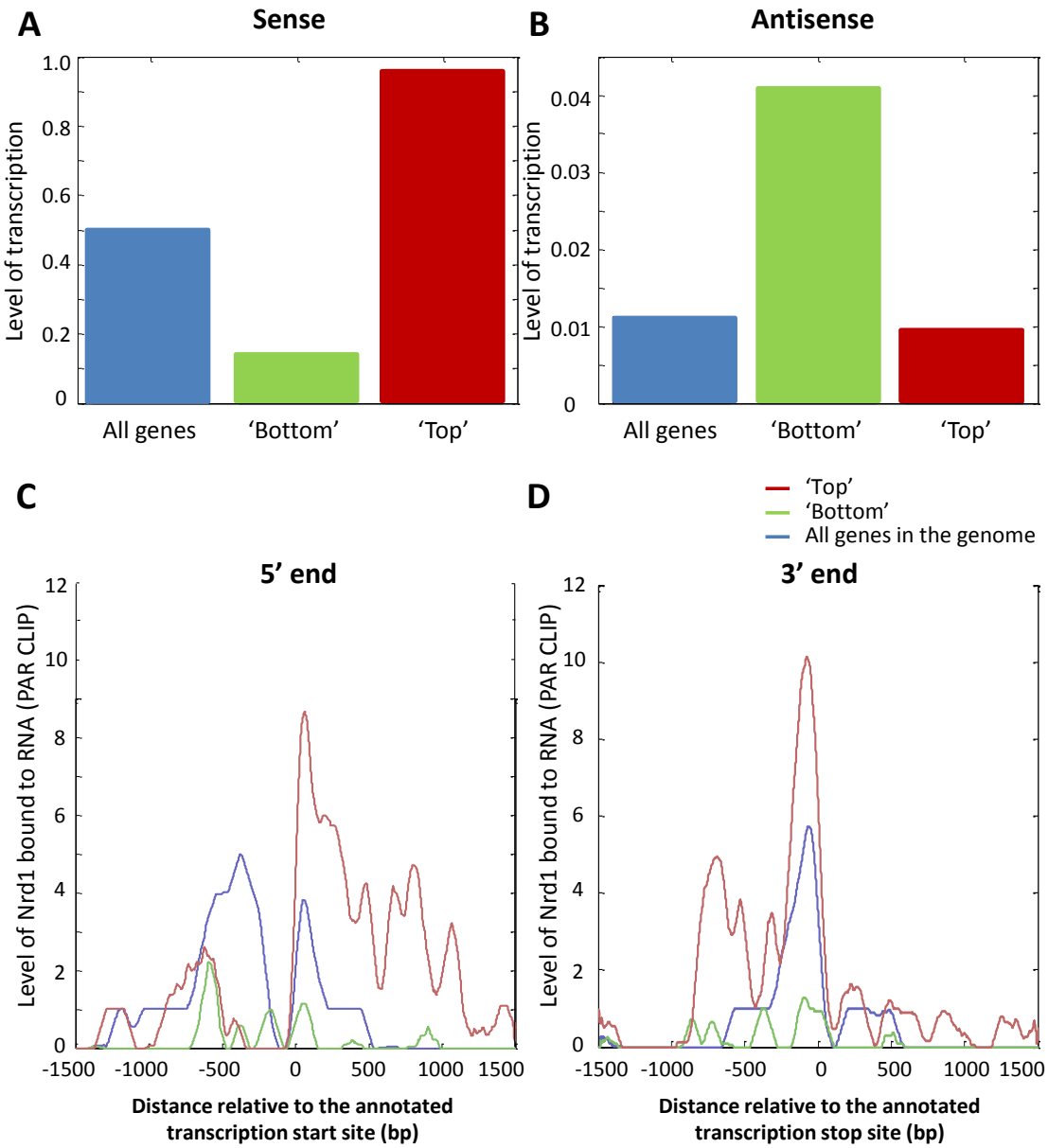


Figure 20. Genes with the 200 smallest ratios of H3K4me3 (K14A/WTH3) are enriched in antisense transcripts and depleted for Nrd1.

(A) The average wildtype levels of sense transcription across all genes in the genome and the genes in the 'top' and 'bottom' K14A-affected classes using the NET-seq dataset (Churchman and Weissman, 2011) (S. Murray). **(B)** As (A) except for antisense transcription in the first 300 bp of the genes. **(C)** The average wildtype distribution of Nrd1 bound to RNA transcribed from the 5' ends of all genes in the genome (blue) and the genes in the 'top' (red) and 'bottom' (green) K14A-affected classes using the PAR CLIP dataset (Creamer et al. 2011) (S. Murray). **(D)** As (C) except at the 3' ends of genes.

To evaluate whether differential H3K4me3 levels may be responsible for the distinct distributions of Nrd1, and therefore antisense transcription, in the 'top' and 'bottom' classes, the same analysis was performed using the ChIP-seq wildtype H3K4me3 dataset collected in this study (Figure 21). At the 3' ends of the genes, the levels of H3K4me3 correlate with the levels of Nrd1 in the two classes. However, at the 5' ends of the genes, the high levels of Nrd1 on the transcripts from the 'top' class of genes do not correspond to similarly high levels of H3K4me3. Therefore, at least at the 5' ends of the genes in the 'top' class, Nrd1 binding is not strictly dependent on H3K4me3. However, the unusually high level of H3K4me3 at the 3' ends of genes in the 'top' class may partially explain the high level of Nrd1 and consequent low levels of antisense transcription of these genes.

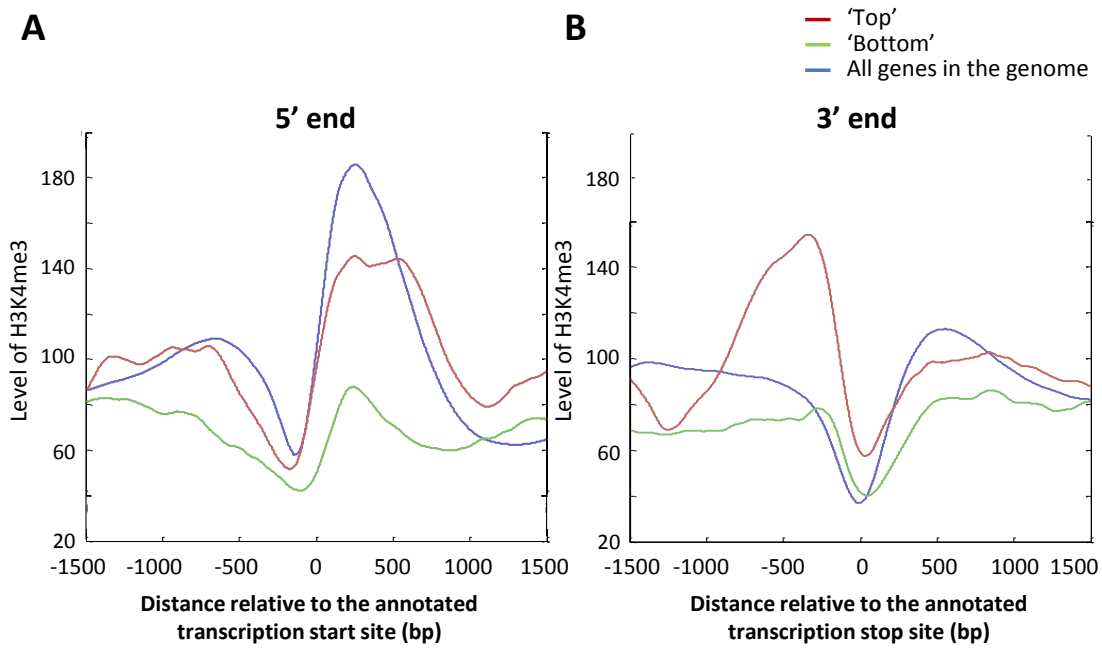


Figure 21. Genes with the top 200 ratios of H3K4me3 (K14A/WTH3) have high levels of H3K4me3 at their 3' ends.

(A) The average distribution of H3K4me3 at the 5' ends of all genes in the genome and the genes in the 'top' and 'bottom' K14A-affected classes using the WTH3 ChIP-seq dataset (S. Murray). **(B)** As (A) except at the 3' ends of genes.

Unfortunately it is not known how the relative levels of sense and antisense transcription change, if at all, in the H3K14 mutant strains. It is tempting to speculate that the lower amount of H3K4me3 on genes in the K14A strain may decrease the efficiency of Nrd1-dependent termination, allowing an increase in antisense transcription originating from the 3' ends of genes. It would be predicted that this effect would be more substantial at the genes in the 'bottom' class that lose more H3K4me3 in the K14A strain than the 'top' class. This increase in antisense read-through transcription could possibly account for the unexpected small relative increases in total transcripts at the 5' ends of the genes with the lowest ratios of H3K4me3 in the K14A and K14Q strains compared to those with the highest ratios (Figure 17 and Figure 18). However, it is also possible that the relationship between H3K4me3 and antisense transcription may operate in the reverse direction. There is an anti-correlation between the level of antisense transcription into a gene promoter and the extent of the 5' H3K4me3 peak (S. Murray, unpub.). Although it is formally possible that the level of H3K4me3 at the 3' end of an antisense transcript may affect its transcription, it is perhaps more likely that antisense transcription may affect H3K4me3. Therefore the reciprocal level of antisense transcription of the genes in the 'top' and 'bottom' classes may itself be responsible for the differential effects of the H3K14 mutation on H3K4me3. This could be a topic of future investigation.

3.3. Discussion

Previous work in this laboratory (D. Clynes) and by others (Nakanishi et al. 2008) has identified a crosstalk between H3K14 and H3K4me3. The work in this chapter has aimed to further investigate this crosstalk, in particular which aspect of H3K14 that is important for H3K4me3 and the potential gene-specific effects of this crosstalk.

3.3.1. H3K14 is not absolutely required for H3K4me3

In the previous study into the H3K14-H3K4me3 crosstalk, Nakanishi et al. (2008) observed undetectable levels of H3K4me3 in three different H3K14 substitution strains. Because all three of these strains lacked H3K14ac, a modification that is known to co-localise with H3K4me3 at the promoters and 5' ends of active genes (Liu et al. 2005; Pokholok et al. 2005) and is present on the same histone H3 tails as H3K4me3 (Hazzalin and Mahadevan 2005; Jiang et al. 2007), they proposed that H3K14ac was required for H3K4me3. There is also evidence in the literature that this relationship might be reciprocal since H3K4me3 interacts both with the PHD finger of Yng1, a component of the NuA3 complex (Martin et al. 2006b; Taverna et al. 2006), and the SAGA subunit, Sgf29 (Vermeulen et al. 2010; Bian et al. 2011). Both of these HAT complexes have been shown to acetylate H3K14 *in vivo* (Grant et al. 1999; Martin et al. 2006a; Jiang et al. 2007). The results in this chapter do not fully support either of these findings. Firstly, the levels of H3K4me3 varied greatly between the three H3K14 substitution strains, with near-wildtype levels of H3K4me3 in the K14R strain *i.e.* in the absence of H3K14ac. Secondly, wildtype global levels of H3K14ac were found in three strains that lacked H3K4me3 or all forms of H3K4 methylation (R2A, *spp1Δ,* and K4A respectively). The reasons for the discrepancies between the results in this chapter and the published data are unknown.

The task now is to try to reconcile the findings in this study with the published results. The link between H3K4me3 and H3K14ac is well-established: these modifications co-localise genome-wide

(Liu et al. 2005; Pokholok et al. 2005) and occur on the same histone tails (Hazzalin and Mahadevan 2005; Jiang et al. 2007). However, this does not necessarily mean that these modifications are co-dependent and it is possible that they could be regulated independently. Arguing against this, there are several reports in the literature of a dependence on H3K4me3 for H3K14 acetylation (Martin et al. 2006b; Taverna et al. 2006; Jiang et al. 2007; Vermeulen et al. 2010; Bian et al. 2011). The data in support of the H3K4me3-H3K14ac crosstalk are quite compelling and it is unclear why an effect of H3K4me3 removal on H3K14ac was not detected in this study. It is possible that the antibody against H3K14ac was not sensitive enough to detect changes in global levels, which might be small if they only occur at a subset of genes. Future experiments should include repeats of the Western blots with serial dilutions to ensure that the levels of H3K14ac are within the linear range of detection by the antibody. Additionally, a ChIP experiment could be performed that might enable identification of any gene-specific crosstalk. It is still possible that no net change in H3K14ac may be observed in the absence of H3K4 methylation since this modification, in addition to recruiting HATs, can also recruit HDACs that act on histone H3 (Kim and Buratowski 2009; Wang et al. 2011).

The evidence for the reverse crosstalk between H3K14ac and H3K4me3 is less substantial. Nakanishi et al. (2008) could not detect H3K4me3 by Western blot in any of the three H3K14 substitution strains and therefore proposed that H3K14ac, which is absent from all three of these mutants, is required for H3K4me3. However, the authors failed to consider the function of the unmodified or the lowly abundant dimethylated lysine 14, states which are also lost from the H3K14 mutant strains. Because their result with the H3K14 mutants is called into question by the findings of this study, their conclusions must also be reviewed. The current findings with the K14R strain are reproducible using several independently constructed strains in different backgrounds and with two separately sourced commercial H3K4me3-specific antibodies. In support of the K14R result, a differential effect between arginine and glutamine substitutions on histone H3

methylation has been observed previously. Psathas et al. (2009) mutated six lysine residues in the histone tail to glutamine. This resulted in decreased H3K36me3 whereas substitution with arginine did not.

Additional evidence that claims to support a role for H3K14ac in H3K4me3 regulation is the decreased global levels of H3K4me3 in the *gcn5Δ* strain (Govind et al. 2007; Jiang et al. 2007). However, Gcn5 acetylates a number of histone residues other than H3K14 (Grant et al. 1999; Jiang et al. 2007) and non-histone proteins, such as Rsc4 (Charles et al. 2011) and participates in many cellular processes including nucleosome assembly (Burgess et al. 2010), DNA replication (Espinosa et al. 2010) and DNA damage repair (Burgess et al. 2010). In particular, the function of Gcn5 in cell cycle progression, which results in the slow growth phenotype of *gcn5Δ* yeast (Zhang et al. 1998) could very likely cause indirect effects on H3K4me3. The fact that H3K4me3 is still reduced by deletion of *GCN5* in the K14R strain strongly suggests that the reason for the decrease in H3K4me3 in the *gcn5Δ* strain is not solely a result of reduced H3K14ac.

There is a small possibility that the function of H3K14 acetylation can be recapitulated by the arginine substitution. Whilst the differences in charge between the neutral acetylated lysine and the positively-charged arginine render it unlikely that arginine can fulfil the same function as H3K14ac in the positive or negative regulation of effector protein binding, H3K14ac may be required to induce a structural change in the histone H3 tail, which could also be promoted by K14R. A study on histone H4 tail acetylation has indicated that the degree of acetylation positively influences the α-helical content of the tail, resulting in its shortening (Wang et al. 2000). This phenomenon is applicable to the histone H3 tail, which is known to adopt α-helical structures *in vitro* (Banères et al. 1997) and can also contain multiple acetylated lysine residues. The structure of the H3 tail is likely to be important for the binding of protein effectors or large complexes like COMPASS that can influence H3K4 methylation.

Another, perhaps more likely, interpretation is that the near-wildtype levels of H3K4me3 in the K14R strain suggest that the positively-charged, unmodified lysine residue may be sufficient for the majority of H3K4me3 in the wildtype strain. (It should be noted though that the K14R strain does have reduced levels of H3K4me3 compared to wildtype, and so arginine cannot fully substitute for H3K14 in the promotion of H3K4me3.) Thus, whilst H3K4me3 and H3K14ac may co-localise on histone tails, dynamic acetylation of H3K14 may potentially be required to allow H3K4 to be trimethylated during the window when H3K14 is unmodified. Experiments in mouse CH3 10T½ cells have shown that continuous dynamic turnover of acetylation is a feature of all histone H3 methylated on lysine 4 (Hazzalin and Mahadevan 2005). The proposed requirement for dynamic acetylation is supported by the severe effect of the glutamine substitution on H3K4me3. Glutamine is frequently used as an acetyl-lysine mimic and so it can be inferred that the loss of temporal and spatial regulation of acetylation at H3K14 would be detrimental to H3K4me3. Together, the different results with the K14R and K14Q strains suggest that the unmodified and acetylated H3K14 residues may have different functions in the modulation of H3K4me3 levels, an idea that will be revisited in the main discussion (Chapter 7).

An alternative hypothesis is that H3K4me3 may be affected indirectly via changes in transcription in the H3K14 mutant strains. The loss of both H3K14ac and H3K18ac in the absence of H3K14 may result in a decrease in transcription since these modifications are known to correlate with transcript levels (Liu et al. 2005; Pokholok et al. 2005). Indeed, the levels of the transcripts tested in this study dropped to approximately 60-90 % in the H3K14 mutant strains compared to the wildtype strain. Since H3K4me3 has been shown to be deposited co-transcriptionally (Krogan et al. 2003a; Ng et al. 2003a), a reduction in transcription could result in H3K4 methylation defects in the H3K14 mutant strains. However, all three H3K14 substitution strains lack H3K14ac and H3K18ac but yet have varying levels of H3K4me3. Furthermore, at several genes the transcript levels but not H3K4me3 are more reduced by the K14R substitution than the alanine or glutamine

substitutions. Therefore, an indirect effect of H3K14 on H3K4me3 via a reduction in transcription is unlikely to be the main reason for the global decrease in H3K4me3 in these strains.

3.3.2. H3K14 is required for H3K18ac

The mechanism by which H3K14 promotes H3K18ac is not known. One working hypothesis is that H3K14ac, lost in the three H3K14 mutant strains, interacts with the Gcn5 bromodomain and recruits this HAT to chromatin to acetylate H3K18, one of its main substrates (Jiang et al. 2007). Several studies would support this idea. The Gcn5 bromodomain interacts with acetylated H3, especially when acetylated at H3K14 (Hassan et al. 2002, 2007), but also H4K16ac (Owen et al. 2000; Li and Shogren-Knaak 2009), to recruit SAGA to nucleosome arrays *in vitro*. Furthermore, Sterner et al. (1999) observed a decrease in the degree of histone H3 acetylation by SAGA in a strain with a Gcn5 bromodomain deletion and Li and Shogren-Knaak (2009) demonstrated a cooperativity of SAGA-mediated acetylation on neighbouring histone H3 tails that was dependent on both the prior acetylation of one tail and the Gcn5 bromodomain. Attempts in the current study to assess the levels of chromatin-bound Gcn5 in the presence or absence of its bromodomain in the wildtype and K14A strains were unsuccessful. Because of the proximity between the site of mutation and the modification affected, an alternative but related hypothesis is that H3K14, in either its modified or unmodified state, forms part of the binding site for the H3K18 HAT, and that mutation of this residue disrupts HAT binding. *In vitro* peptide pull-down assays may help to investigate this possibility.

3.3.3. H3K14 substitution differentially affects H3K4me3 levels on genes that are expressed in opposite metabolic states

The ChIP-seq experiment measuring H3K4me3 in the wildtype and K14A strains demonstrated that mutation of H3K14 decreases H3K4me3 to different extents at individual genes. (Interestingly there are a few genes for which the levels of H3K4me3 were increased by the K14A mutation. All

of these are at dubious or uncharacterised ORFs.) When the genes with H3K4me3 that was decreased the least and most by the K14A mutation were analysed, it was found that they segregated into distinct functional groups (Table 9). The genes with H3K4me3 that was least affected by the K14A mutation ('top' class) encoded proteins involved in ribosome biogenesis and translation whereas the genes with H3K4me3 that was most decreased by the K14A mutation ('bottom' class) encoded proteins that were involved in processes such as the response to environmental stresses, thiamine and sulphur compound biosynthesis, and carbohydrate catabolism. The genes in these two classes of K14A-affected genes overlapped significantly with those oppositely regulated by the environmental stress response and those transcribed in different phases of the yeast metabolic cycle. The results from this analysis therefore imply that the gene classes differentially affected by H3K14 are subject to different regulation, dependent on the intracellular metabolic status.

K14A-affected class	Predominant GO terms	Expression during the ESR	Expression phase in YMC
'Top'	Translation	Repressed	Oxidative (growth)
'Bottom'	Stress response, carbohydrate transport and catabolism	Induced	Reductive/charging (quiescent-like)

Table 9. Summary of the features of genes in the two differentially affected K14A classes.

The predominant gene ontology (GO) terms, expression during the environmental stress response (ESR) and phase of expression during the yeast metabolic cycle (YMC) of the 200 genes with H3K4me3 least ('top') and most ('bottom') decreased by mutation of H3K14.

The transcripts from genes in the 'top' class peak when cells are undergoing oxidative metabolism and growth but are repressed in response to environmental stress. Therefore, in the experimental conditions used, these genes are likely to be actively transcribed. The transcription of the genes required for growth is known to be positively regulated by Gcn5/SAGA-mediated histone acetylation (Huisinga and Pugh 2004; Robert et al. 2004), agreeing with the high global and gene-

specific levels of histone H3 and H4 acetylation (except H3K56ac and H4K16ac) during the oxidative phase of the yeast metabolic cycle (Cai et al. 2011). It is therefore curious that these genes, with high levels of H3K14ac, show the least dependence on H3K14 for H3K4me3, further questioning the requirement of H3K14ac for H3K4me3. As described, the genes in the 'top' class have a high level of sense transcription in the wildtype strain. Since there is a known correlation between the levels of transcripts/transcription and the amount of H3K4me3 on a gene, and there is evidence that H3K4me3 can be deposited co-transcriptionally (Krogan et al. 2003a; Ng et al. 2003a), it is therefore possible that the high level of sense transcription of these genes allows H3K4me3 to be maintained co-transcriptionally in the absence of H3K14 in the mutant strains. The levels of sense transcription in the H3K14 mutant strains would need to be obtained to add weight to this hypothesis.

In contrast, genes in the 'bottom' class are expressed under less favourable conditions for growth. It is therefore likely that these genes were transcribed at very low levels in this experiment. Under these experimental conditions, there is a high level of antisense transcription in the 5' regions of these genes in the wildtype strain. Unfortunately, the levels of strand-specific transcription in the K14A strain are not known. However, it is possible that the high level of antisense transcription in the wildtype strain may make these genes more susceptible to H3K4me3 loss upon mutation of H3K14 than the genes in the 'top' class. Furthermore, the low level of sense transcription of these genes removes the potential for co-transcriptional deposition of H3K4me3, which may expose this modification to more severe losses in the absence of H3K14 than the genes in the 'top' class with high sense transcription. However, the genome-wide Pearson correlation between the ratio of H3K4me3 (K14A/WTH3) and the wildtype level of sense transcription is 0.158 and so the level of transcription in the wildtype strain cannot fully explain the differential effects of the K14A mutation on H3K4me3. It is plausible that the K14A mutation could cause changes in the relative

levels of transcription at individual genes and that this could result in the non-uniform effects of this mutation on H3K4me3.

In an extension to the above idea, it is possible that mutation of H3K14 changes the relative lengths of the phases of the YMC and that this relative change in phase duration could cause the differential effects on the genes expressed in the separate phases. For instance, if the length of the oxidative phase increased relative to the reductive phases in the K14A strain, even more of the cells in the population would be in the oxidative phase at any one time than for the wildtype strain. Therefore, at a population level, the transcription of the genes that peak during the OX phase would be increased relative to those in the reductive phases and consequently the potential for co-transcriptional H3K4me3 deposition would also increase. Conversely, a shortened duration of the reductive phases in the K14A strain would result in lower transcription of the reductive phase genes, causing decreased co-transcriptional H3K4me3 deposition on these genes relative to those peaking in the oxidative phase. Furthermore, the high level of antisense transcription of the reductive phase genes when they are repressed (Machné and Murray 2012) may exacerbate the reduction in H3K4me3 (S. Murray, unpub.). The fact that deletion of or catalytically mutating *GCN5*, which results in strong reductions in several histone H3 acetylation marks including H3K14ac and H3K18ac (modifications that are also lost in the H3K14 mutant strains), inhibits metabolic cycling could support this hypothesis (Cai et al. 2011).

Given the hypothesised longer duration of the OX phase in the K14A strain, it might be expected that the levels of H3K4me3 on the OX phase genes might *increase* relative to the wildtype strain. Except for a few genes, an increase in H3K4me3 was not observed in the K14A strain. This would indicate that there are two H3K14-dependent factors that influence levels of H3K4me3. Firstly, the mutation of H3K14 causes a global decrease in the level of H3K4me3. The possible mechanisms for this crosstalk will be explored in Chapter 5. Secondly, the potential K14A-

mediated shift in the YMC from reductive to oxidative phases could differentially affect the extent of the H3K4me3 loss on individual genes so that H3K4me3 is reduced less on the OX phase genes than the R/C phase genes. Additionally, the different modes of regulation described for the genes in the 'top' and 'bottom' K14A-affected classes may alter the susceptibility of H3K4me3 to mutation of H3K14.

Experiments would need to be performed to gain evidence to support the proposed H3K14-mediated influence on YMC phase duration. The ideal experiment would be to assess the YMC in the wildtype and K14A strains via changes in dissolved oxygen concentration in a continuously slow-growing culture in a chemostat. However, the lack of access to a working chemostat precludes this approach. There are other indirect ways to evaluate the proportion of cells in the oxidative phase. One possibility is to analyse the expression of selected genes that are known to peak in the two phases by measuring transcript levels. This would have to be done in a strand-specific manner to avoid the interpretation issues encountered with reverse transcription in this project. An alternative suggested measure of cells in the oxidative phase is their increased mitochondrial activity, resulting in a greater mitochondrial membrane potential relative to cells in the reductive phases (Slavov et al. 2012). The dye $DiOC_6$ can be used as an indicator of mitochondrial membrane potential and the prediction of this hypothesis would be that the cells from the K14A strain would stain more strongly with $DiOC_6$ than the wildtype strain. Lastly, Slavov et al. (2012) have demonstrated that cells in the reductive phase are more resistant to heat shock than those in the oxidative phase. A prediction of the K14A YMC hypothesis would therefore be that the K14A strain would be more sensitive to heat shock than the wildtype strain. Although no differential sensitivity to heat shock was displayed in this project (data not shown), the low severity of the heat shock applied (42°C, 15 min) and the subsequent drop plate assays may not have been sensitive enough to observe a decrease in cell viability. An experiment exposing cells to increasing lengths of heat shock at 50°C followed by assessment of viability in liquid culture using

the Bioscreen spectrophotometer should be performed to assess any differences in viability after stress between the wildtype and K14A strains.

It should be noted that there are some caveats to this hypothesis. There is still debate in the literature as to whether the YMC occurs during the very rapid growth occurring under standard laboratory conditions (Slavov and Botstein 2011; Slavov et al. 2012). The Botstein lab have demonstrated that the duration of the oxidative phase increases relative to the reductive phases with the growth rate of the culture (Slavov and Botstein 2011). This means that at high growth rates, a very low proportion of the population is in the reductive phase, causing the detection of cells in this phase to be difficult. Therefore it may be possible that fast-growing cultures do undergo metabolic cycling but that this cannot be detected. Following on from this idea of the duration of the oxidative phase increasing with growth rate, it would be predicted that if the K14A strain spent longer in the oxidative phase that this strain would have a higher growth rate than the wildtype strain. Whilst the K14A (and K14Q) strains did have slightly shorter doubling times than the wildtype strain, this difference was not statistically significant. Additionally, an assumption has been made that peaking of transcripts during the YMC is due to an increase in transcription rather than a decrease in RNA degradation. This may not be completely valid since mRNAs are differentially targeted for degradation by the TOR complex, a major regulator of growth that controls many of the processes occurring periodically during the YMC (Albig and Decker 2001). However, despite these caveats, H3K14-dependent regulation of the YMC is an attractive hypothesis that can explain the differential effects of H3K14 mutation on genes with diverse cellular functions.

Chapter 4.

Investigating the function of H3P16

4. Investigating the function of H3P16

4.1. Introduction

Work in the previous chapter has demonstrated that the residue at position 14 in the histone H3 tail has the ability to influence the levels of a relatively distant modification, H3K4me3. It is possible that other residues in this region of the tail may also regulate H3K4me3. It was decided to concentrate the investigation on H3P16 because it is close to H3K14 and as discussed below, via the ability of the peptidyl-prolyl bond to undergo *cis-trans* isomerisation, may be capable of exerting long-range effects on histone H3 modifications.

4.1.1. Proline can undergo cis-trans isomerisation

Proline, as the sole cyclical amino acid, is the only residue that can undergo *cis-trans* isomerisation of the peptide bond it shares with its neighbouring residue due to the lower steric hindrance between neighbouring C_α atoms upon formation of the *cis* peptide bond when one of the participating residues is proline. However, even with a proline residue, formation of the *cis* peptidyl-prolyl bond is still less energetically favourable than the *trans* peptidyl-prolyl bond and so *cis-trans* isomerisation is a slow process (in the order of minutes). However, there are proteins *in vivo* that aid this conversion by lowering the activation energy: these are termed peptidyl-prolyl *cis-trans* isomerases (PPIases) (reviewed by Lu et al. 2007). The dihedral angle differs by 180° between the *cis* and *trans* states and so this isomerisation has the potential to dramatically affect the surrounding protein structure (Lu et al. 2007).

4.1.2. Modifications of the histone tail affect its structure

Histone tails are often thought of as unstructured regions since they are too flexible to appear in the X-ray crystal structure of the nucleosome (Luger et al. 1997a). However, other techniques show that this assumption may not be true. For example, roughly 50 % of the histone H3 and H4

N-terminal tails were determined by circular dichroism to adopt an α-helical conformation when bound to DNA in the nucleosome (Banères et al. 1997). The histone H3 tail contains three proline residues: H3P16, P30 and P38. Since proline is known to disrupt α-helices (Gunasekaran et al. 1998), these residues may therefore act as hinge points in the tail, with *cis-trans* isomerisation of the peptidyl-prolyl bond having the potential to dramatically affect the relative trajectories of the flanking helices. A similar property has been ascribed to a conserved proline in elastin-like polypeptides (Glaves et al. 2008).

There is some evidence that post-translational modifications of histone tails can influence their structure. For example, a study on histone H4 has indicated that acetylation increases the α-helical content of the tail, resulting in tail shortening (Wang et al. 2000). Furthermore, *in silico* dynamics simulations have also hinted at a less disordered structure of the tail in its modified form (Sanli et al. 2011 and H. Patterton, pers. comm.). Interestingly, simulations predict that H3P16 adopts the *cis* conformation when the H3 tail (residues 1-43) contains H3K4me3, H3K9ac, H3K14ac and H3K36me3 (H. Patterton, pers. comm.).

4.1.3. Proline can affect deposition of neighbouring modifications

The work described above has suggested that modifications on the histone H3 tail can affect its structure, potentially by affecting the *cis-trans* isomerisation of the three peptidyl-prolyl bonds. There is also evidence to suggest that this relationship may be reciprocal since the binding of some histone modifying enzymes is predicted to be altered by modification-induced changes in backbone dihedral angles (Sanli et al. 2011). This creates the possibility whereby one modification can induce a conformational alteration in the H3 tail structure that promotes or lessens the binding ability of an enzyme depositing or removing a separate modification.

A more direct example of a proline residue affecting a histone modification is the crosstalk between H3P38 and H3K36me3 (Nelson et al. 2006). The authors mutated H3P38 to valine and observed a decrease in H3K36me3 that was not a result of antibody epitope effects. They identified a PPIase, Fpr4, which binds to the histone H3 tail and promotes *cis-trans* isomerisation of H3P30 and H3P38 (but not H3P16). Using a catalytic mutant of Fpr4, which presumably resulted in increased H3P38*trans* (and/or H3P30*trans*), they demonstrated that this assumed structure adopted by the H3 tail was more favourable for Set2-mediated H3K36 methylation *in vivo*. Interestingly, this study also showed that H3K36 methylation abrogated the Fpr4-mediated isomerisation of H3P38 on peptides *in vitro*, suggesting that there is once again reciprocity in the structure-modification crosstalk.

Like the histone tails, the RNA polymerase II subunit Rpb1 C-terminal domain (CTD) is also highly flexible and does not appear in the crystal structures of this complex (Cramer et al. 2001; Armache et al. 2005). However, residual structural elements have been identified in this domain (reviewed by Meinhart et al. 2005). The CTD heptapeptide repeat contains two proline residues at positions 3 and 6. Three structures of an RNA polymerase II CTD peptide in complex with domains from interacting binding partners, including the termination factor Pcf11, show that prolines 3 and 6 are in the *trans* configuration (reviewed in Meinhart et al. 2005). Bound proline 6 can also be in the *cis* conformation, such as when interacting with the serine 2 phosphatase Ssu72 (Xiang et al. 2010; Werner-Allen et al. 2011). Therefore, the conformations of the proline residues are likely to be important for the binding of protein cofactors, including those that modify the CTD. CTD proline isomerisation is catalysed by the PPIase Ess1 (Pin1 in humans) (Hani et al. 1995; Wu et al. 2000; Krishnamurthy et al. 2009). *ESS1* is an essential gene in yeast (Hanes et al. 1989) and has been shown to influence the phosphorylation state of the RNA polymerase II CTD in two ways. Firstly, binding of the CTD by Pin1, inhibits the Fcp1-mediated dephosphorylation of serine 2, although this inhibition possibly operates by steric hindrance rather than Ess1 PPIase activity (Xu

et al. 2003; Palancade et al. 2004). Secondly, Ess1 is required for Ssu72-mediated dephosphorylation of serines 5 and 7 (Krishnamurthy et al. 2009; Singh et al. 2009; Xiang et al. 2010; Werner-Allen et al. 2011; Bataille et al. 2012; Ma et al. 2012). By modulating the ability to dephosphorylate the CTD and therefore the binding of cofactors that recognise specific CTD phosphorylation states, Ess1 can affect many transcription-based processes. These examples on the RNA polymerase II CTD, with a similar flexible structure to the histone H3 tail, demonstrate the ability of proline residues to influence the deposition of neighbouring modifications.

4.1.4. Studying the effects of proline in vivo and in vitro

Because of the spontaneous, albeit slow inter-conversion between the *cis* and *trans* states, the study of peptidyl-prolyl isomerisation *in vivo* and *in vitro* is challenging. *In vitro*, the isomerisation state of the peptidyl-prolyl bond in a peptide can be fixed by hydroxylation of the C-4 position in the pyrrolidine ring (Taylor et al. 2005). *In vivo*, however, this approach is not possible and, without knowing the specific PPIase, the only current method for studying the importance of *cis-trans* isomerisation of a peptidyl-prolyl bond is to mutate the proline residue. Substitution with any other amino acid would prevent *cis-trans* isomerisation, but valine has been used in previous studies due to its similar mass to proline (Nelson et al. 2006). The disadvantage of both of these approaches is that it is impossible to distinguish the effect of loss of *cis-trans* isomerisation from any effects caused by either hydroxylation of the pyrrolidine ring or mutation of the residue.

4.1.5. Aims

The aim of this chapter is to investigate the function of H3P16 in the histone H3 tail. Given the documented involvement of a proline residue in histone crosstalk and the potential for structural changes mediated by proline isomerisation, the ability of H3P16 to influence the levels of histone H3 modifications, in particular H3K4me3 will be tested and the potential for crosstalk with the neighbouring H3K14 residue will be explored.

4.2. Results

4.2.1. H3P16 is important for the interaction of proteins with adjacent sites

As mentioned, *in silico* molecular dynamics simulations indicate that the modifications on the histone H3 tail can influence its structure, with H3P16 predicted to be in the *cis* conformation when the tail is decorated with 'active' modifications (H. Patterton, pers. comm.). Therefore, it was decided to test whether the reverse was true: whether the structures of the histone H3 tail adopted upon *cis-trans* isomerisation of H3P16 could affect tail modifications, potentially via the altered binding of effector proteins.

Two assays were designed in order to probe whether the removal of the proline at position 16 in the H3 tail affected the binding of proteins known to associate with that region of the tail. Firstly, the binding of the Spt7 bromodomain to the H3P16-adjacent H3K18ac was assessed by surface plasmon resonance (SPR). Bromodomains are characterised acetyl-lysine binding domains and are found in a number of chromatin-associated proteins (Taverna et al. 2007). Full-length and C-terminally truncated Spt7 are components of the SAGA and SLIK complexes respectively (Pray-Grant et al. 2002; Sterner et al. 2002; Wu and Winston 2002). The function and binding specificity of the Spt7 bromodomain are currently unknown (Gansheroff et al. 1995; Sterner et al. 1999) but this domain does preferentially interact with hyperacetylated histones *in vitro* (Hassan et al. 2007). The binding constants for the interactions between the Spt7 bromodomain and various modified and substituted histone H3 peptides containing residues 11-26 were measured by SPR (Figure 22A). The Spt7 bromodomain specifically recognised acetylated H3K18, as evidenced by the lack of binding to the unmodified peptide. Interestingly, substitution of H3P16 with valine increased the binding of the Spt7 bromodomain to the H3K18-acetylated H3 peptide. Furthermore, the P16V substitution was able to rescue the abrogation of binding caused by substitution of H3K14 with arginine.

Another assay to test the importance of H3P16 in protein binding makes use of two commercial antibodies that are raised against separate modifications adjacent to H3P16: H3K14ac and H3K18ac. The dot blots in Figure 22B show the specific binding of these antibodies to the acetylated over the unmodified H3 peptides. However, mutation of H3P16 to valine abolished binding of both of these antibodies to their cognate epitopes. These two examples demonstrate that H3P16 can indeed modulate the binding of protein effectors to the surrounding region of the histone H3 tail but interestingly this effect can be either positive or negative and presumably depends on the nature of the binding interface between the protein effector and the histone H3 tail. It is tempting to speculate that the differences in the P16V-mediated effect may reflect the predominant isomerisation state of the proline residue during these interactions. However, it should be noted that the valine substitution employed in these assays does not allow conclusions to be drawn as to whether the binding phenotypes are a result of loss of isomerisation and altered structure at that region of the H3 tail or simply due to the replacement of the proline by a different residue.

A

H3 peptide	K_d (M)
Unmodified	-
K18ac	1.3×10^{-4}
P16V-K18ac	6.0×10^{-5}
K14R-K18ac	-
K14R-P16V-K18ac	1.0×10^{-4}

B

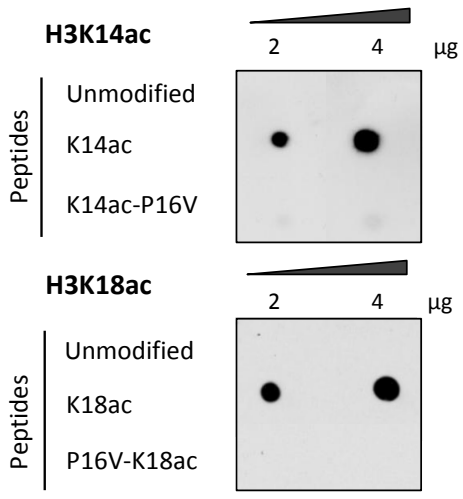


Figure 22. The involvement of H3P16 in the binding of proteins to adjacent sites on the histone H3 tail.

(A) Binding constants as measured by surface plasmon resonance of purified recombinant Spt7 bromodomain to the indicated H3 peptides (I. Boubriak). **(B)** Dot blots showing the abrogation of H3K14ac and H3K18ac antibody binding to 2 and 4 µg of P16-substituted acetylated H3 peptides.

4.2.2. H3P16 affects the global levels of H3K4me3

Having established the fact that H3P16 can influence the binding of protein effectors, it was next explored whether mutation of H3P16 affects the modifications on the histone tail, in particular methylation of H3K4. To this end, whole cell extracts were prepared from the wildtype and P16V strains and probed for H3K4me2 and H3K4me3 by Western blot (Figure 23A). Whereas global levels of H3K4me2 were unaffected by mutation of H3P16, H3K4me3 was reduced. This finding was confirmed by ChIP at *FMP27* (Figure 23B). A similar decrease in H3K4me3 was observed in a P16A strain (M. Sale, data not shown). It is interesting to note that both H3K14 and H3P16 participate in trimethyl-specific crosstalk with H3K4, raising the possibility that these two crosstalk mechanisms may be linked.

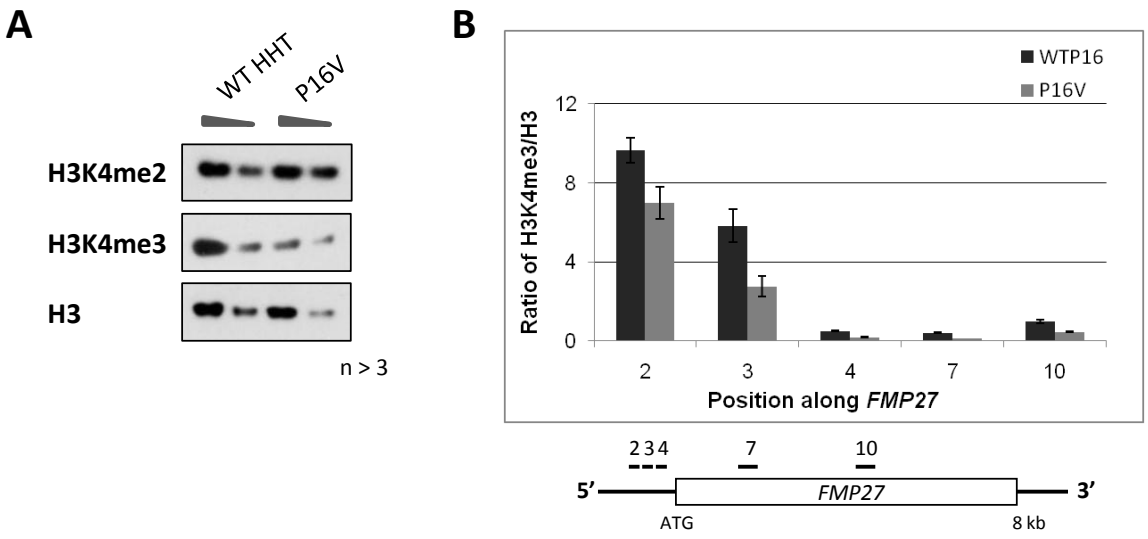
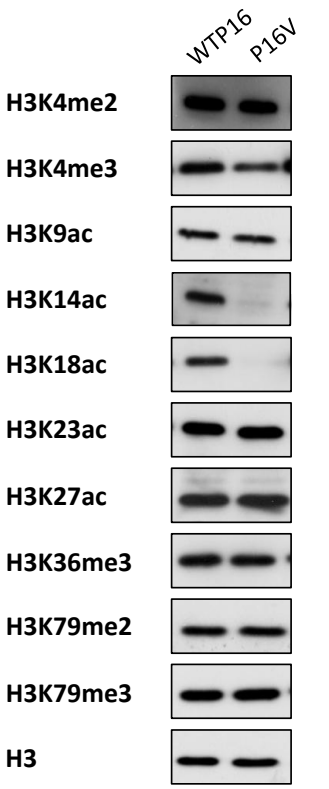


Figure 23. Mutation of H3P16 affects H3K4me3.

(A) A Western blot showing levels of H3K4me2 and H3K4me3 in the wildtype (WT HHT) and P16V strains. Histone H3 levels are shown as a loading control. Two-fold dilutions of each extract were loaded to ensure the experiment was in the linear range of detection. **(B)** A ChIP-qPCR experiment showing the levels of H3K4me3 at *FMP27* in the H3P16 substitution strain relative to the WTP16 strain. Primer positions are indicated on the locus map. Signals are normalised to levels of H3. Results shown are representative of at least three experiments. Error bars show standard errors of the real-time PCR reaction.

Following the observation of the effect of H3P16 substitution on H3K4me3, the global levels of other modifications on the histone H3 tail were also assessed in the P16V strain (Figure 24A). As expected from the epitope effects described above, there was no detectable H3K14ac or H3K18ac in the P16V strain. However none of the other H3 modifications tested was affected by substitution of H3P16. Because of the inability to detect H3K14ac and H3K18ac in the P16V strain by antibody-based methods, mass spectrometry was instead employed to determine whether these modifications were in fact present in the P16V strain. As detailed in Figure 24B, a preliminary mass spectrometry experiment identified several histone H3 tryptic peptides that were acetylated on H3K14 and/or H3K18, indicating that, unlike H3K14, H3P16 is not required for H3K18ac. Further experiments would enable determination of the relative levels of H3K14ac and H3K18ac in the wildtype and P16V strains. However, due to the fact that the P16V mutation is in some of the same tryptic peptides as the modifications and that this mutation may affect the ionisation efficiency and/or the flight of the peptide, for normalisation purposes these peptide mixtures would need to be spiked with known concentrations of heavy-labelled mutated peptides of exactly the same sequence and length as the tryptic peptide of interest (Kettenbach et al. 2011). Since there are multiple H3K14- and H3K18-containing peptides due to the variable histone modifications influencing the efficiency of the tryptic cleavage of the protein, this analysis would be challenging.

A



n = 2

B

H3 residue number	Observed mass (av)	Mr (experimental, av)	Mr (calculated)	ppm (av)	Sequence (in the P16V strain)	Modification	Ion score (av)
9-17	473.2771	944.5396	944.5403	-0.75	R.KSTGGKAVR.K	Ac (K9 or K14)	19
9-17	494.2820	986.5494	986.5509	-2	R.KSTGGKAVR.K	2 Ac (K9 and K14)	73
10-17	409.2287	816.4429	816.4454	-3	K.STGGKAVR.K	Ac (K14)	34
10-18	394.2771	944.5403	944.5403	0	K.STGGKAVRK.Q	Ac (K14 or K18)	11
18-26	451.4957	1013.5990	1013.5981	0.5	R.KQLASKAAR.K	Ac (K18 or K23)	25
18-26	528.8119	1055.6093	1055.6087	1	R.KQLASKAAR.K	2 Ac (K18 and K23)	69

Figure 24. Assessing the effect of H3P16 on other histone H3 modifications.

(A) Western blots (performed by M. Sale) showing the levels of the indicated histone H3 post-translational modifications in the wildtype (WTP16) and P16V strains. Histone H3 levels are shown as a loading control. **(B)** Mass spectrometry data showing the presence of H3K14ac and H3K18ac in the P16V strain. Each row displays a separate histone H3 tryptic peptide. The corresponding H3 residue number, average observed mass, average experimental molecular weight (Mr), calculated Mr, average parts per million (ppm, difference between experimental and calculated mass), peptide sequence, modification and average ion score are shown for each peptide.

4.2.3. H3P16 also has gene-specific effects on H3K4me3

The crosstalk between H3P16 and H3K4me3 was investigated further by looking for gene-specific effects by ChIP (Figure 25A). As with the H3K14-H3K4me3 crosstalk, mutation of H3P16 decreased H3K4me3 to varying extents at different genes. A genome-wide ChIP-sequencing experiment was performed to expand this analysis and signals were normalised as described for the K14A strain in the previous chapter and the Materials and Methods. Genes were ranked (from largest to smallest) according the ratio of H3K4me3 in the P16V strain relative to the WTH3 strain and the top and bottom 200 genes were used for further analysis (Figure 25B) and are listed in Table 27 and Table 28 (Appendix).

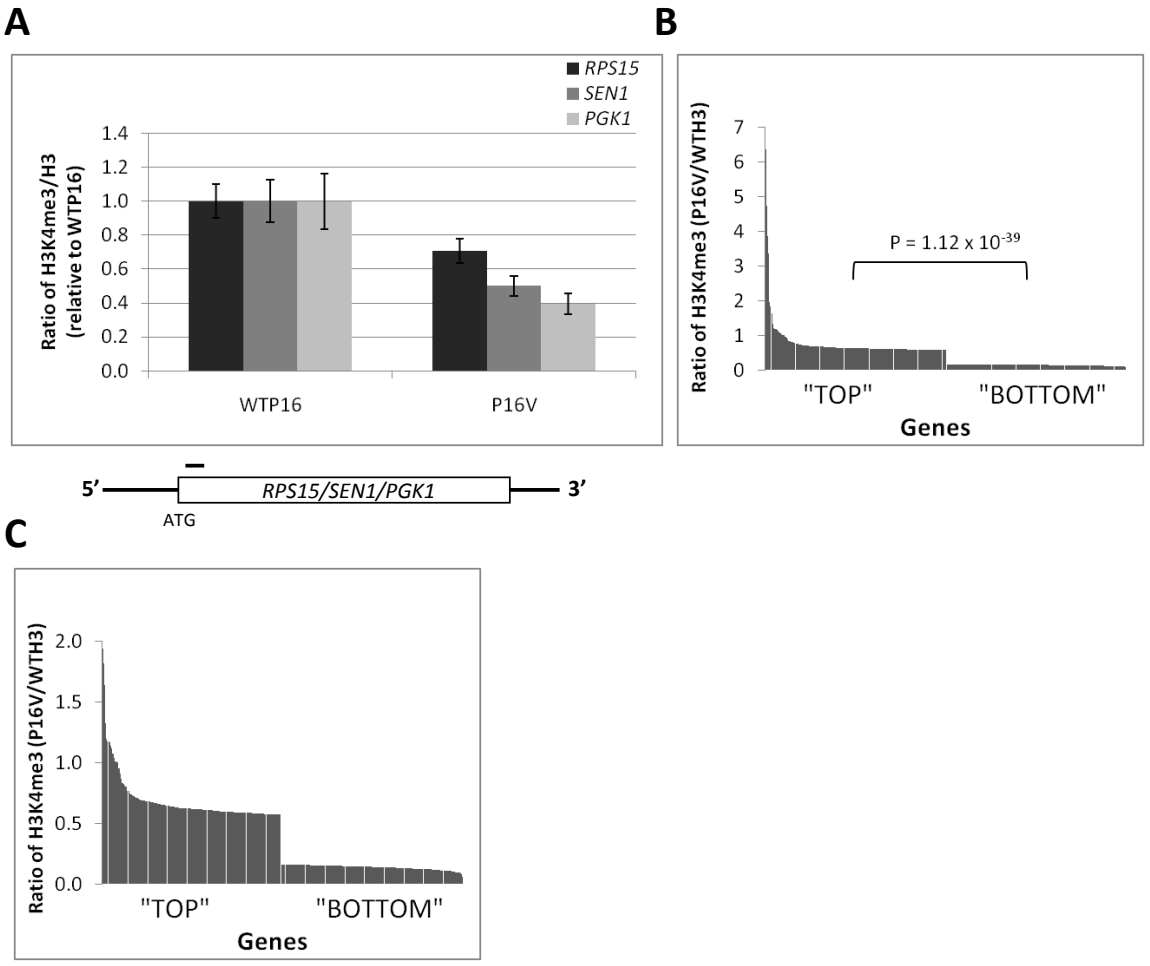


Figure 25. Mutation of H3P16 affects H3K4me3 to different extents at individual genes.

(A) A ChIP-qPCR experiment showing the levels of H3K4me3 at the 5' end of three genes (*RPS15, SEN1* and *PGK1*) in the H3P16 substitution strain relative to the WTP16 strain. Primer positions are indicated on the locus map. Signals are normalised to levels of H3. Error bars show standard errors of the real-time PCR

reaction. **(B)** The levels of H3K4me3 in the P16V relative to the WTH3 strain at the top and bottom 200 genes after ranking according to the ratio of the average H3K4me3 across the transcription unit (P16V/WTH3) as measured by ChIP-sequencing. Signals are not normalised to levels of H3. P values are derived from student's t-test analysis. **(C)** As (B) after removal of the genes with the four highest H3K4me3 ratios to facilitate visualisation of the ratios of the 'bottom' class of genes.

4.2.4. Mutation of H3P16 differentially affects H3K4me3 on genes with distinct cellular functions

The GO terms for the genes in the 'top' and 'bottom' classes of P16V-affected genes are listed in Table 10 and Table 11. As with the K14A experiment, the predominant GO terms in the 'top' class of P16V-affected genes are those associated with translation and ribosome biogenesis. In contrast, the GO terms for the genes in the 'bottom' class of P16V-affected genes are less similar to the K14A 'bottom' class terms, featuring sporulation and meiosis, in addition to some of the stress response pathways in the equivalent K14A list.

Biological process	**Count**	**%**	**P value**
translation	39	19.8	2.90E-10
ribosome biogenesis	21	10.7	1.10E-05
regulation of cellular protein metabolic process	15	7.6	3.20E-05
regulation of translation	14	7.1	4.30E-05
ribonucleoprotein complex biogenesis	21	10.7	6.80E-05
posttranscriptional regulation of gene expression	14	7.1	8.40E-05
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	9	4.6	2.00E-04
rRNA export from nucleus	6	3	1.30E-03
rRNA processing	13	6.6	2.30E-03
RNA export from nucleus	7	3.6	9.60E-03
intracellular transport	21	10.7	1.30E-02
ncRNA processing	14	7.1	1.40E-02
RNA localisation	8	4.1	1.40E-02
ncRNA metabolic process	15	7.6	2.10E-02
establishment of RNA localisation	7	3.6	2.20E-02
RNA transport	7	3.6	2.20E-02
nucleocytoplasmic transport	8	4.1	2.60E-02
nuclear export	7	3.6	3.50E-02
RNA processing	17	8.6	4.30E-02
ribosome assembly	5	2.5	4.60E-02
nucleobase, nucleoside, nucleotide and nucleic acid transport	7	3.6	4.80E-02
negative regulation of cellular protein metabolic process	3	1.5	4.80E-02

Table 10. Gene ontology terms associated with the genes with the top 200 H3K4me3 ratios (P16V/WTH3).

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the top 200 ratios of H3K4me3 (P16V/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

Biological process	Count	%	P value
sporulation	23	11.7	6.50E-10
ascospore formation	16	8.2	1.00E-08
reproductive developmental process	17	8.7	9.20E-08
reproductive process in single-celled organism	17	8.7	7.10E-07
sexual reproduction	19	9.7	1.70E-05
spore wall biogenesis	7	3.6	3.00E-04
cell wall assembly	7	3.6	3.40E-04
response to toxin	7	3.6	3.80E-04
M phase of meiotic cell cycle	13	6.6	4.90E-04
response to temperature stimulus	14	7.1	6.40E-04
thiamine and derivative biosynthetic process	5	2.6	8.90E-04
oxidation reduction	18	9.2	1.10E-03
thiamine and derivative metabolic process	5	2.6	1.30E-03
cellular response to heat	12	6.1	1.50E-03
chromosome organisation involved in meiosis	4	2	3.60E-03
synapsis	4	2	3.60E-03
response to abiotic stimulus	16	8.2	5.90E-03
monosaccharide metabolic process	9	4.6	8.10E-03
water-soluble vitamin biosynthetic process	6	3.1	8.70E-03
aromatic compound biosynthetic process	5	2.6	9.90E-03
cell wall biogenesis	7	3.6	1.10E-02
carbohydrate transport	5	2.6	1.20E-02
hexose metabolic process	8	4.1	1.50E-02
sulphur compound biosynthetic process	6	3.1	1.60E-02
vitamin B6 biosynthetic process	3	1.5	1.80E-02
sulphur metabolic process	7	3.6	1.90E-02
cellular aldehyde metabolic process	4	2	2.10E-02
cell wall organisation	11	5.6	4.10E-02
transmembrane transport	11	5.6	4.10E-02
propionate catabolic process, 2-methylcitrate cycle	2	1	4.20E-02

Table 11. Gene ontology terms associated with the genes with the bottom 200 H3K4me3 ratios (P16V/WTH3).

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the bottom 200 ratios of H3K4me3 (P16V/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

As with the analysis of the genes with H3K4me3 most and least affected by mutation of H3K14, genes in the 'top' and 'bottom' classes of H3P16-affected genes were compared with those whose expression was altered during the yeast metabolic cycle and the environmental stress response. Unlike the K14A 'top' class of genes (Chapter 3), there was no significant bias towards the oxidative phase expression of genes with the top 200 ratios of H3K4me3 (P16V/WTH3) (Table 12). However, genes in the P16V 'bottom' class did significantly overlap with genes expressed during the reductive/charging phase of the YMC, as for the K14A 'bottom' class of genes.

	Expression phase in the yeast metabolic cycle		
	Oxidative	**Reductive/building**	**Reductive/charging**
All genes in the Tu et al. study (n=6575)	1017 (15.5 %)	975 (14.8 %)	1509 (23.0 %)
Genes with top 200 ratios of H3K4me3 (P16V/WTH3)	37 (18.5 %)	13 (6.5 %) $P = 1.4 \times 10^{-4}$	23 (11.5 %) $P = 4 \times 10^{-5}$
Genes with bottom 200 ratios of H3K4me3 (P16V/WTH3)	17 (8.5 %) $P = 1.4 \times 10^{-3}$	17 (8.5 %) $P = 3.2 \times 10^{-3}$	**57 (28.5 %)** $P = 1.6 \times 10^{-2}$

Table 12. Overlap between the genes with the top and bottom 200 ratios of H3K4me3 (P16V/WTH3) and the genes temporally expressed during the yeast metabolic cycle.

Yeast metabolic cycle expression data taken from Tu et al. (2005). The percentages of genes in the 'top' and 'bottom' classes of genes overlapping with each phase are shown in brackets. Bold percentages signify enrichment of genes in that class for a particular feature compared to the rest of the genome. P values were calculated from a simulated probability distribution curve. The absence of P values denotes non-significant results ($P > 5 \times 10^{-2}$).

As in the K14A strain, genes in the 'top' and 'bottom' classes of P16V-affected genes overlapped significantly with those repressed and induced during the environmental stress response respectively (Table 13). Furthermore, the stratification according to SAGA/TFIID-regulated genes again yielded a significant bias towards genes in the 'bottom' class of P16V-affected genes to be regulated by SAGA (Table 14).

	Response to environmental stress	
	Induced	**Repressed**
All genes in the Gasch et al. study (n=6152)	278 (4.5 %)	605 (9.8 %)
Genes with top 200 ratios of H3K4me3 (P16V/WTH3)	0 (0.0 %) $P < 1 \times 10^{-5}$	**56 (28.0 %)** $P < 1 \times 10^{-5}$
Genes with bottom 200 ratios of H3K4me3 (P16V/WTH3)	**22 (11.0 %)** $P = 2.0 \times 10^{-5}$	1 (0.5 %) $P < 1 \times 10^{-5}$

Table 13. Overlap between the genes with the top and bottom 200 ratios of H3K4me3 (P16V/WTH3) and the genes induced and repressed during the common environmental stress response.

Environmental stress response expression data taken from Gasch et al. (2000). The percentages of genes in the 'top' and 'bottom' classes of genes overlapping with each phase are shown in brackets. Bold percentages signify enrichment of genes in that class for a particular feature compared to the rest of the genome. P values were calculated from a simulated probability distribution curve. The absence of P values denotes non-significant results ($P > 5 \times 10^{-2}$).

	Transcriptional regulation		
	SAGA	**TFIID**	**SAGA and TFIID**
All genes in the Huisinga and Pugh study (n=6226)	577 (9.3 %)	5130 (82.4 %)	165 (2.7 %)
Genes with top 200 ratios of H3K4me3 (P16V/WTH3)	2 (1.0 %) $P < 1 \times 10^{-5}$	131 (65.5 %) $P = 2.0 \times 10^{-4}$	2 (1.0 %) $P = 4.4 \times 10^{-2}$
Genes with bottom 200 ratios of H3K4me3 (P16V/WTH3)	**28 (14.0 %)** $P = 2.1 \times 10^{-3}$	91 (45.5 %) $P < 1 \times 10^{-5}$	4 (2.0 %) $P < 1 \times 10^{-5}$

Table 14. Overlap between the genes with the top and bottom 200 ratios of H3K4me3 (P16V/WTH3) and the genes regulated by SAGA and/or TFIID.

SAGA/TFIID data taken from Huisinga and Pugh (2004). The percentages of genes in the 'top' and 'bottom' classes of genes overlapping with each phase are shown in brackets. Bold percentages signify enrichment of genes in that class for a particular feature compared to the rest of the genome. P values were calculated from a simulated probability distribution curve. The absence of P values denotes non-significant results ($P > 5 \times 10^{-2}$).

4.2.5. H3K14 and H3P16 regulate H3K4me3 similarly on the same genes

There were therefore several indications that H3K14 and H3P16 affected a similar set of genes in the same ways. Indeed, the Pearson correlation between the ratios of H3K4me3 in the K14A and P16V strains compared to the wildtype strain was 0.843. To investigate the degree of overlap further, analyses were performed with the 'top' and 'bottom' classes of H3K14- and H3P16-affected genes and the results are summarised in Figure 26. As revealed by the high positive correlation between the ratios of H3K4me3 on genes in the K14A and P16V strains, there is a strong overlap between the genes that are similarly affected by mutation of H3K14 and H3P16, with 57.5 % and 45 % of the 200 genes in the 'top' and 'bottom' classes respectively overlapping between the strains. The genes that fall into one of the non-overlapping sub-classes are defined as being dominantly regulated by one of the two histone residues. For example, the genes that have the top 200 ratios of H3K4me3 (K14A/WTH3) that do not feature in the 'top' class for the P16V-affected genes are termed 'K14A-dominant top'. It cannot be said, however, that these genes are not affected by the P16V mutation but it is evident that they are not affected to the same extent as by the K14A mutation. The GO terms associated with the genes in all six 'top' and 'bottom' overlapping and non-overlapping sub-classes were obtained and summarised in Figure

26. Interestingly, the terms associated with genes in the ‘K14A-dominant’ and ‘K14A and P16V’ sub-classes are very similar but the genes in the ‘P16V-dominant’ sub-classes participate in distinct biological processes. The genes in both the ‘top K14A-dominant’ and ‘top K14A and P16V’ sub-classes predominantly encode products that are involved in translation or ribosome biosynthesis but the most significant GO terms for the ‘top P16V-dominant’ sub-class are ‘vesicle-mediated transport’ and ‘negative regulation of protein metabolism’ (Table 15). A similar distinction was observed with the ‘bottom’ classes of genes: genes in the ‘bottom K14A-dominant’ and ‘bottom K14A and P16V’ sub-classes both encode proteins involved in metabolism and response to stress whereas genes in the ‘bottom P16V-dominant’ sub-class encode proteins participating in sporulation and meiosis (Table 16). Furthermore, the ‘P16V-dominant’ genes do not significantly overlap with genes periodically expressed during the yeast metabolic cycle or regulated by SAGA, although the overlap with the genes oppositely regulated during the environmental stress response is preserved. This analysis therefore suggests that, whilst there is a strong overlap between the genes with H3K4me3 affected in the same manner by mutation of H3K14 and H3P16, at a subset of functionally-related genes the P16V mutation has a greater effect on H3K4me3 than the K14A mutation relative to the rest of the genes in the genome. It is interesting that the genes with H3K4me3 most and least affected by the K14A mutation are not functionally distinct from those also affected in the same manner by H3P16, as implied by the highly overlapping functions for the genes in the ‘K14A-dominant’ and ‘K14A and P16V’ sub-classes.

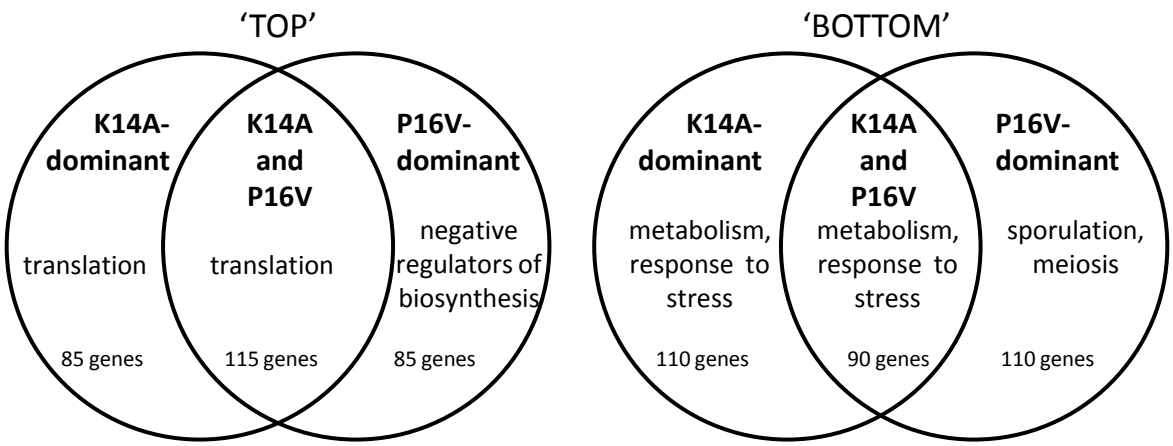


Figure 26. Overlap between genes with H3K4me3 similarly affected by H3K14 and H3P16.

A diagram illustrating the overlap between the K14A and P16V 'top' and 'bottom' classes of 200 genes (not to scale). The 'top' and 'bottom' classes are further divided into 'K14A-dominant', 'K14A and P16V' and 'P16V-dominant'. The number of genes and the predominant gene ontology term(s) for each sub-class are displayed.

Biological process	Count	%	P value
negative regulation of cellular protein metabolic process	3	3.5	1.30E-02
vesicle-mediated transport	9	10.6	2.60E-02
negative regulation of macromolecule metabolic process	7	8.2	3.90E-02
regulation of cellular protein metabolic process	6	7.1	4.30E-02

Table 15. Gene ontology terms associated with the genes with the top 200 H3K4me3 ratios of P16V/WTH3 but not K14A/WTH3.

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the top 200 ratios of H3K4me3 (P16V/WTH3 but not K14A/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

Biological process	Count	%	P value
sporulation	16	15	4.20E-09
ascospore formation	11	10.3	2.10E-07
reproductive developmental process	11	10.3	3.30E-06
reproductive process in single-celled organism	11	10.3	1.20E-05
M phase of meiotic cell cycle	10	9.3	1.20E-04
sexual reproduction	12	11.2	1.20E-04
chromosome organisation involved in meiosis	4	3.7	4.80E-04
synapsis	4	3.7	4.80E-04
spore wall biogenesis	5	4.7	1.10E-03
cell wall assembly	5	4.7	1.20E-03
M phase	10	9.3	8.20E-03
meiosis I	5	4.7	1.10E-02
carboxylic acid catabolic process	4	3.7	1.90E-02
propionate catabolic process, 2-methylcitrate cycle	2	1.9	2.10E-02
oxidation reduction	9	8.4	3.20E-02
meiotic sister chromatid cohesion	2	1.9	4.20E-02
cell wall organisation	7	6.5	4.90E-02

Table 16. Gene ontology terms associated with the genes with the bottom 200 H3K4me3 ratios of P16V/WTH3 but not K14A/WTH3.

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the bottom 200 ratios of H3K4me3 (P16V/WTH3 but not K14A/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

4.2.6. Sense and antisense transcription of the P16V-affected genes

The Churchman and Weissman (2011) dataset was used to compare wildtype levels of sense and antisense transcription of the P16V-affected genes. Like the K14A-affected genes, the P16V 'bottom' class genes have lower sense and higher antisense transcription compared to the genes in the P16V 'top' class (Figure 27A, B). However, although the genes in the 'P16V-dominant' sub-classes have similar levels of sense transcription to the corresponding P16V (all 200) classes, the genes in the 'P16V-dominant top' sub-class are slightly *enriched* for antisense transcription compared to the genes in the 'P16V-dominant bottom' class and the average of all genes in the genome (Figure 27C, D).

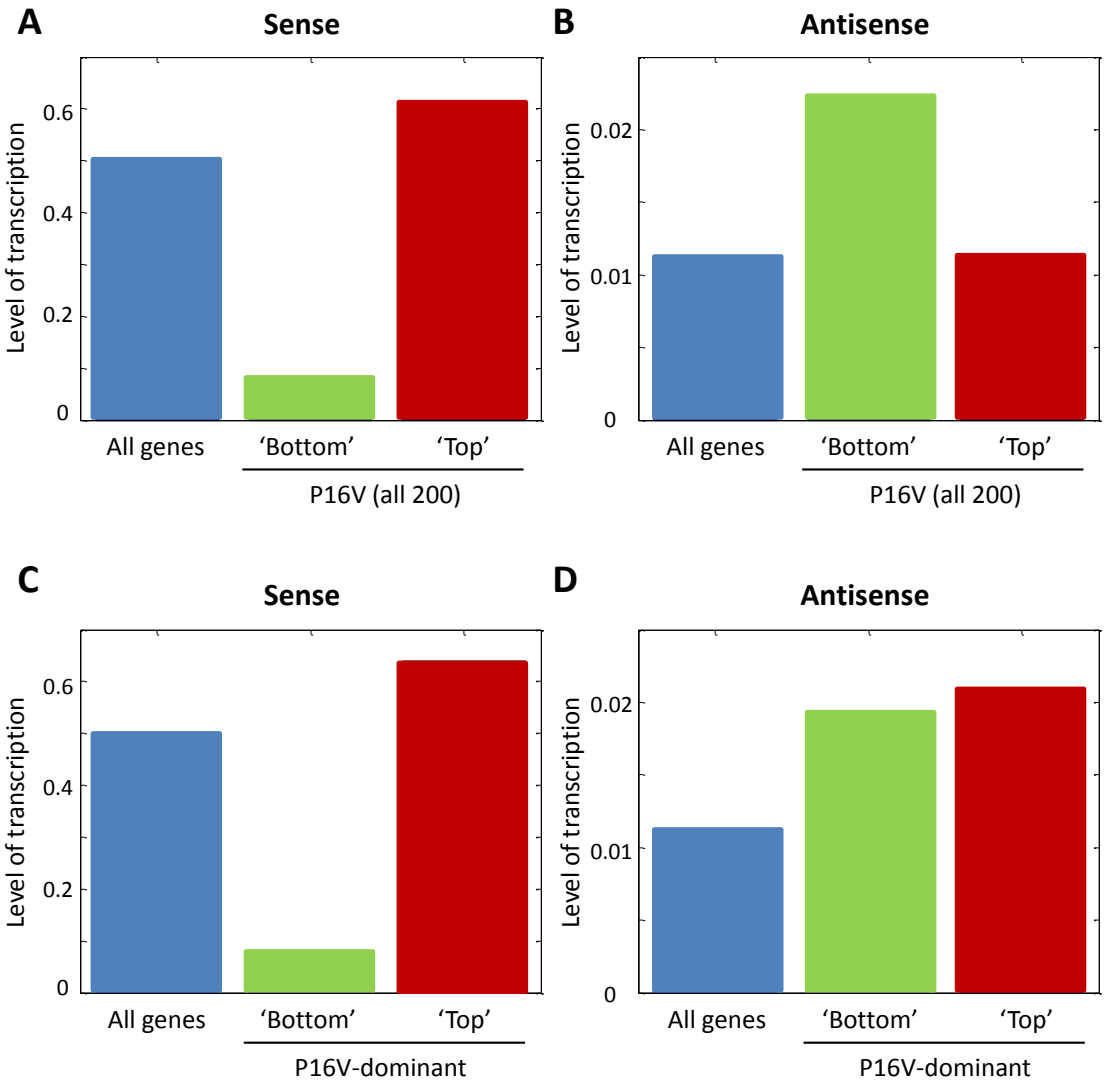


Figure 27. Sense and antisense transcription of P16V-affected genes.

(A) The average wildtype levels of sense transcription across all genes in the genome and the genes in the 'top' and 'bottom' P16V-affected classes using the NET-seq dataset (Churchman and Weissman, 2011) (S. Murray). **(B)** As (A) except for antisense transcription in the first 300 bp of the genes. **(C)-(D)** As (A) and (B) except for the genes in the 'P16V-dominant' sub-class (genes that are not also in the equivalent K14A-affected class).

The Nrd1 levels on RNA produced from the P16V-affected genes were analysed using the Nrd1 PAR CLIP dataset (Creamer et al. 2011). Compared to the average distribution on all 200 genes in the P16V-affected 'top' class, the predominant 3' Nrd1 peak on the 'P16V-dominant top' class of genes is shifted upstream (Figure 28A, B). This lower remaining level of Nrd1 at the extreme 3' end of the 'P16V-dominant top' genes may account for the higher level of antisense transcription reaching the gene promoter in the 'P16V-dominant top' class. As in the previous chapter, the link between Nrd1-dependent transcription termination and H3K4me3 was explored (Terzi et al. 2011). The genes in the 'P16V-dominant top' sub-class had a slightly lower level of H3K4me3 at their 3' ends compared to all 200 genes in the P16V 'top' class (Figure 28C, D). Thus, this slightly lower level of H3K4me3 could contribute to the reduced recruitment of Nrd1 to the 3' ends of the genes in the 'P16V-dominant top' sub-class, resulting in higher levels of antisense transcription into the 5' gene regions. Additionally, there is a second, more 5'-localised peak of H3K4me3 in the 'P16V-dominant top' class, the effect of which is not known. Unfortunately, how the sense and antisense transcription of these genes is affected by the P16V mutation has not yet been identified.

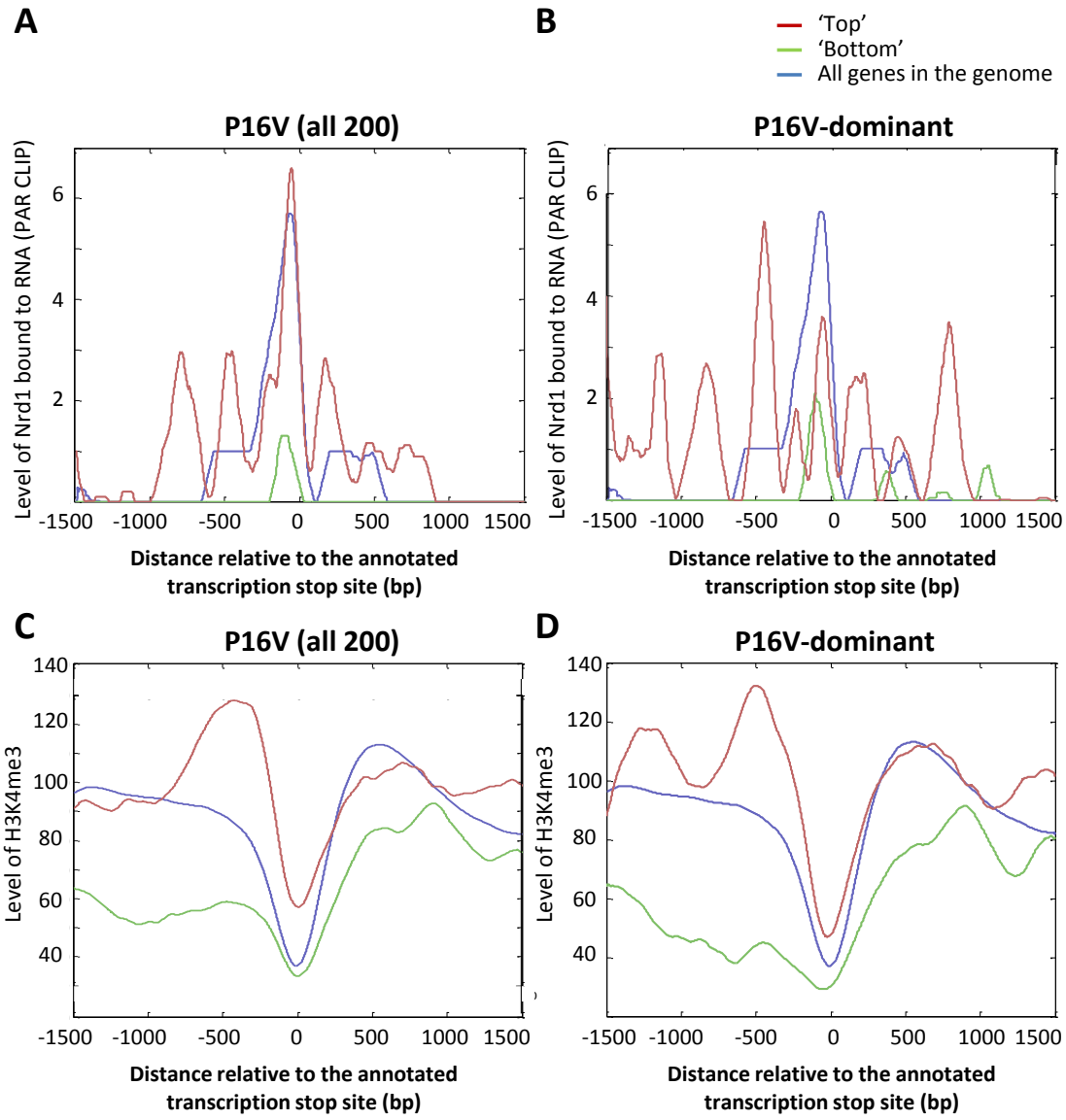


Figure 28. Nrd1 and H3K4me3 distributions at the 3' ends of P16V-affected genes.

(A) The average wildtype distribution of Nrd1 bound to RNA at the 3' ends of all genes in the genome and the 200 genes in the 'top' and 'bottom' P16V-affected classes using the Nrd1 PAR CLIP dataset (Creamer et al. 2011). (S. Murray) **(B)** As (A) except for the genes in the 'P16V-dominant' sub-class (genes that are not also in the equivalent K14A-affected class). **(C)-(D)** As (A) and (B) for the average distribution of H3K4me3 at the 3' ends of all genes in the genome using the WTH3 H3K4me3 ChIP-seq dataset (this study).

4.2.7. Crosstalk between H3K14 and H3P16

To investigate if H3K14 and H3P16 were influencing H3K4me3 through a common pathway, an epistasis experiment was performed: H3K14 H3P16 double mutant strains were created and global levels of H3K4me2 and H3K4me3 were assessed (Figure 29). As expected, the global levels of H3K4me2 did not vary substantially between the double mutant strains tested since mutation of H3K14 or H3P16 alone did not affect this modification. The global levels of H3K4me3 in the K14A P16V and K14Q P16V strains were more similar to the K14A and K14Q strains respectively than the P16V strain whereas the K14R P16V strain had levels of H3K4me3 that were closer to the P16V than the K14R strain. Together these findings imply that there may not be a simple linear pathway between H3K14 and H3P16 controlling H3K4me3. Without further experimentation, it would be difficult to conclude more from these results other than the fact that H3K14 and H3P16 are likely to operate in a branched pathway that converges on overlapping mechanisms controlling H3K4me3. When analysing these results, it should be remembered that not all genes are affected to the same extent by mutation of H3K14 and H3P16 and so the effects on a global level may be difficult to interpret. ChIP experiments could be performed to better assess the small changes in H3K4me3 between the mutant strains and to observe any gene-specific effects of the H3K14-P16 crosstalk on H3K4me3 at individual genes in the top/bottom 'K14A-dominant', 'K14A and P16V' and 'P16V-dominant' classes.

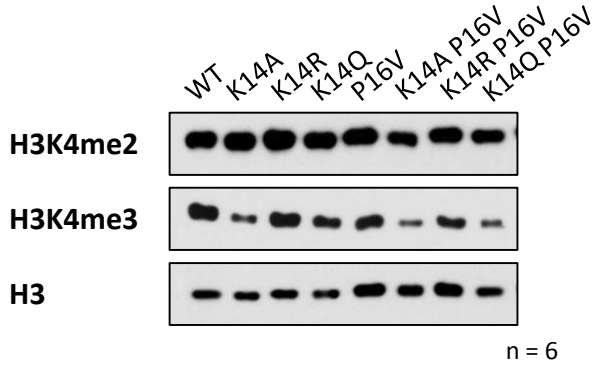


Figure 29. Crosstalk between H3K14 and H3P16.

A Western blot showing levels of H3K4me2 and H3K4me3 in the wildtype (WT HHT), H3K14/H3P16 single and double mutant HHT strains. Histone H3 levels are shown as a loading control.

4.3. Discussion

The field is becoming increasingly aware of the need to consider the effect of local protein structures in previously viewed unstructured flexible protein tails. The bond between proline and its neighbouring residue can undergo *cis-trans* isomerisation, which gives it the potential to dramatically alter the relative trajectories of the flanking regions and thus local protein structure. Several groups working on the RNA polymerase II C-terminal domain (CTD), which has a similar flexible structure to the histone H3 tail, are focusing on the regulation of the isomerisation of the two proline residues in the heptapeptide repeat and how this affects the binding of protein effectors that regulate post-translational modification of the CTD to influence transcriptional processes. The work in this chapter has explored the function of one of the three histone H3 tail proline residues, H3P16, in the control of a relatively distal modification, H3K4me3.

4.3.1. H3P16 is important for proper modification of the histone H3 tail

Mutation of H3P16 to valine caused a decrease in the level of H3K4me3. This effect was specific to the trimethyl state since H3K4me2 was unaffected in the P16V strain. Two assays have shown that mutation of H3P16 to valine can both increase and decrease the binding of proteins to neighbouring modifications *in vitro* on peptides. Therefore, this decrease in H3K4me3 could have been due to an altered binding of a protein that regulates H3K4me3 levels to the histone H3 tail in the P16V mutant. If H3P16 is part of the recognition site for a protein effector, it is easy to imagine how its mutation could directly influence the binding of this protein. However, as well as directly affecting protein binding to influence H3K4me3, H3P16 has the potential to indirectly affect H3K4me3 via structural changes in the histone H3 tail. Although it is impossible to distinguish between the effect of H3P16 substitution and the loss of the ability to isomerise at that residue, it is tempting to speculate that, since H3K4me3 is reduced by the P16V substitution, H3P16 in *cis* is required for optimal H3K4me3. When the peptidyl-prolyl bond is in the *cis* conformation, the H3 tail may adopt more of a 'bent' configuration and this could decrease the

distance in space between the two H3P16-flanking regions. Conversely, with H3P16 peptidyl-prolyl bond in *trans*, the H3 tail would be more 'straight' and the flanking regions would be more distally located. If a region within the C-terminal half of the H3 N-terminal tail forms part of a docking site for the Set1 enzyme/complex that modifies H3K4, a residue that is distant in terms of primary sequence, isomerisation of H3P16 to *cis* could increase the proximity of these two regions and potentially the ability to modify H3K4. It is also possible that H3K4 may require an interaction with the other histone H3 tail in the nucleosome for trimethylation. Proline isomerisation may help position these tails for efficient trimethylation. Experiments investigating a mechanism for the H3P16-H3K4me3 crosstalk will be discussed in the next chapter.

It is unknown how the isomerisation of the H3P16 peptidyl-prolyl bond is regulated. Work in this laboratory using a candidate-based genetic approach failed to identify a peptidyl-prolyl isomerase (PPIase) that acts on H3P16, as measured by changes in H3K4me3 (data not shown). This indirect measure of proline isomerisation relies on the correctness of the above hypotheses and may well be sub-optimal for identification of PPIase activity but it highlights the difficulty of studying proline isomerisation *in vivo*. However, since *cis-trans* isomerisation can occur spontaneously, albeit slowly, in relatively unstructured regions such as the histone H3 tail, it is conceivable that there is no PPIase that acts on this residue. Potentially the slow inter-conversion between *cis* and *trans* proline conformations could act as a molecular timer to limit certain processes for which the optimum protein effector binding surface is one of the two states (Lu et al. 2007). Additionally, it is possible that there may be proteins that specifically recognise and bind to certain conformations of the H3 tail and in so doing, stabilise those conformations.

4.3.2. Mutation of H3P16 differentially affects H3K4me3 on individual genes

The ChIP-sequencing experiment revealed that H3K4me3 was reduced by a greater extent on some genes compared to others in the P16V strain. The gene ontology analysis of the 200 genes

with H3K4me3 most and least affected by the P16V mutation demonstrated that the genes in these two classes are functionally distinct. The genes whose levels of H3K4me3 were least reduced by the P16V mutation are genes involved in translation and ribosome biogenesis whereas the functions of the genes with H3K4me3 most reduced by mutation of H3P16 encode proteins functioning in more disparate processes such as sporulation, some stress responses and carbohydrate transport. However, this latter class is enriched in genes that are expressed in the reductive/charging phase of the yeast metabolic cycle, induced during the common environmental stress response and regulated by the SAGA complex. In contrast, the former class is not significantly enriched for genes expressed temporally during the yeast metabolic cycle but is enriched for genes repressed during the stress response. Because the genes in these two P16V-affected classes are expected to be transcribed under opposite cellular metabolic conditions, it is likely that their modes of regulation are also different. This could explain the relative dependencies on H3P16 for proper H3K4me3.

4.3.3. Both H3K14 and H3P16 crosstalk with H3K4me3 – do they influence each other?

As discussed in Chapter 3, H3K14 can also specifically affect H3K4me3. The high correlation between the levels of H3K4me3 on genes in the K14A and P16V strains indicates that these two residues similarly influence H3K4me3. It is therefore possible that these residues may operate in a common pathway: H3K14, in its modified or unmodified form, may influence the isomerisation state of H3P16 to regulate H3K4me3, and/or H3P16 isomerisation could regulate the modification state of H3K14 to influence H3K4me3. Unfortunately it is not possible to assess the isomerisation state of H3P16 *in vivo* but there is some *in vitro* evidence to suggest that the modification state of H3K14 may contribute to H3P16 isomerisation. Previous work in the laboratory using an indirect *in vitro* assay (Fischer et al. 1984; Nelson et al. 2006) indicated that H3K14ac increases the proportion of H3P16*trans* in peptides. Furthermore the results of the Spt7 bromodomain binding assay may also infer a crosstalk between these two residues. The increased binding of the Spt7

bromodomain to the P16V-K18ac peptide over the K18ac peptide could imply that the Spt7 bromodomain has a preference for binding to H3P16*trans*-K18ac. The abrogation of binding by the K14R substitution and subsequent rescue by the P16V mutation could indicate that the K14R mutation increases the proportion of H3P16*cis*, which is less preferential for binding. However, caution must be taken when interpreting the results with the P16V mutation since the effects of the loss of isomerisation potential cannot be distinguished from the effects of substitution of this residue. Nevertheless, the evidence, albeit indirect, that H3K14ac increases the proportion of H3P16*trans* is consistent with the effect of the K14R substitution, mimicking the unmodified lysine, potentially increasing the proportion of H3P16*cis*. Further studies on the effect of H3K14 modification on the isomerisation state of the peptidyl-prolyl H3A15-H3P16 bond could make use of an NMR-based approach.

Unfortunately the reverse experiments, assessing whether H3P16 can influence the modification state of H3K14, are also difficult to perform due to the epitope effects described with the antibody raised against H3K14ac in the P16V strain. From the crystal structures of the *Tetrahymena* Gcn5 protein in complex with acetyl-CoA and histone H3 peptides of different lengths, it can be predicted that H3P16 would be in the active site of the yeast enzyme when acetylating both H3K14 and H3K18 (Rojas et al. 1999; Clements et al. 2003). Since van der Waals interactions have been identified between the *Tetrahymena* Gcn5 active site residues and H3K14 and H3P16 that are important for H3K14 substrate selectivity (Clements et al. 2003), it is likely that mutation of H3P16 would affect acetylation of H3K14 by Gcn5. A preliminary mass spectrometry experiment indicated that H3K14ac is detectable in the P16V strain and it would be interesting to discover whether the levels of H3K14ac are affected by the P16V mutation.

The epistasis experiment, aimed at elucidating any common mechanism for H3K4me3 regulation by these two residues produced results that were difficult to interpret: the K14A/K14Q H3K4me3 phenotypes were dominant in the K14A/K14Q P16V strains whereas the P16V phenotype was dominant in the K14R P16V strain. The seemingly conflicting results may stem from the different effects of the three substitutions of H3K14 on the regulation of H3K4me3 and potentially the isomerisation state of H3P16. Therefore, in the double mutant strains, the effects of the three H3K14 substitutions and the P16V mutation may combine differently to result in diverse levels of H3K4me3.

Chapter 5.

Investigating a mechanism for the crosstalk between H3K14/H3P16 and H3K4me3

5. Investigating a mechanism for the crosstalk between H3K14/H3P16 and H3K4me3

5.1. Introduction

The mechanism by which H3K14 influences H3K4me3 is unknown. When this crosstalk was originally described, the authors proposed that prior acetylation of H3K14 (H3K14ac) was required for H3K4me3 but did not suggest further mechanistic insight (Nakanishi et al. 2008). The work in this thesis questions the absolute requirement of H3K14ac for H3K4me3 (Chapter 3). However, irrespective of the debate surrounding the requirement for H3K14ac, the potential components of the crosstalk pathway can be examined. The lower levels of H3K4me3 in the H3K14 mutant strains may result from decreased H3K4 methylation, increased H3K4 demethylation or a combination of both. Mutation of H3K14 could directly affect the ability to methylate and/or demethylate H3K4 or could indirectly affect H3K4me3 via other known or unknown crosstalk pathways influencing H3K4me3.

5.1.1. H3K4 methylation is deposited by the Set1 methyltransferase complex

Set1 is the only enzyme that methylates H3K4 in yeast (Briggs et al. 2001; Roguev et al. 2001; Nagy et al. 2002; Krogan et al. 2002) and operates as part of a complex of proteins termed COMPASS (Miller et al. 2001; Roguev et al. 2001). Individual COMPASS subunit deletions have identified a function for each component in the regulation of the state of H3K4 methylation. Swd1 and Swd3 form a heterodimer (Murton et al. 2010) and are both required for the integrity of the COMPASS complex (Dehé et al. 2006). Therefore, in the absence of either of these subunits, all H3K4 methylation is lost (Morillon et al. 2005; Schneider et al. 2005). Sdc1 and Bre2 also form a heterodimer (South et al. 2010) and are required for H3K4me3 and proper H3K4me2 (Morillon et al. 2005; Schneider et al. 2005; Dehé et al. 2006). Swd2 is an essential protein and is found in both

the COMPASS and the cleavage and polyadenylation factor (CPF) complexes. For these reasons, its effects on H3K4 methylation have been difficult to determine but a role for Swd2 has been described in the coupling of H2BK123 monoubiquitination and H3K4me3 (Lee et al. 2007; Vitaliano-Prunier et al. 2008).

Of most relevance to this work, Spp1, like H3K14 and H3P16, is specifically required for proper H3K4me3, with only small reductions in H3K4me2 upon *SPP1* deletion (Morillon et al. 2005; Schneider et al. 2005). The mechanism by which Spp1 promotes H3K4me3 is unclear. It has been suggested that Spp1 may influence the catalytic activity of Set1 to favour trimethylation through the tyrosine switch model (Takahashi et al. 2009). However, Spp1 is not required as part of a reconstituted COMPASS complex for H3K4me3 on free histone H3 *in vitro* (Takahashi et al. 2011) and residual amounts of H3K4me3 remain in the absence of *SPP1 in vivo* (Schneider et al. 2005; Takahashi et al. 2009; Mersman et al. 2011). These findings do not support an absolute requirement of Spp1 for Set1-mediated H3K4 trimethylation but perhaps indicate that Spp1 acts to increase the efficiency of Set1 trimethylase activity in the chromatin environment.

Spp1 has a PHD finger near its N-terminus. PHD fingers are characterised methyl-binding domains and are frequently found in chromatin-associated proteins (Taverna et al. 2007). The Spp1 PHD finger interacts with H3 peptides methylated on H3K4, but not on H3K36 or H3K79, *in vitro* (Shi et al. 2007). Furthermore, the Spp1 PHD finger interacts more strongly with H3K4me2/3 than H3K4me1 in this assay. This interaction has led to speculation that Spp1 may promote H3K4me3 either by binding H3K4me2 and recruiting or stabilising the COMPASS complex on the histone H3 tail to trimethylate H3K4 or by protecting H3K4me2/3 from demethylation by Jhd2 (Kirmizis et al. 2007). Alternatively, Spp1 may bind H3K4me2/3 on one histone H3 tail and recruit COMPASS to allow methylation of lysine 4 on the other histone H3 tail in the same nucleosome.

5.1.2. H3K4 methylation is removed by the Jhd2 demethylase

The JmjC domain-containing protein Jhd2 is responsible for demethylation of H3K4 (Ingvarsdottir et al. 2007; Liang et al. 2007; Seward et al. 2007; Tu, S et al. 2007) and can remove all three H3K4 methylation states *in vivo* (Huang et al. 2010). The regulation of Jhd2 activity is unknown: Jhd2 functions as a monomeric protein and does not purify with any stably-associated protein partners (Liang et al. 2007). However, it is known that Jhd2 protein levels are tightly controlled *in vivo* by Not4-dependent polyubiquitination and subsequent proteasomal degradation (Laribee et al. 2007; Mersman et al. 2009). Like Spp1, Jhd2 also has a PHD domain, although unlike Spp1, the Jhd2 PHD finger does not interact with peptides methylated on H3K4, H3K36 or H3K79 *in vitro* (Shi et al. 2007). While the Jhd2 PHD domain is required for Jhd2 association with chromatin *in vivo*, this interaction does not require the first 28 residues of the histone H3 tail (Huang et al. 2010). The same study also showed that the Jhd2 PHD finger is also necessary for demethylase activity *in vivo*, perhaps due to its role in positioning the enzyme correctly for demethylation.

5.1.3. A number of residues participate in crosstalk with H3K4

As mentioned above, H3K14 may influence H3K4me3 indirectly by affecting levels of other histone modifications that regulate H3K4me3 levels. Two modifications known to affect H3K4me3 are asymmetric dimethylation of H3R2 and monoubiquitination of H2BK123. H3R2me2a prevents the binding of Spp1 to H3K4me2 within peptides *in vitro* and in chromatin *in vivo* (Kirmizis et al. 2007). This is thought to be responsible for the inhibition of H3K4me3 by H3R2me2a in yeast.

H2BK123 monoubiquitination is required for H3K4me2/3 (Dover et al. 2002; Sun and Allis 2002) but not H3K4me1 (Dehé et al. 2005; Shahbazian et al. 2005). The mechanism by which this crosstalk operates is not clear. Two groups have documented the involvement of the COMPASS subunit Swd2, although there is some debate as to the exact role of this protein in the crosstalk (Lee et al. 2007; Vitaliano-Prunier et al. 2008). An alternative hypothesis was proposed by

Chandrasekharan et al. (2010). The authors identified a helix in the H2B C-terminal tail that is required for the docking of COMPASS and subsequent H3K4 trimethylation but not H2BK123 monoubiquitination. They suggested that the function of H2Bub is to stabilise the nucleosome (Chandrasekharan et al. 2009) and so preserve the COMPASS docking site on H2B to promote H3K4me3. The idea of multiple residues from different histones forming a docking site for methyltransferases has also been proposed for Set2, which requires residues from histones H3, H4 and H2A for proper H3K36 methylation (Du et al. 2008; Du and Briggs 2010; Hainer and Martens 2011; Endo et al. 2012) and Dot1, which requires the interaction with H2Bub and a basic patch on histone H4 for proper H3K79 methylation (Fingerman et al. 2007; Oh et al. 2010).

5.1.4. Aims

The work in this chapter aims to identify components that participate in the crosstalk between H3K14/H3P16 and H3K4me3. The relative activities of the Jhd2 demethylase and the Set1 methyltransferase (in concert with Spp1) in the wildtype and H3K14 mutant strains will be examined. The potential for the involvement of other crosstalk pathways in the H3K14-H3K4me3 crosstalk will also be explored.

5.2. Results

5.2.1. Regulation of H3K4me3 by opposing methylation and demethylation

As detailed in the introduction to this chapter, H3K4 methylation is deposited by the Set1 methyltransferase within the COMPASS complex, with the individual COMPASS subunits influencing the methylation state. As demonstrated by Morillon et al. (2005) and Schneider et al. (2005), and confirmed in Figure 30A, all the COMPASS subunits are required for proper H3K4me3, although a residual level of H3K4me3 remains in the absence of *SPP1*. In contrast, near-wildtype levels of H3K4me2 were found in the *spp1Δ* strain. Next it was investigated whether additional deletion of the H3K4 demethylase, *JHD2*, in the COMPASS deletion strains could rescue H3K4me3 (Figure 30B). Only a very slight increase in H3K4me3 was observed in the absence of *JHD2,* in agreement with other studies (Ingvarsdottir et al. 2007; Liang et al. 2007; Tu, S et al. 2007). The small decrease in H3K4me2 in the *jhd2Δ* strain may be due to the absence of H3K4me3 demethylation to produce H3K4me2. Deletion of *JHD2* failed to recover H3K4me2 or H3K4me3 in the *bre2Δ, sdc1Δ, set1Δ, swd1Δ* or *swd3Δ* strains, indicating that these subunits are necessary for H3K4me3 deposition. Whilst there was no increase in H3K4me2 in the *spp1Δ jhd2Δ* strain compared to the *spp1Δ* strain, a longer exposure of the Western blot to film shows a substantial increase in global H3K4me3 upon deletion of *JHD2* in the *spp1Δ* strain (lanes 9 and 10). This confirms the *in vitro* result that H3K4me3 can be deposited by COMPASS in the absence of *SPP1* (Takahashi et al. 2011) and supports a role for Spp1 in increasing the efficiency of Set1-mediated trimethylation of H3K4 rather than being absolutely required for H3K4me3 deposition. Additionally, this result demonstrates that there is a balance between the actions of methylation by Set1 and demethylation by Jhd2 that controls the levels of H3K4me3 in the wildtype strain (Figure 30C).

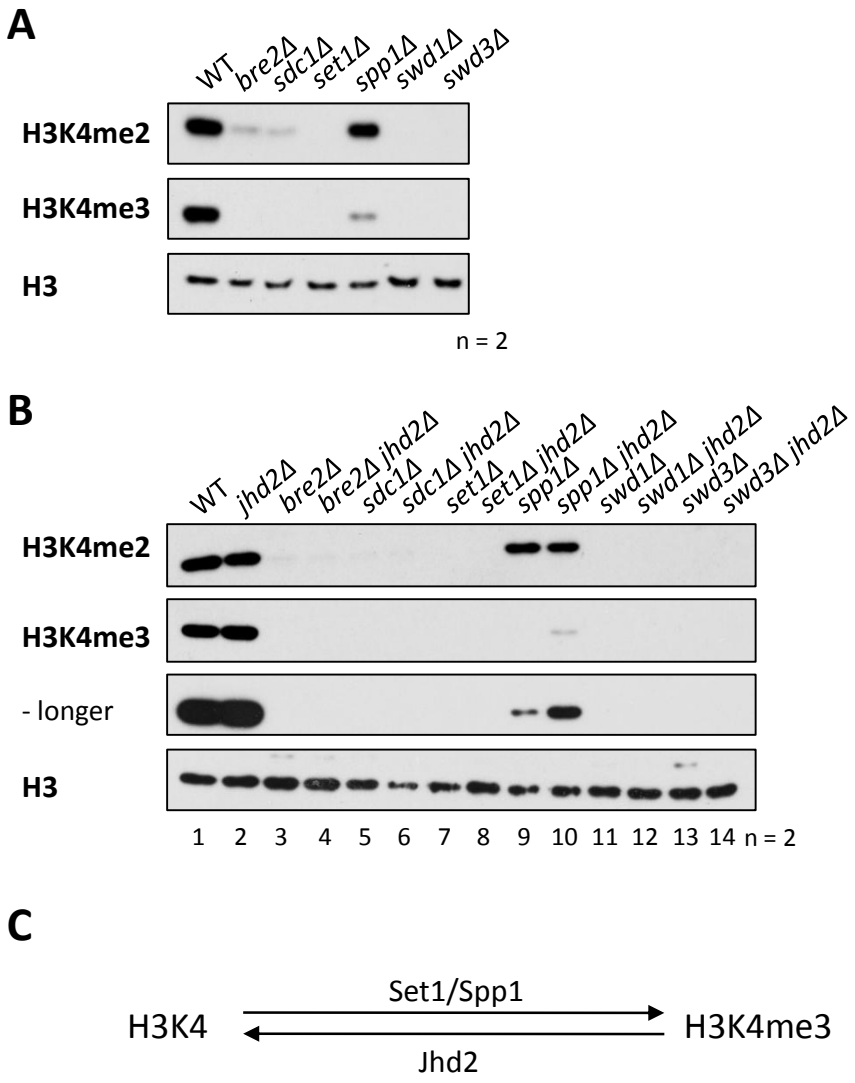


Figure 30. Control of H3K4me3 by COMPASS and Jhd2.

Western blots (performed by A. Nair) showing the levels of H3K4me2 and H3K4me3 in the wildtype (BY4741) and **(A)** the individual COMPASS subunit deletion strains and **(B)** the COMPASS deletion strains in the presence or absence of *JHD2.* A longer exposure of H3K4me3 is shown to enable visualisation of the H3K4me3 signals from the *spp1Δ* and *spp1Δ jhd2Δ* strains. Histone H3 levels are included as a loading control. **(C)** A diagram illustrating the balance between methylation by Set1 (with Spp1) and demethylation by Jhd2 in the control of H3K4me3 in the wildtype strain.

To add to this idea of a balance between methylation and demethylation, the endogenous *JHD2* and *SPP1* genes were placed under the control of the *GAL1* promoter. This system allowed control over the expression of *JHD2* and *SPP1* by altering the carbon source in the media. When these strains were grown in galactose-containing media, the overexpression of *JHD2* and *SPP1* was induced and the resultant Jhd2 and Spp1 proteins far exceeded their respective wildtype levels

(data not shown). (As an aside, it is interesting to note that this method of *JHD2* overexpression was able to overcome the normal proteasomal regulation of Jhd2 protein levels.) Strains with every combination of *JHD2* and *SPP1* deletion and overexpression were constructed and assessed for global levels of H3K4me2 and H3K4me3 (Figure 31A). The results with the single and double deletions of *JHD2* and *SPP1* (lanes 1-4) repeated the observations in Figure 30B. Overexpression of *JHD2* (lane 5), as indicated by the detection of the N-terminal HA tag on Jhd2, did not affect H3K4me2 but caused a substantial drop in the levels of H3K4me3, in agreement with Huang et al. (2010). The authors proposed that the lack of effect of *JHD2* overexpression on H3K4me2 was due to the demethylation of H3K4me3 to H3K4me2, resulting in an increase in H3K4me2, cancelling out the decrease in H3K4me2 caused by its demethylation to H3K4me1. In contrast, overexpression of *SPP1* (lane 6) did not affect H3K4me2 or H3K4me3, possibly due to control by or limiting levels/activity of other COMPASS components. (A lack of increase in H3K4me3 was also observed upon overexpression of *SET1*, either alone or in combination with overexpression of *SPP1* (data not shown).) Surprisingly, both H3K4me2 and H3K4me3 levels were reproducibly lower in the double *JHD2* and *SPP1* overexpression strain (lane 7) when compared to either of the single overexpression strains (lanes 5 and 6), implying that the greater cellular level of Spp1 somehow sensitises H3K4me2/3 to demethylation by Jhd2. As expected, there was undetectable H3K4me3 in the *pGAL1-3HA-JHD2 spp1*Δ strain (lane 8) but interestingly, H3K4me2 was also lowered in this strain when compared to either of the single mutants (lanes 3 and 5). The fact that overexpression of *JHD2* did not affect H3K4me2 except in the strain in which *SPP1* had also been deleted agrees with the 'masking' hypothesis put forward by Huang et al. In the absence of *SPP1*, there is less H3K4me3 to demethylate to H3K4me2 upon *JHD2* overexpression. Therefore this smaller increase in H3K4me2 can no longer cancel out the decrease in H3K4me2 upon its demethylation to H3K4me1, which consequently leads to an overall decrease in global H3K4me2 levels. Alternatively, it is tempting to suggest that Spp1, potentially through its PHD domain, may function directly to protect H3K4me2 from demethylation by Jhd2. Deletion of *JHD2* in the *SPP1*

overexpression strain only slightly increased/did not affect H3K4me2 or H3K4me3 levels (lane 9), as predicted from the near-wildtype H3K4 methylation in the single mutants (lanes 2 and 6).

The function of Spp1 in the protection from demethylation by Jhd2 was explored further. Spp1 has a PHD finger and speculation has been made about whether this may bind to and shield H3K4 methylation from Jhd2 (Kirmizis et al. 2007; Shi et al. 2007), an idea that would be supported by the result in the *pGAL1-3HA-JHD2 spp1Δ* strain. The observation of decreased H3K4me2/3 in the double *SPP1* and *JHD2* overexpression strains confuses this hypothesis but indicates that both too little and too much Spp1 protein sensitises H3K4me3 to demethylation by Jhd2. The Spp1 PHD finger is located near the N-terminus of the protein. To assess whether this domain is important for exacerbating the effects of *JHD2* overexpression on H3K4me3, a crude experiment was performed in which a version of Spp1 lacking the first 76 residues (Spp1ΔN, therefore deleting the PHD finger) was overexpressed in strains with wildtype, deleted or overexpressed *JHD2* (Figure 31B). Truncation of Spp1 did not appear to rescue the reduction in H3K4me3 caused by the combined overexpression of *SPP1* and *JHD2*. However, Spp1ΔN was less stable than the wildtype protein, as indicated by the lower levels of HA signal on the Western blot (compare lanes 7-9 with lanes 4-6), which may be responsible for the lower H3K4me3 levels in all the Spp1ΔN strains. Therefore no conclusions can be made about the function of this domain in the double *JHD2/SPP1* overexpression strains. The use of Spp1 PHD finger point mutants may remove these stability issues in future experiments. Whilst the exact functions of Spp1 in H3K4 methylation remain uncertain, the work in this chapter so far clearly demonstrates the importance of striking the correct balance between methylation and demethylation for the establishment of proper H3K4 methylation.

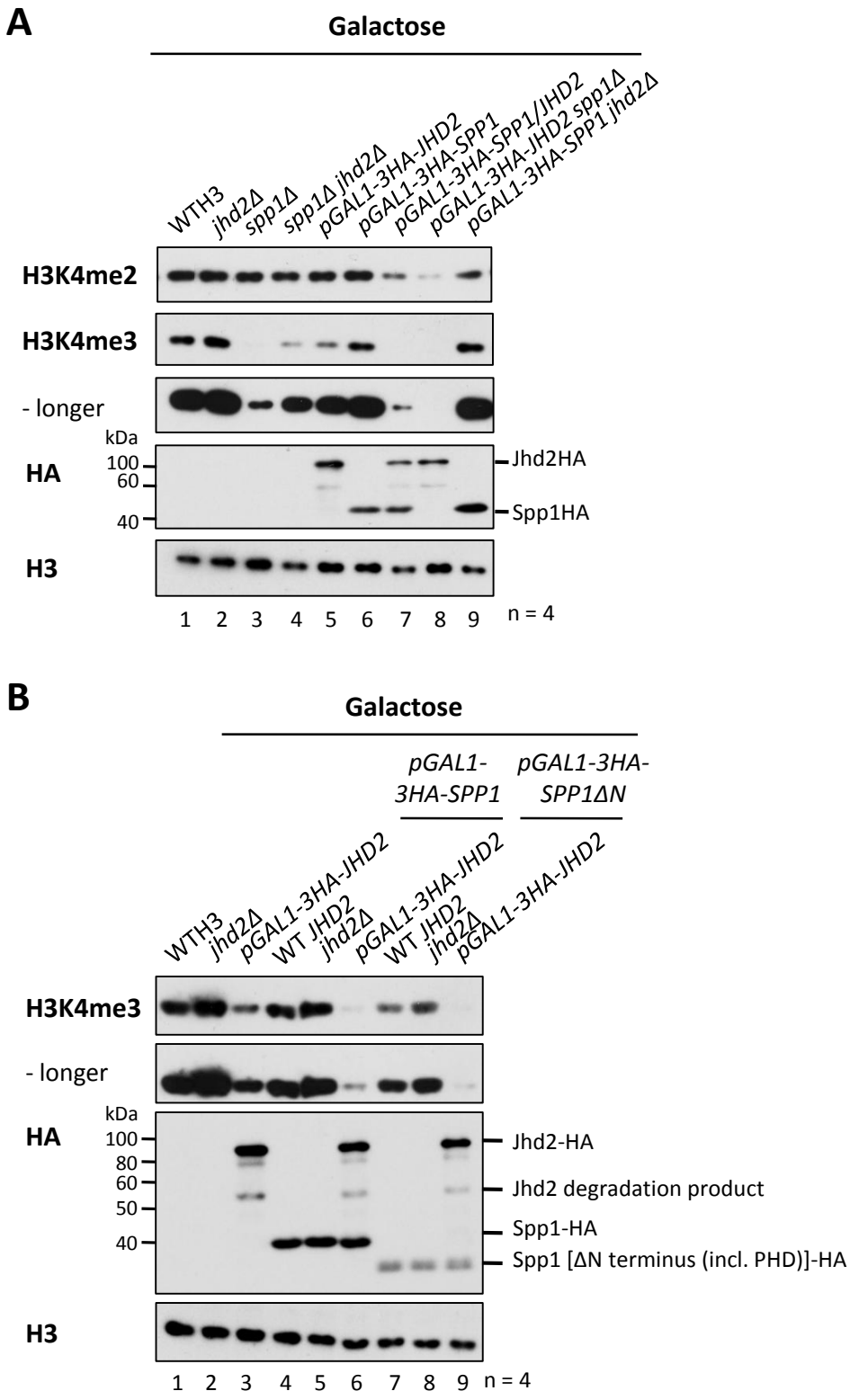


Figure 31. The balance between methylation and demethylation controls H3K4me3 levels.

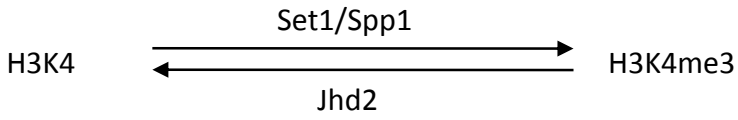
(A) A Western blot showing the levels of H3K4me2, H3K4me3 and HA in the wildtype strain (WTH3) and strains with every combination of *JHD2* and *SPP1* deletion and overexpression (using the *pGAL1-3HA* system). *pGAL1-3HA-SPP1/JHD2* denotes overexpression of *SPP1* and *JHD2* in the same strain. All strains were grown in galactose-containing media to induce overexpression in the *pGAL1-3HA* strains. A longer exposure of H3K4me3 is shown to enable visualisation of the low levels of H3K4me3 in the *spp1Δ* and

pGAL1-3HA-SPP1/JHD2 strains. Histone H3 levels are included as a loading control. **(B)** As (A) except with the strains indicated. Spp1ΔN is lacking the first 76 amino acids, including the PHD finger (residues 20-76).

5.2.2. Investigating the balance between H3K4 methylation and demethylation in the H3K14 mutants

One way in which H3K14 may influence H3K4me3 is to control the balance between methylation and demethylation so that in the H3K14 mutants, this balance is shifted more in favour of demethylation over methylation. This idea is depicted in Figure 32. The data presented in Figure 30 and Figure 31, showing the higher sensitivity of H3K4me3 than H3K4me2 to alterations in Jhd2 and Spp1 levels, agrees with the H3K14 trimethyl-specific effect. The sensitivity of H3K4me3 to changes in Jhd2 and Spp1 was therefore investigated in the H3K14 substitution strains.

WTH3 strain:



H3K14 mutant strains:

Figure 32. H3K14 may influence the balance between H3K4 methylation and demethylation.

A diagram illustrating the idea that H3K14 may be required to maintain the correct balance between H3K4 methylation by Set1/Spp1 and demethylation by Jhd2 and that in the H3K14 mutants, H3K4 demethylation is favoured over methylation.

5.2.3. Investigating the activity of Jhd2 in the H3K14 mutant strains

It is possible that there is less H3K4me3 in the H3K14 mutant strains simply as a result of higher Jhd2 protein levels in these strains. To examine this hypothesis, the DNA sequence for an HA tag was added to the 3' end of endogenous *JHD2* gene, enabling monitoring of Jhd2 protein levels. However, this explanation was ruled out by the finding that there were wildtype levels of Jhd2 in the H3K14 substitution strains (Figure 33).

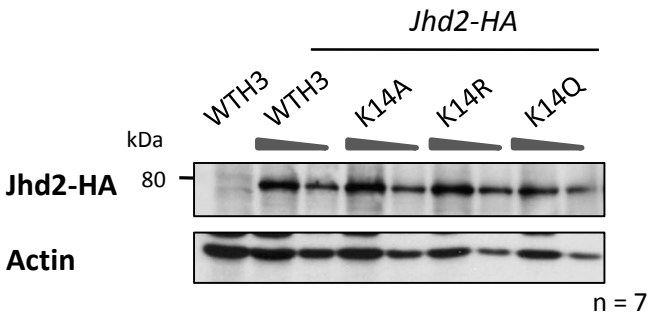


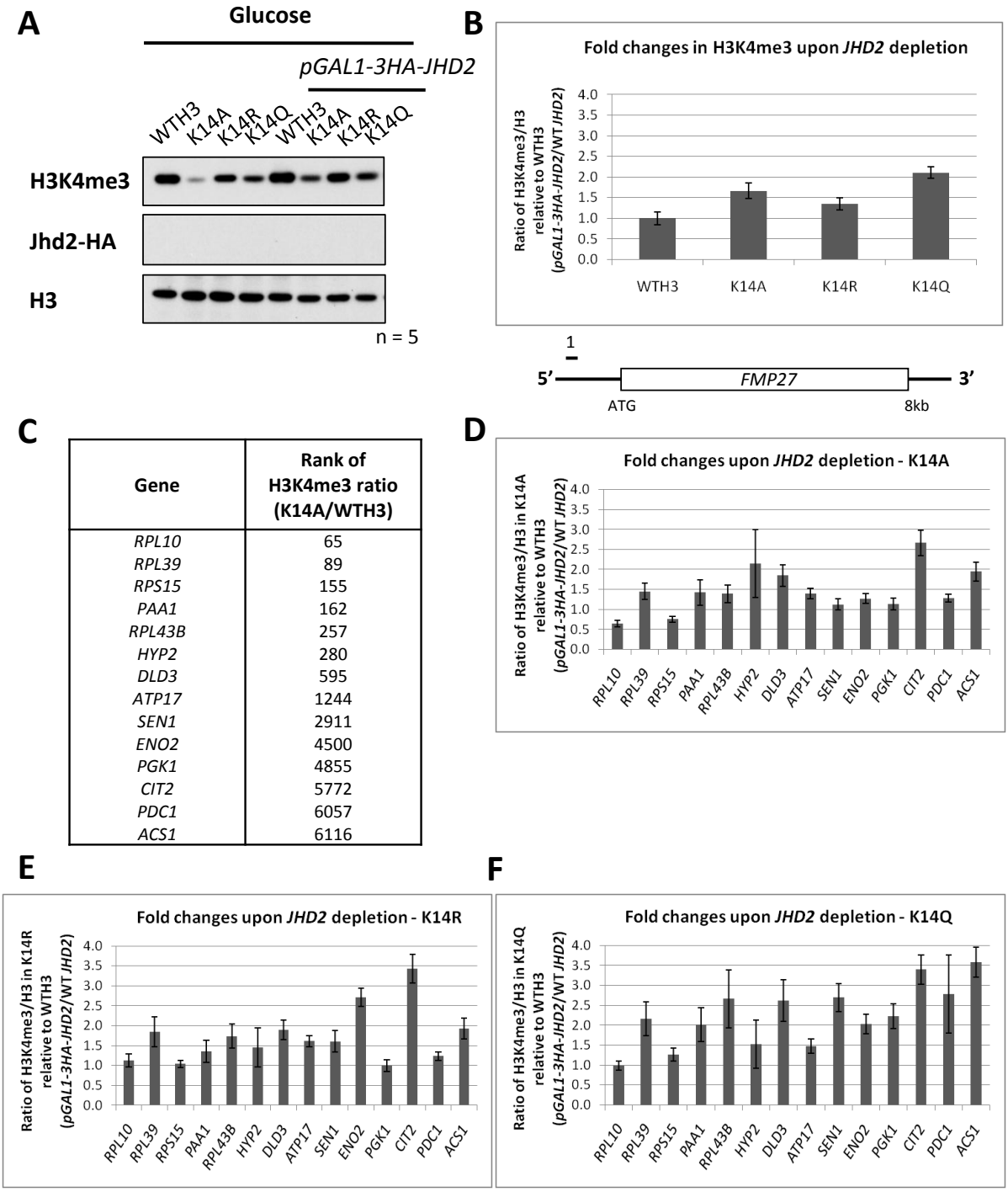
Figure 33. Jhd2 protein levels are not altered in the H3K14 substitution strains.

A Western blot showing levels of Jhd2 via its C-terminal HA tag in the wildtype and H3K14 substitution strains. An untagged strain (WTH3) was included as a negative control. Actin levels are shown as a loading control.

Next, the levels of H3K4me3 were assessed in the wildtype and H3K14 substitution strains in the absence of Jhd2*,* as achieved by growing the strains containing the *pGAL1-3HA-JHD2* constructs in glucose (Figure 34A). The depletion of Jhd2 protein was confirmed by undetectable HA signals on a Western blot, even after a long exposure to film. Depletion of Jhd2 caused a small increase in H3K4me3 in the presence of a wildtype copy of histone H3, as noted in Figure 30. Interestingly, increases in H3K4me3 were also observed in the H3K14 mutant strains, although it should be noted that depletion of Jhd2 did not restore H3K4me3 to wildtype levels in these strains. The same results were obtained when using strains in which *JHD2* had been deleted (data not shown). In order to produce more quantitative data that would allow evaluation of the relative increases in H3K4me3 upon Jhd2 depletion in the strains with wildtype or mutant H3K14, a ChIP experiment was performed (Figure 34B). At *FMP27*, the fold increases in H3K4me3 between the *pGAL1-3HA-JHD2* (glucose, Jhd2-depleted) relative to the WT *JHD2* strains were greater in the H3K14 mutant strains than for the strain with a wildtype copy of histone H3. Furthermore, the fold increases for the K14A and K14Q strains were greater than for the K14R strain, reflecting the relative levels of H3K4me3 in these strains. These results indicate that Jhd2 is having more effect on H3K4me3 at *FMP27* in the K14A and K14Q strains than the wildtype and K14R strains.

To expand this analysis, the fold increases in H3K4me3 upon Jhd2 depletion in the wildtype H3 and H3K14 mutant strains were obtained at the 5' ends of a number of other genes. The genes chosen and the ranking of their ratios of H3K4me3 (K14A/WTH3) from the ChIP-seq experiment (Chapter 3) are shown in Figure 34C. Figure 34D-F show the fold increases in H3K4me3 upon Jhd2 depletion relative to the wildtype histone H3 strain for the K14A, K14R and K14Q strains respectively. There is considerable variation between the effects of Jhd2 depletion at individual genes but there is a slight trend for H3K4me3 at the 5' ends of genes with the largest H3K4me3 ratios (K14A/WTH3) (those least affected by mutation of H3K14), to increase less than those with the smallest H3K4me3 ratios upon depletion of Jhd2. This may suggest that the reason why some genes are affected more than others by the mutation of H3K14 is partly due to differential Jhd2 action.

It is plausible that a maximal level of H3K4me3 is reached upon depletion of Jhd2 in the presence of wildtype histone H3 and that the fold increase in H3K4me3 is greater in the H3K14 mutant strains simply because there is a larger capacity for H3K4me3 to increase in these strains. However, this possibility can be discounted because at some genes, for example at *RPL43B*, Jhd2 depletion combined with the K14R mutation results in higher absolute levels of H3K4me3 than Jhd2 depletion in the presence of a wildtype copy of H3 but yet at these genes a greater fold-increase in H3K4me3 is measured in the K14R strain upon Jhd2 depletion.



Gene	Rank of H3K4me3 ratio (K14A/WTH3)
RPL10	65
RPL39	89
RPS15	155
PAA1	162
RPL43B	257
HYP2	280
DLD3	595
ATP17	1244
SEN1	2911
ENO2	4500
PGK1	4855
CIT2	5772
PDC1	6057
ACS1	6116

Figure 34. H3K4me3 increases more upon depletion of Jhd2 in the H3K14 substitution strains.

(A) A Western blot showing levels of H3K4me3 and Jhd2-HA in the wildtype and H3K14 substitution strains in the absence or presence of the *pGAL1-3HA-JHD2* construct. Histone H3 levels are included as a loading control. All strains were grown in glucose-containing media to repress *JHD2* expression in the *pGAL1-3HA-JHD2* strains. **(B)** A ChIP-qPCR experiment at *FMP27* showing the ratios of H3K4me3 in the *pGAL1-3HA-JHD2* strains relative to the WT *JHD2* strains in the WTH3 and H3K14 substitution strains. All fold changes in H3K4me3 are displayed relative to the WTH3 strain. The primer position is indicated on the locus map. Error bars show standard errors of the real-time PCR reaction. Results are representative of two independent experiments. **(C)** The rank of the ratio of the average H3K4me3 across the transcription unit (K14A/WTH3) from the ChIP-seq experiment for each gene assessed in (D)-(F). Genes were ordered from largest to smallest H3K4me3 ratios. **(D)–(F)** As (A) for the K14A, K14R and K14Q strains except qPCR was

performed at the 5' ends of the genes displayed in (C). The primer sequences are listed in the Materials and Methods.

To explore the effect of Jhd2 further, the reverse experiments were performed to assess H3K4me3 levels in the presence and absence of wildtype histone H3 when *JHD2* was overexpressed. As has already been shown in Figure 30, overexpression of *JHD2* causes a substantial loss of H3K4me3. This effect was also observed in the H3K14 mutant strains. However, because of the low starting amounts of H3K4me3 in the H3K14 strains, the further loss of H3K4me3 upon *JHD2* overexpression resulted in undetectable levels of H3K4me3 by Western blot, making it difficult to assess the relative effects of *JHD2* overexpression (Figure 35A). A ChIP experiment was performed to try to quantify these fold decreases in H3K4me3 upon *JHD2* overexpression but again, the levels of H3K4me3 in the H3K14 mutant *pGAL1-3HA-JHD2* (galactose, *JHD2* overexpression) strains were below the limit for accurate qPCR. Instead, another approach was taken to observe the effect of *JHD2* overexpression on H3K4me3 in the H3K14 mutant strains. A time-course of *JHD2* induction was performed following the addition of galactose to the culture media and global H3K4me3 levels were assessed in the WTH3, K14R and K14Q *pGAL1-3HA-JHD2* strains (Figure 35B). (This experiment was not performed with the K14A strain since this strain already had nearly undetectable levels of H3K4me3 before overexpression of *JHD2*.) H3K4me3 levels did not change in the wildtype strain during the time-course (data not shown), signifying that any changes in H3K4me3 observed during the time-course with the *pGAL1-3HA-JHD2* strains were due to the induction of *JHD2*. At t=0 (glucose), the H3K4me3 signal from the WTH3 *pGAL1-3HA-JHD2* strain is greater than that from the wildtype strain because the *GAL1* promoter is inactive in glucose and so this is effectively a *JHD2* delete. The induction of *JHD2* was monitored using the N-terminal HA tag and occurred at equal rates between the strains, independent of the presence of a wildtype copy of histone H3. (A lower HA band was frequently observed and since Jhd2 levels are under tight proteasomal control, this was assumed to be a Jhd2 degradation product.) After 4.5 h in galactose, the H3K4me3 signal had decreased in the

WTH3 *pGAL1-3HA-JHD2* strain. However, this decrease occurred sooner for the K14R and K14Q *pGAL1-3HA-JHD2* strains. It is unlikely that this faster decrease in H3K4me3 is simply due to the lower starting levels of H3K4me3 in the H3K14 mutant strains since the K14R strain has near-wildtype levels of H3K4me3 at t=0 but has already lost around half by 1 h. Therefore, from the experiments with the *JHD2* depletion and overexpression strains, it can be concluded that the H3K14 mutants are more sensitive to Jhd2 action than the wildtype strain. This sensitivity could stem from a number of factors including increased recruitment of Jhd2 to chromatin, better positioning of Jhd2 on the H3 tail or an increase in Jhd2 activity in the H3K14 substitution strains.

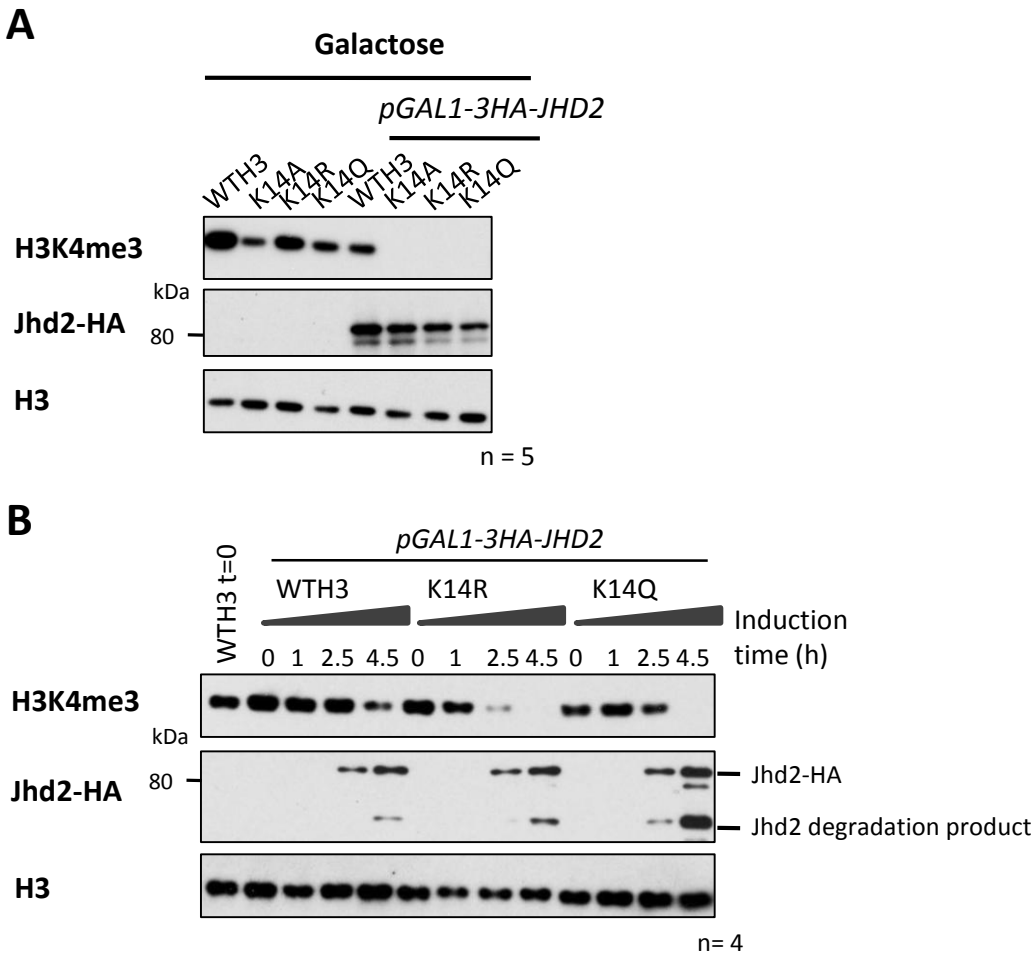
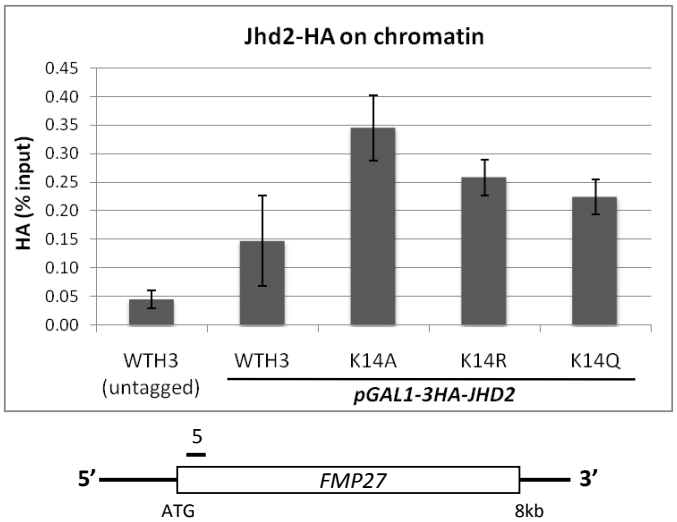


Figure 35. H3K4me3 is more sensitive to *JHD2* overexpression in the H3K14 mutant strains.

(A) A Western blot showing levels of H3K4me3 and Jhd2-HA in the wildtype and H3K14 substitution strains in the absence or presence of the *pGAL1-3HA-JHD2* construct. Histone H3 levels are included as a loading control. All strains were grown in galactose-containing media to induce *JHD2* overexpression in the *pGAL1-3HA-JHD2* strains. **(B)** A Western blot showing the levels of H3K4me3 and Jhd2-HA during a time-course of *JHD2* induction. At t=0, cells were transferred from glucose- to galactose-containing media. Aliquots were taken at the times indicated (h). Histone H3 levels are included as a loading control.

A preliminary experiment was performed to try to measure the relative amounts of chromatin-associated Jhd2 in the wildtype and H3K14 mutant strains. Previous studies have struggled to immunoprecipitate Jhd2 with chromatin since the demethylase is maintained at low cellular concentrations by Not4-dependent polyubiquitination and subsequent proteasomal degradation (Ingvarsdottir et al. 2007; Mersman et al. 2009). To overcome these difficulties, MG132, a proteasome inhibitor was used in combination with a deletion of *PDR5*, a multidrug transporter, with a view to increasing the cellular levels of Jhd2. However, this strategy failed to elicit any noticeable increase in global Jhd2 protein levels by Western blot (data not shown) and so was not pursued. The only study to date that successfully immunoprecipitated Jhd2 with DNA made use of the higher levels of Jhd2 caused by its overexpression and a double crosslinking protocol with dimethyl adipimidate followed by formaldehyde (Huang et al. 2010). This latter protocol adjustment did not yield an improvement but a signal was detected over the untagged control when *JHD2* was overexpressed (Figure 36A). A caveat to this approach is that *JHD2* may localise differently when present at non-wildtype levels. In this experiment, the levels of Jhd2 on chromatin were higher in the H3K14 mutant strains than the wildtype strain at *FMP27*. The levels of Jhd2 on chromatin were also assessed at the 5' ends of *PAA1, HYP2* and *CIT2*, ordered according to their ratios of H3K4me3 in the K14A relative to the WTH3 strain (from largest to smallest, Figure 36B). Unexpectedly, the K14R strain, with near-wildtype levels of H3K4me3, had the highest levels of Jhd2 at these genes. Because there was not more Jhd2 at *CIT2,* which had the smallest K14A/WTH3 H3K4me3 ratio of the three genes tested, than at *HYP2* or *PAA1*, increased Jhd2 recruitment to chromatin is unlikely to be the main the mechanism for the differential sensitivities of these genes to mutation of H3K14 (Figure 36C). This experiment would need to be repeated before any more substantial conclusions were made about the altered recruitment of Jhd2 to chromatin in the wildtype and H3K14 mutant strains. However, even if this preliminary result does not repeat, it would not rule out contributions of H3K14 to the exact positioning of Jhd2 on the histone H3 tail or regulation of Jhd2 activity.

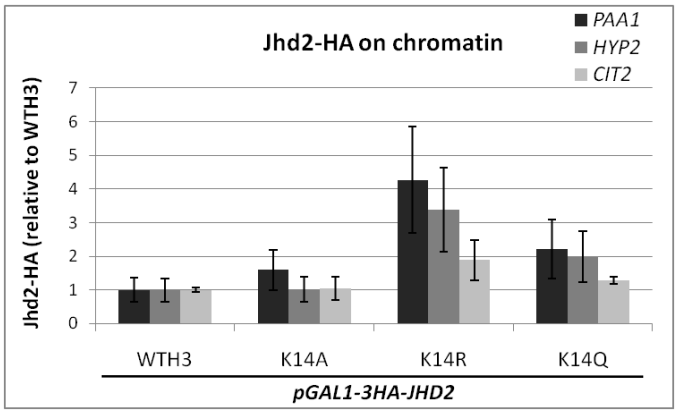
A



B

Gene	Rank of H3K4me3 ratio (K14A/WTH3)
PAA1	162
HYP2	280
CIT2	5772

C



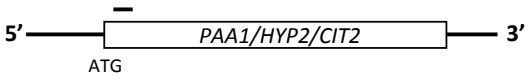


Figure 36. Jhd2 chromatin association increases in the H3K14 substitution strains.

(A) A preliminary ChIP-qPCR experiment at *FMP27* showing the levels of Jhd2-HA on chromatin in the wildtype and H3K14 mutant *pGAL1-3HA-JHD2* strains. Strains were grown in galactose-containing media to induce *JHD2* overexpression. An untagged strain is included as a negative control. The primer position is indicated on the locus map. Error bars show standard errors of the real-time PCR reaction. **(B)** The rank of the ratio of the average H3K4me3 across the transcription unit (K14A/WTH3) from the ChIP-seq experiment for each gene assessed in (C). Genes were ordered from largest to smallest H3K4me3 ratios. **(C)** The levels of Jhd2-HA on chromatin at the 5' ends of *PAA1*, *HYP2* and *CIT2* in the H3K14 mutant strains relative to the WTH3 strain. The primer positions are indicated on the locus map and sequences are listed in the Materials and Methods. Error bars show standard errors of the real-time PCR reaction.

5.2.4. Investigating the function of Spp1 in the H3K14 mutant strains

The Jhd2 data presented so far demonstrate that the H3K14 substitution strains are more sensitive to Jhd2 action than the strain with a wildtype copy of histone H3. However, the failure to restore H3K4me3 to wildtype levels in the H3K14 mutant strains upon depletion/deletion of Jhd2 indicates that increased Jhd2 action cannot be the sole reason for the lower levels of H3K4me3 in the H3K14 mutant strains. The balance of methylation/demethylation may be shifted in the H3K14 strains not only by increased demethylation but also by decreased methylation. Since mutation of H3K14 predominantly causes a trimethyl-specific effect on H3K4, the function of Spp1 in the H3K14 strains was probed.

Like Jhd2, neither the levels of Spp1 protein nor COMPASS (as represented by Swd1, a subunit required for COMPASS integrity) are altered in the H3K14 or H3P16 mutant strains (Figure 37A). The *pGAL1-3HA-SPP1* strains were grown in glucose-containing media to deplete Spp1. Depletion of Spp1 resulted in decreased H3K4me3 in the presence and absence of a wildtype copy of histone H3 (Figure 37B). The relative global levels of H3K4me3 are likely to be maintained in the K14A, K14R and K14Q strains, as suggested by the residual H3K4me3 signal in the K14R strain in the longer exposure of the Western blot to film. Most interestingly, depletion of Spp1 in the H3K14 mutants had a greater effect on H3K4me2 than the WTH3 strain. The same result was obtained after deletion (rather than depletion) of *SPP1* in the H3K14 mutant strains (data not shown). This could again support the idea that the deposition of H3K4me3, promoted by Spp1, masks the loss of H3K4me2 caused by increased demethylation, either due to *JHD2* overexpression or due to the greater Jhd2 activity in the H3K14 substitution strains.

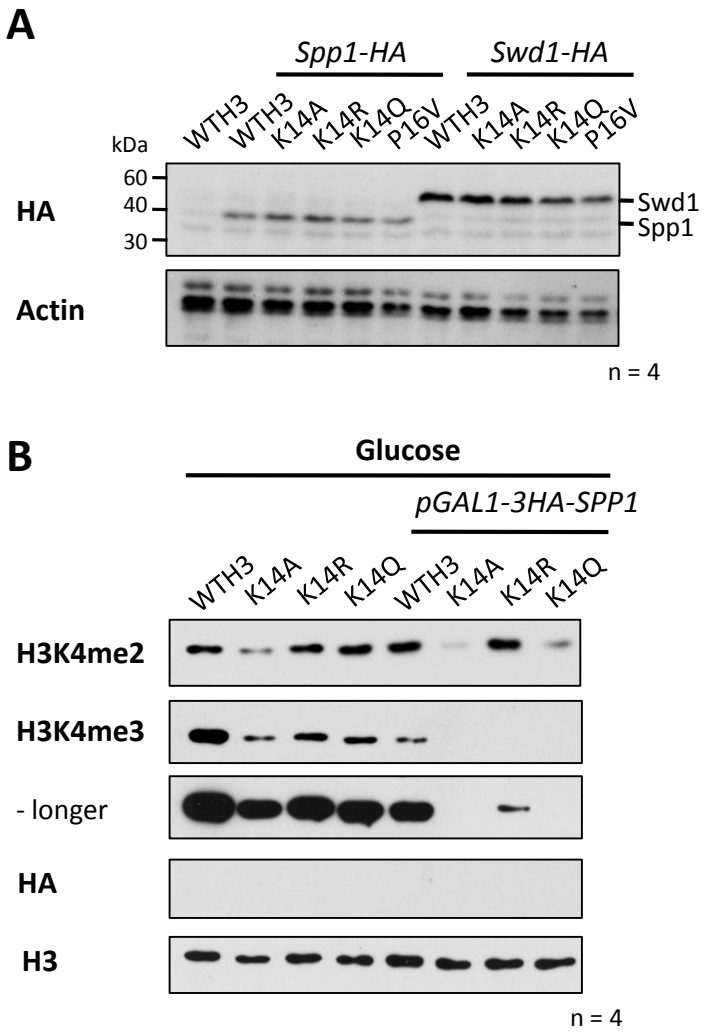


Figure 37. Investigating the involvement of Spp1 in the H3K14-H3K4me3 crosstalk.

(A) A Western blot showing levels of Spp1 and Swd1 via their C-terminal HA tags in the wildtype and H3K14 substitution strains. An untagged strain (WTH3) was included as a negative control. Actin levels are shown as a loading control. **(B)** Western blots showing levels of H3K4me3 and Spp1-HA in the wildtype and H3K14 substitution strains in the absence or presence of the *pGAL1-3HA-SPP1* construct. A longer exposure of H3K4me3 is shown to enable visualisation of the low levels of H3K4me3 in the K14R *pGAL1-3HA-SPP1* strain. Histone H3 levels are included as a loading control. All strains were grown in glucose-containing media to repress *SPP1* expression.

Many attempts were made to measure the relative levels of chromatin-bound Spp1 and Swd1 in the wildtype and H3K14 mutant strains by ChIP but no reproducible result was obtained. Experiments with the H3P16 mutant strain proved more successful (Figure 38A). There was less Spp1 on chromatin in the P16V than the WTP16 strain. Furthermore, the reductions in Spp1 occupancy roughly matched the decreases in H3K4me3 at these three loci (Figure 38B, reproduced from Figure 23). The ability of H3P16 to influence the binding of Spp1 to chromatin

supports the results in Chapter 4 showing the altered binding of H3K14ac and H3K18ac antibodies and the Spt7 bromodomain to sites in the vicinity of H3P16 and adds weight to the hypothesis that a proline at position 16 in the histone H3 tail may be important for the local tail structure and protein binding. More experiments need to be performed with the P16V strain to investigate any possible increased effect of Jhd2 when compared to the wildtype strain.

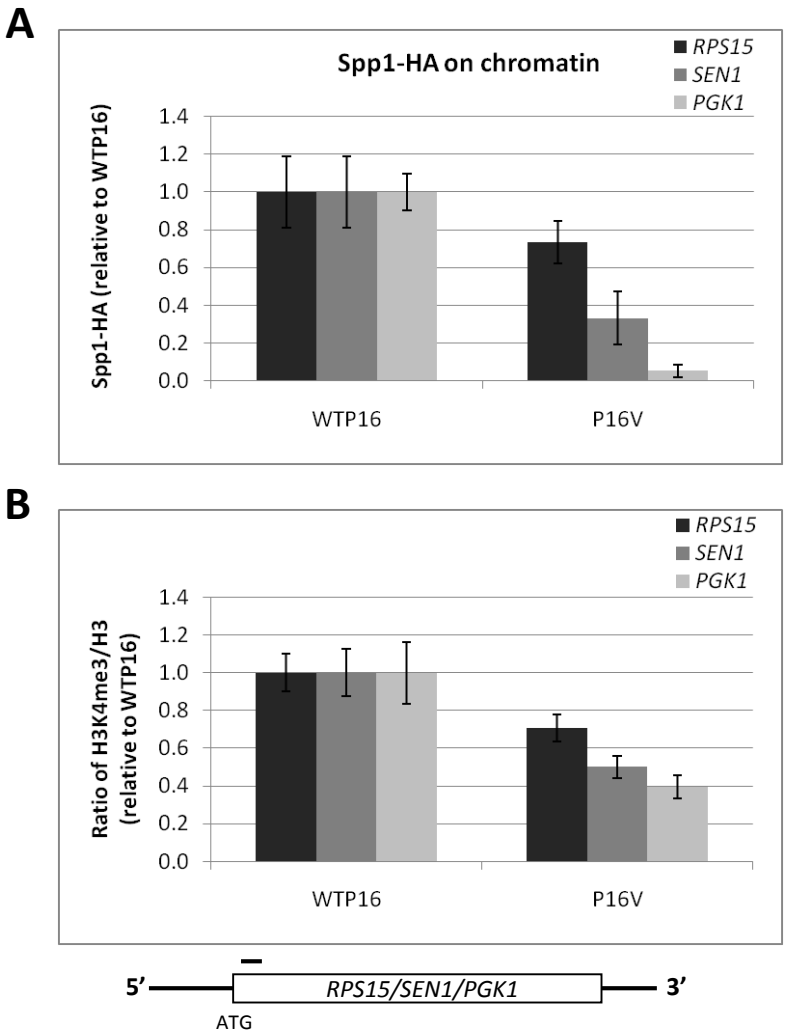


Figure 38. Spp1 chromatin association is decreased in the H3P16 substitution strain.

The levels of **(A)** Spp1-HA and **(B)** H3K4me3 (reproduced from Figure 23) on chromatin at the 5' ends of *RPS15, SEN1* and *PGK1* in the H3P16 mutant strains relative to the WTP16 strain. Primer positions are indicated on the locus map and primer sequences are listed in the Materials and Methods. H3K4me3 signals are normalised to levels of histone H3. Error bars show standard errors of the real-time PCR reaction. Results are representative of at least two independent experiments.

5.2.5. Other mechanisms by which the H3K14-H3K4me3 crosstalk may operate

As mentioned in the introduction to this chapter, there are a number of described crosstalk pathways that can regulate the levels of H3K4me3. Therefore it is possible that the H3K14-H3K4me3 crosstalk may operate indirectly through these previously described pathways. The Western blot shown in Chapter 3, Figure 13 demonstrated that, apart from H3K4me3, mutation of H3K14 to three different residues did not differentially affect the levels of other tested modifications on the histone H3 tail. However, the most characterised H3K4me3-regulating pathways are those involving H3R2 methylation and H2BK123 monoubiquitination, which were not tested in this original experiment. Antibodies were obtained against H3R2me2a (Abcam), which is suggested to repress H3K4me3 by inhibition of Spp1 binding (Kirmizis et al. 2007), and global levels of H3R2me2a were assessed in the wildtype and H3K14 substitution strains by Western blot and ChIP. Unfortunately, the quality of this antibody was very poor, with a high background and no enrichment of signal over the R2A negative control, preventing the attainment of results from these experiments (data not shown).

H2BK123 monoubiquitination is required for H3K4me2/3 (Dover et al. 2002; Sun and Allis 2002) and so it is possible that either H2B ubiquitination influences H3K14 to regulate H3K4me3 or vice versa. To distinguish between these possibilities, the levels of H2B ubiquitination were assessed in the H3K14 mutant strains by probing for histone H2B and looking for a higher molecular weight band (representing ubiquitinated H2B). However, once again, the low quality of the antibodies raised against histone H2B precluded this experimental approach. Instead, two of the enzymes required for H2B ubiquitination, *RAD6*, the E2 ubiquitin-conjugating enzyme, and *BRE1*, the E3 ubiquitin ligase, were deleted in the WTH3 and H3K14 mutant strains (Figure 39). There was no detectable H3K4me2/3 in these double mutant strains and moreover no recovery of H3K4me2/3 upon deletion of *JHD2.* These results demonstrate that H2B ubiquitination acts upstream of or in a separate pathway from H3K14 in the control of H3K4me3. Given the H3K4me3-specific effect of

H3K14 mutation and the H3K4me2/3 effect of H2B ubiquitination, it is more likely that H2BK123 monoubiquitination operates in a separate pathway from H3K14 to regulate H3K4me3 levels.

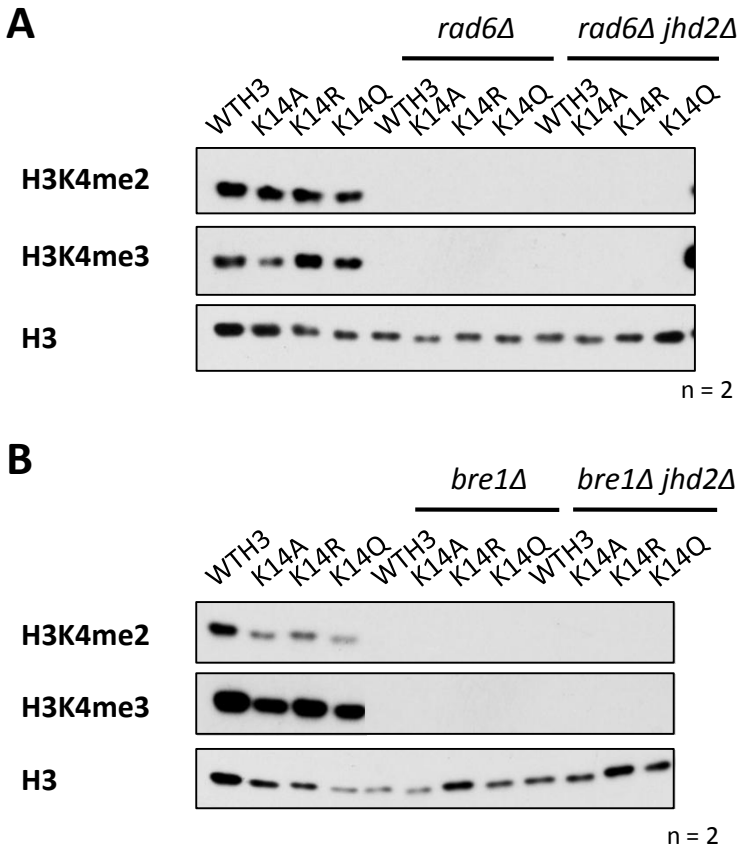


Figure 39. Investigating the involvement of H2BK123 monoubiquitination in the H3K14-H3K4me3 crosstalk.

Western blots showing the levels of H3K4me2 and H3K4me3 in the wildtype and H3K14 mutant strains with additional deletions of **(A)** *RAD6* and *JHD2* and **(B)** *BRE1* and *JHD2.* Histone H3 levels are included as a loading control.

5.3. Discussion

5.3.1. The lower levels of H3K4me3 in the H3K14 substitution strains result from an altered balance between methylation and demethylation of H3K4

The experiments with the *SPP1* and *JHD2* deletion and overexpression strains in the presence of a wildtype copy of histone H3 led to the idea of balanced methyltransferase and demethylase action determining the levels of H3K4me3. This work underpins the hypothesis that the balance between methylation and demethylation is tipped in favour of demethylation over methylation in the H3K14 mutant strains, resulting in a substantial decrease in the levels of H3K4me3, with a small effect on H3K4me2. H3K4me3 was more sensitive to *JHD2* overexpression in the absence of H3K14 and furthermore, a greater fold increase in H3K4me3 was observed upon depletion of *JHD2* in the H3K14 mutant strains when compared to the strain with a wildtype copy of histone H3. Preliminary ChIP data hinted that the levels of chromatin-associated Jhd2 were higher in the H3K14 mutant strains. However this latter finding would need to be repeated before more substantial conclusions were made. Published evidence showing the lack of dependence of Jhd2 chromatin association on the histone H3 tail might argue against an involvement for a residue in the first 28 residues of the histone H3 tail for the recruitment of Jhd2 to chromatin, although this study assayed bulk chromatin association by cell fractionation followed by Western blot rather than by ChIP (Huang et al. 2010). It is distinctly possible, however, that there are two modes of Jhd2 recruitment to allow H3K4 demethylation. The first may recruit the protein to chromatin; this may not be dependent on the histone H3 tail but is dependent on the Jhd2 PHD domain. The second may involve a more precise positioning of Jhd2 on or near the histone H3 tail to allow subsequent demethylation of H3K4. This latter positioning may depend on a number of residues or modifications and could be influenced directly or indirectly by H3K14.

The work exploring the relative activities of the Set1 methyltransferase in the wildtype and H3K14 mutant strains was more challenging. Because Set1 operates as part of the COMPASS complex, it

was difficult to isolate the trimethyl-specific Set1 function, promoted by Spp1. Overexpression of *SPP1* had little effect despite a large increase in the level of Spp1 protein, presumably due to another limiting factor for H3K4 trimethylation. A similar result was obtained when *SET1* was overexpressed in this study, an effect that had previously been documented by Halbach et al. (2009). However, in this case the authors observed increased *SET1* mRNA upon *SET1* overexpression but only detected modest increases in Set1 protein levels. They proposed that co-translational assembly of the COMPASS complex allows the stabilisation of nascent Set1 protein and consequently enables regulation of Set1 protein levels to match the availability of other COMPASS components. A later study confirmed the importance of COMPASS subunits in determining the levels of Set1 protein since deletion of *SWD1* or *SWD3* did not affect the level of *SET1* mRNA but did cause a near-complete loss of Set1 protein (Mersman et al. 2011). Therefore the tight control of the relative cellular protein levels of the COMPASS subunits makes it difficult to analyse their individual functions using overexpression.

Curiously, overexpression of both *JHD2* and *SPP1* resulted in lower global levels of H3K4me2 and H3K4me3 than overexpression of either gene alone. Unfortunately, the deletion of the Spp1 N-terminal domain (including the PHD finger) resulted in lower Spp1 and H3K4me3 levels, presumably due to decreased stability of the truncated protein. Therefore it was not possible to determine whether the PHD finger was involved in this negative effect of *SPP1* when overexpressed in concert with *JHD2*. Potentially the excess levels of Spp1 protein sequester COMPASS away from its site of action or render it less efficient as a methyltransferase. Whatever the function of Spp1, these results demonstrate the importance of proper levels of Spp1 protein for H3K4 methylation.

Deletion of *SPP1* decreased H3K4me3 in the presence and absence of a wildtype copy of histone H3, but the very low levels of H3K4me3 in the H3K14 mutant strains upon *SPP1* deletion

prohibited analysis of the relative changes caused by *SPP1* deletion in the wildtype and H3K14 mutant strains. Chromatin immunoprecipitation of Spp1 or Swd1 failed to give a reliable result in the H3K14 mutant strains. However, substitution of H3P16 resulted in reproducibly lower Spp1 on chromatin than in the wildtype strain, in agreement with the ability of this residue to influence the binding of other proteins in its vicinity (Chapter 4). The fact that mutation of H3P16 decreased chromatin-associated Spp1 to roughly the same extent as the decrease in H3K4me3 at three individual genes lends support to the involvement of Spp1 in the H3P16-H3K4me3 crosstalk. More experiments need to be performed with the P16V strain to further assess the contributions of Spp1 and Jhd2 to the H3P16-H3K4me3 crosstalk.

The interaction of COMPASS with the nucleosome could operate similarly to that of Set2. Several residues from histones H3, H4 and H2A are required for H3K36me2/3 and these are proposed to form a docking site for Set2 (Du et al. 2008; Du and Briggs 2010; Hainer and Martens 2011; Endo et al. 2012). COMPASS, as a large protein complex, could also be expected to make contact with multiple residues in the nucleosome, some of which may be required for H3K4 methylation. Chandrasekharan et al. (2010) have already proposed that the C-terminal helix in histone H2B is a docking site for COMPASS and several other residues in the nucleosome core are required for H3K4me3 (Nakanishi et al. 2008). Potentially, H3K14/H3P16 forms another point of contact between COMPASS and the nucleosome. However, the failure to detect a reproducible decrease in chromatin-associated Spp1 or Swd1 in the H3K14 mutant strains does not mean that this residue is not important for COMPASS activity. Indeed, whilst mutation of six lysine residues in the histone H3 tail to glutamine caused reduced H3K36me2/3, these mutations did not decrease Set2 recruitment to chromatin *in vivo* but did affect Set2 catalytic activity *in vitro* (Psathas et al. 2009). Interestingly, this study also observed different effects on Set2 activity dependent on which residue was engineered in place of the endogenous lysines in the tail: substitution of the six lysines with glutamine reduced H3K36 methylation but substitution with arginine did not.

Therefore, different substitutions at histone H3 position 14 have the potential to differentially influence Set1 catalytic activity. Because of the ability of H3P16 to influence the structure of the H3 tail, mutation of this residue may have a more severe effect on the interaction between COMPASS and the nucleosome that results in reduced recruitment of Spp1 to chromatin to cause decreased H3K4me3.

5.3.2. Differential alteration of the methylation/demethylation balance may cause the variations in H3K4me3 K14A/WTH3 ratios between genes

Following on from the notion that the lower levels of H3K4me3 in the H3K14 substitution strains are caused by changes in the relative activities of H3K4 methylation and demethylation, it is possible that differential shifting of the balance between methylation and demethylation causes the variable effects of H3K14 substitution at individual genes. Therefore, at the genes in the 'top' class, with the 200 largest H3K4me3 mutant/wildtype ratios, the balance between H3K4 methylation and demethylation is closer to wildtype than at the 'bottom' class of genes, with the 200 smallest H3K4me3 ratios, where the balance is tipped more in favour of demethylation over methylation. There is some evidence to support this idea. Firstly, the fold increases in H3K4me3 upon *JHD2* depletion are greater at the genes with lower H3K4me3 K14A/WTH3 ratios than for those less affected by H3K14 substitution. Secondly, the levels of Spp1 on chromatin in the P16V strain vary with the ratios of H3K4me3 in the P16V relative to the WTP16 strain at the three genes tested. More experiments at a greater number of genes would validate this hypothesis.

5.3.3. Set1 has two RRM domains – a link to regulation of H3K4 methylation by RNA?

Not investigated in this work are the functions of the Set1 RNA recognition motifs (RRMs). Schlichter and Cairns (2005) and Fingerman et al. (2005) independently identified the requirement of an RRM in Set1 for H3K4me3 but not H3K4me2. Furthermore, mutations that abolish RNA binding to the Set1 RRM also specifically diminish H3K4me3 *in vivo*, suggesting an

involvement of RNA in the regulation of H3K4me3 (Schlichter and Cairns 2005). An RNA molecule has been purified with a complex containing a subset of COMPASS components, including Set1. However, this was found to be the *SET1* mRNA and its interaction with this complex did not depend on the Set1 RRMs (Halbach et al. 2009). Schlichter and Cairns speculated that nascent mRNA stimulates Set1 via the RRM to trimethylate H3K4 near the 5' ends of genes. This hypothesis was questioned by a study that identified a second RRM, RRM2, downstream of the previously characterised RRM1 (Trésaugues et al. 2006). Isolated RRM1 could not bind RNA *in vitro* but required the presence of RRM2. Furthermore, mutation of the residues important for *in vitro* RNA binding within RRM2 did not affect H3K4 methylation *in vivo.* Instead, the H3K4me3-specific effect of the RRM1 mutation may have been due to the destabilisation of Set1 caused by mutation of this domain (Trésaugues et al. 2006). However, the attractiveness of the idea of an RNA-mediated regulation of Set1 could warrant its further investigation.

5.3.4. Other mechanisms by which the H3K14-H3K4me3 crosstalk may operate

Rather than affecting H3K4me3 directly, H3K14 may influence the levels of other modifications that can regulate the levels of H3K4me3. Alternatively, these previously characterised crosstalk pathways may control H3K4me3 through H3K14. H3R2me2a inhibits H3K4me3 by interfering with Spp1 binding (Kirmizis et al. 2007). Unfortunately, the quality of the antibodies raised against H3R2me2a precluded the exploration of the involvement of H3K14 in this crosstalk. The other well-investigated crosstalk pathway required for H3K4me3 is the *trans*-tail crosstalk involving H2BK123 monoubiquitination (Dover et al. 2002; Sun and Allis 2002). The levels of H2Bub could not be monitored in the H3K14 mutant strains due to poor quality antibodies but double mutant strains were created to test the involvement of H3K14 in the H2Bub-mediated effect on H3K4me3. Mutation of H3K14, even to arginine, could not recover H3K4me3 in the *rad6Δ* or *bre1Δ* strains that lack H2Bub. If H3K14 were acting upstream of H2Bub in this pathway then it would be expected to observe the relative K14A/K14R/K14Q levels of H3K4me3 in the double

mutant strains. Because no recovery was obtained in double mutant strains, H2Bub must either act upstream or in a separate crosstalk pathway from H3K14. Since loss of H2Bub and H3K14 affect different H3K4 methylation states (H2Bub affects H3K4me2/3 whereas H3K14 affects predominantly H3K4me3), it is more likely for H2Bub and H3K14 to act in separate pathways converging on H3K4me3. This could be more firmly established if an antibody of sufficient quality could be obtained to enabling monitoring of H2Bub levels in the H3K14 mutant strains.

A factor that should be considered in future studies is the effect of H3K4ac. This modification is localised to the 5' ends of genes, upstream of H3K4me3 and H3K4me2 (Guillemette et al. 2011). The authors proposed that H3K4me2/3 is required to prevent the spreading of H3K4ac into the coding regions of genes. It is possible that, in the H3K14 mutants, H3K4ac is deregulated and can overcome the H3K4me2/3-mediated restriction of its genomic location and spread into the coding regions. This could either be caused by the H3K14 mutant-induced reduction in H3K4me3 or intriguingly, could be responsible for the reduction in H3K4me3 and, to a smaller extent, H3K4me2 (depending on the extent of H3K4ac spread into the gene). However, an argument against the latter possibility is that modulation of H3K4ac levels does not result in altered global H3K4 methylation (Guillemette et al. 2011). A preliminary experiment was performed to assess H3K4ac levels in the H3K14 mutant strains, but the commercially available antibody raised against H3K4ac was not specific for this modification by Western blot or by ChIP.

Chapter 6.

Characterisation of H3K18me1

6. Characterisation of H3K18me1

6.1. Introduction

As noted in Chapter 3, mutation of H3K14 resulted in undetectable H3K18ac. With the discovery of H3K18me1 in yeast by mass spectrometry (Garcia et al. 2007), an extra layer of complexity was added to this crosstalk.

6.1.1. Discovery of H3K18me1

Garcia *et al.* employed the use of sensitive mass spectrometry, allowing the identification of previously unknown modifications. They detected H3K18me1 in yeast, mouse and humans but not *Tetrahymena thermophila*. The same group went on to characterise this modification further by measuring the dynamics of H3K18me1 on histones from HeLa cells (Zee et al. 2010). $^{13}CD_3$-methionine was introduced into the culture media and cell aliquots were taken over a number of days and analysed by high-resolution mass spectrometry. Incorporation of a heavy methyl group caused a +4 Da increase in histone mass compared to the histones with an unlabelled methyl group. This shift in size enabled calculation of the half-maximal time of a methyl group on each residue. H3K18me1 had a half-maximal time of 1.207 days. This was compared to the half-maximal times of previously characterised 'activating' and 'repressive' modifications, and was deemed to be more similar to the half-maximal times for modifications with 'repressive' functions in chromatin. The dynamics of H3K18me1 were also similar to those observed for histone H3 ($t_{1/2}$= 1.298 days). From this analysis, the authors proposed that H3K18me1 may have a repressive effect in chromatin, possibly by antagonising the positively-acting H3K18ac.

6.1.2. Previously characterised acetyl/methyl switch at lysine residues

Acetyl/methyl switches have been described previously at residues in histone H3. In mammalian cells, the crosstalk between H3K9ac and H3K9 methylation has been well studied. These two

modifications are localised to different regions of the genome (H3K9ac is in euchromatin whereas H3K9 methylation is in heterochromatin) (Wang et al. 2008) and are temporally regulated during the cell cycle-linked induction and repression of DHFR (Nicolas et al. 2003). There is a correlation between the presence of H3K9ac, the absence of H3K9 methylation and the expression of DHFR (and vice versa), suggesting that the switch in modifications at H3K9 in gene promoters may influence the expression of the downstream gene.

Another potential crosstalk between acetylation and methylation of the same residue occurs at H3K36 in *S. cerevisiae* (Morris et al. 2007). However, the evidence for this is purely a genome-wide anti-correlation between the sites of H3K36ac and H3K36 methylation. A more thoroughly tested crosstalk is that between H3K4ac and H3K4me2/3 in *S. cerevisiae* (Guillemette et al. 2011). The authors demonstrated that the crosstalk between acetylation and methylation at H3K4 is more than a simple competition for modification sites: deletion of *SET1* removes all three species of H3K4 methylation, previously covering the length of genes. However, the increase in H3K4ac is predominantly localised to gene promoters in the *set1Δ* strain. Additionally, this effect is not reciprocal: removing or increasing the level of H3K4ac does not affect the levels of H3K4 methylation. Thus, whilst it is true that the same lysine residue cannot be modified simultaneously by two different modifications, it should be remembered that there are two copies of each histone per nucleosome and that these could potentially harbour different modifications. It appears therefore that there are more complex mechanisms than simple competition governing the relationship between acetylation and methylation of H3K4 and potentially other histone residues.

6.1.3. Characteristics of methyltransferases

Lysine methyltransferases catalyse the addition of up to three methyl groups to the target lysine using S-adenosyl-methionine as a methyl donor. There are a number of different domains/motifs

that are commonly found in methyltransferases. Three of the four yeast histone methyltransferases identified to date contain the catalytic SET domain: Set1, Set2 and Set5. This domain is comprised of 130-140 amino acids and was first identified in *Drosophila* (its name is derived from the three first discovered SET proteins in *Drosophila*: Su(var)3–9, Enhancer of zeste, and trithorax). However, Dot1, the remaining characterised yeast histone methyltransferase, does not contain a SET domain. *In silico* studies have identified a list of putative methyltransferases based on sequence similarity to known methyltransferase domains (Petrossian and Clarke 2009). These proteins can be used as candidates when searching for a methyltransferase for a new modification.

6.1.4. Characteristics of demethylases

Six demethylases have been identified in yeast and these all contain the catalytic JmjC domain, conserved from yeast to humans (Klose et al. 2006). The demethylase activity is dependent on Fe^{2+} and α-ketoglutarate cofactors, which are both coordinated by the JmjC domain. In contrast to the yeast methyltransferases, the demethylases have more overlapping functions, for example both Rph1 and Jhd1 remove H3K36 methylation (Kim and Buratowski 2007; Tu, S et al. 2007).

6.1.5. Aims

The main aim of this chapter is to characterise the recently discovered modification, H3K18me1. This characterisation will include its distribution over genes, a genome-wide analysis of localisation, and a search for the methyltransferase, demethylase and other factors influencing the levels of this modification. The possible crosstalk between H3K18ac and H3K18me1 will be explored further and the effect of H3K14 mutation on H3K18me1 will be described.

6.2. Results

6.2.1. Characterisation of the first antibody raised against H3K18me1

In order to study H3K18me1, an antibody was raised against this modification. A monomethylated peptide was designed corresponding to the sequence of the histone H3 N-terminal tail surrounding lysine 18 (cGKAPR**K(me)**GLASKAARK, Figure 40A). This was used to raise a polyclonal antibody in rabbit. Two preparations of the antibody were obtained: the serum from the final bleed and an affinity-purified preparation. The affinity-purified antibody specifically bound the monomethylated peptide over the unmethylated peptide on a dot blot (Figure 40B) and both antibody preparations were specific by Western blot to synthetic histones monomethylated on lysine 18 (Figure 40C). However, when the affinity-purified antibody was tested against whole cell extracts from wildtype yeast and strains in which lysine 18 had been replaced by alanine (K18A) or the H3 N-terminal tail had been deleted (H3 Δ1-28), there was no specificity for lysine 18, or indeed H3 (Figure 40D). The serum gave very high background when analysing whole cell extracts by Western blot and also showed no specificity (data not shown). However, when the antibody was tested for specificity by ChIP at *FMP27*, a constitutively active gene, the affinity-purified antibody gave no signal but the serum specifically immunoprecipitated a modification on H3K18, as evidenced by the low signal observed in the K18A strain (Figure 40E). Therefore ChIP was employed to assess H3K18me1 levels in all further experiments with this antibody.

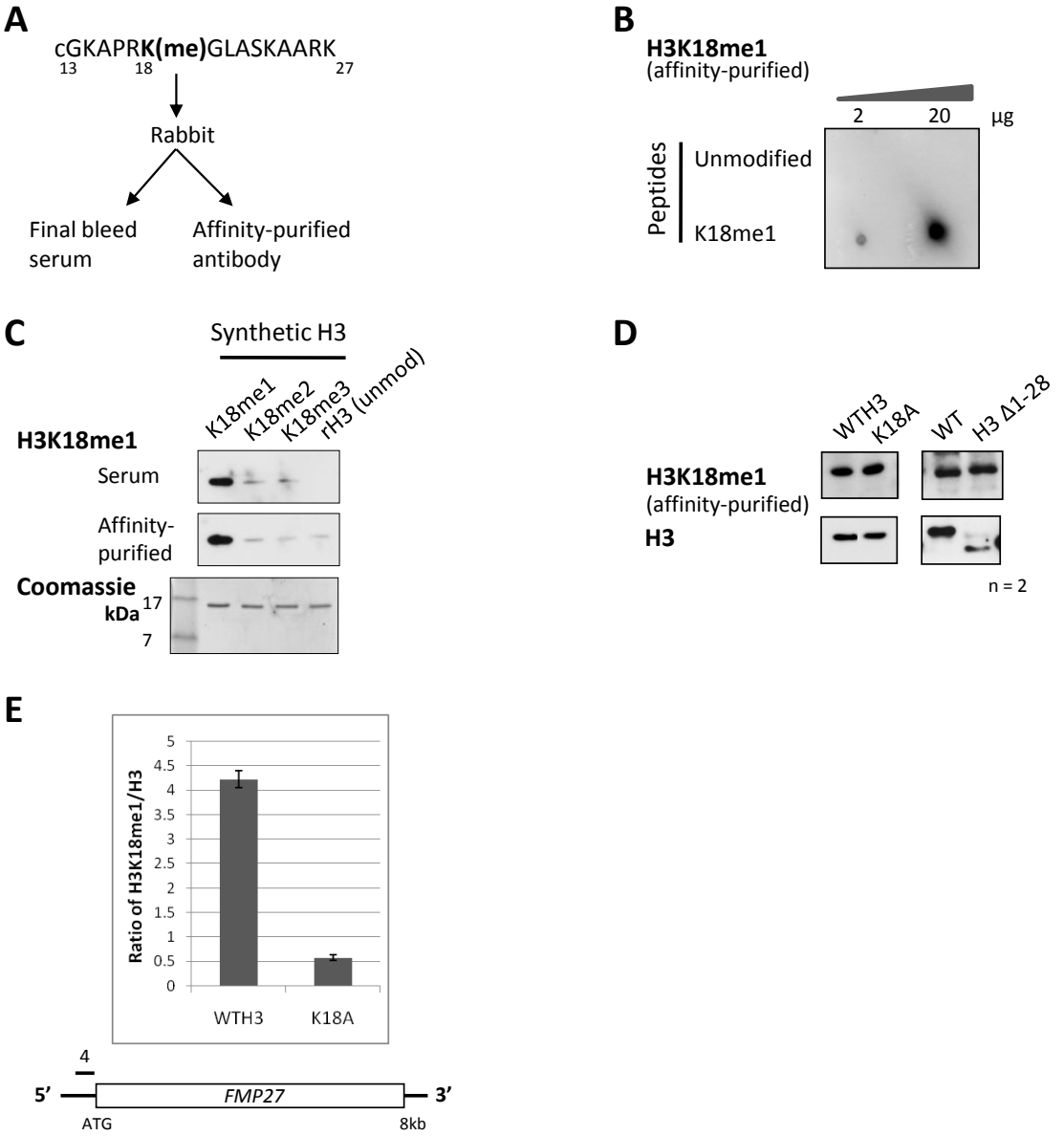


Figure 40. Characterisation of the first antibody raised against H3K18me1.

(A) The methylated peptide sequence used to raise the antibody in rabbit. Numbers indicate the corresponding residue in histone H3. An N-terminal cysteine (c) was included to facilitate peptide immobilisation in subsequent experiments. **(B)** A dot blot showing the specificity of the affinity-purified antibody for the mono-methylated over the unmodified H3 peptide at two amounts of peptide (2 and 20 µg) (D. Clynes). **(C)** Testing the specificity of both antibody preparations against synthetic histone H3 that has been mono-, di- or trimethylated on lysine 18 (kindly provided by R. Klose). Recombinant H3 (rH3, expressed and purified from *E. coli*) is unmodified. A Coomassie-stained gel is shown as a loading control. **(D)** Western blot analysis of H3K18me1 levels in whole cell extracts from K18A and tail-less H3 strains and their isogenic WT strains (WTH3 and NSY429 respectively). H3 levels are shown as a loading control. **(E)** A ChIP-qPCR experiment showing the specificity of the H3K18me1 antibody serum preparation for a modification on H3K18 at primer 4 on *FMP27* (indicated on the locus map). Signals are normalised to levels of H3. Results shown are representative of at least three experiments. Error bars show standard errors of the real-time PCR reaction.

A clue for the lack of specificity of the antibody raised against H3K18me1 came from an experiment in which whole cell extracts from strains with individual deletions of seven of the SET domain-containing methyltransferases in yeast were tested for levels of H3K18me1 by Western blot (Figure 41A). There was a noticeable loss of the species recognised by the antibody raised against H3K18me1 in the *set7Δ* strain. Set7/Rkm4 is a characterised methyltransferase for the ribosomal protein Rpl42ab (Webb et al. 2008). Since ribosomes run at a similar size to histones on SDS-polyacrylamide gels, it was possible that the H3K18me1 antibody was instead recognising the site on Rpl42ab methylated by Set7. Indeed this was the case, as demonstrated by the slightly smaller size of the protein recognised by the antibody raised against H3K18me1 when compared to the histone H3 loading control and the reduction of this signal in the *rpl42aΔ* strain (Figure 41B). (Rpl42ab is encoded by two identical genes, *RPL42A* and *RPL42B.* The majority of protein is produced from *RPL42A* and so the remaining signal in the *rpl42aΔ* strain is presumably due to the Rpl42ab protein produced from the unaffected identical gene *RPL42B.*) The recognition of a methylation mark on Rpl42ab by the antibody raised against H3K18me1 is surprising since the amino acid sequences surrounding the sites of methylation in histone H3 and Rpl42ab are not similar (Figure 41C). The work with this antibody demonstrates the importance of controlling for antibody specificity *in vivo*.

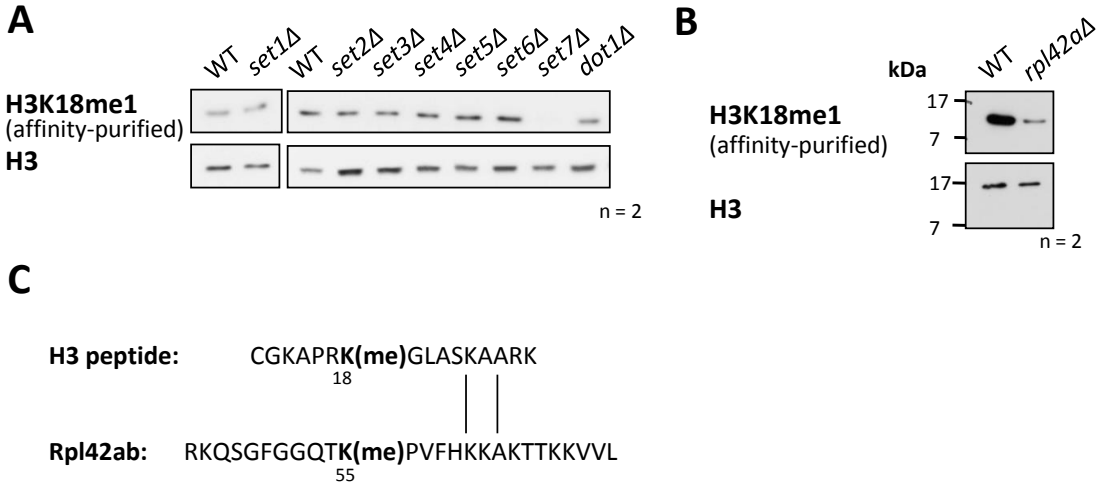


Figure 41. The first antibody raised against H3K18me1 recognises a methyl mark on Rpl42.

(A) A Western blot showing levels of the species recognised by the H3K18me1 antibody in the methyltransferase deletion strains indicated. The *SET1* deletion was performed in the W303 background. The remaining deletions were performed in the BY4741 background. H3 levels are shown as a loading control. **(B)** A Western blot showing the recognition of Rpl42a by the affinity-purified H3K18me1 antibody. H3 levels are shown as a loading control. **(C)** A sequence alignment of the methylated H3 peptide and the region surrounding the Set7 methyltransferase target residue in Rpl42.

6.2.2. Characterisation of the second antibody raised against H3K18me1

Because of the inability to study H3K18me1 using the first antibody by Western blotting, a second antibody was raised against H3K18me1. The monomethylated peptide corresponding to histone H3 used to immunise the rabbit was designed so that it was slightly longer than the first antigen used, with the methylated residue at the centre (TGGKAPR**K(me)**QLASKAARc, Figure 42A). As with the first antibody, the second antibody raised against H3K18me1 was specific for the methylated peptide on a dot blot (Figure 42B) and synthetic histone H3 monomethylated on lysine 18 by Western blot (Figure 42C). However, in contrast to the first antibody, when a whole cell extract from the K18A negative control strain was analysed by Western blot, no signal was observed (Figure 42D). Another experiment with the K18A strain demonstrated that the second H3K18me1 antibody is also specific for a modification on lysine 18 by ChIP (Figure 42E). A titration of increasing antibody amounts was performed and the conditions giving the highest recovery were chosen for use in subsequent ChIP experiments (data not shown). (In the following sections, only

the second H3K18me1 antibody was used for Western blot experiments but either antibody preparation was used for the ChIP experiments – this will be clearly stated in the figure legends.)

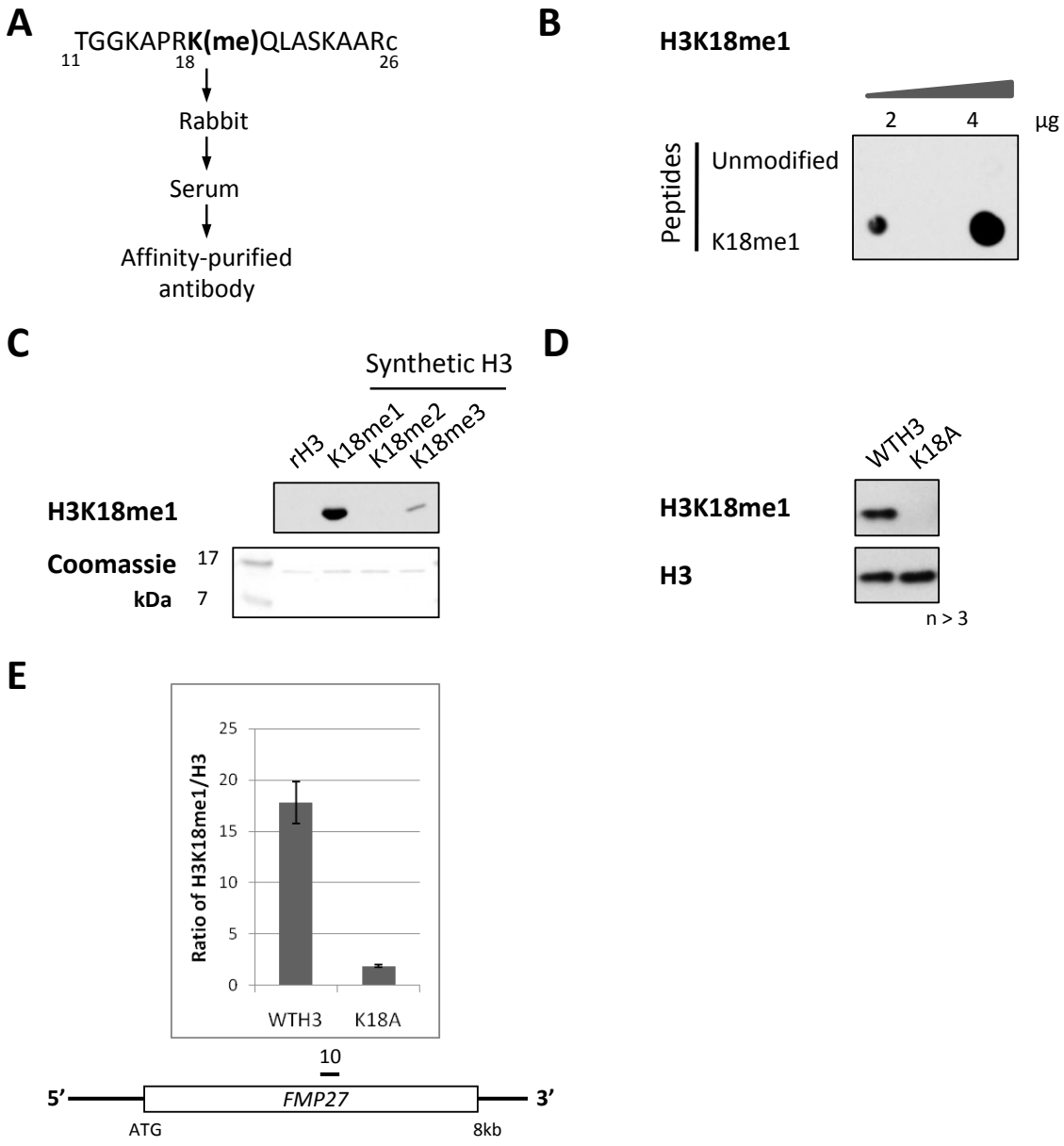


Figure 42. Characterisation of the second antibody raised against H3K18me1.

(A) The methylated peptide sequence used to raise the antibody in rabbit. Numbers indicate the corresponding residue in histone H3. A C-terminal cysteine (c) was included to facilitate peptide immobilisation in subsequent experiments. **(B)** A dot blot showing the specificity of the affinity-purified antibody for the mono-methylated over the unmodified H3 peptide at two amounts of peptide (2 and 4 µg). **(C)** Testing the specificity of the H3K18me1 antibody against synthetic histone H3 that has been mono-, di- or trimethylated on lysine 18 (kindly provided by R. Klose). Recombinant H3 (rH3, expressed and purified from *E. coli*) is unmodified. A Coomassie-stained gel is shown as a loading control. **(D)** Western blot analysis of H3K18me1 levels in whole cell extracts from a wildtype and K18A strain. H3 levels are shown as a loading control. **(E)** A ChIP-qPCR experiment showing the specificity of the H3K18me1 antibody for a modification on H3K18 at primer 10 on *FMP27*. Signals are normalised to levels of H3. Results shown are representative of at least three experiments. Error bars show standard errors of the real-time PCR reaction.

6.2.3. Distribution profile of H3K18me1 across genes

The distribution profile of H3K18me1 was assessed by ChIP across *FMP27*, a long, constitutively transcribed gene. As can be seen from Figure 43A, the distribution of this modification was relatively even across the length of the gene. A similar distribution was observed at *SEN1*, another long, constitutively active gene (data not shown). To investigate whether this was the case for all genes, a ChIP-sequencing experiment was performed. As detailed in the Materials and Methods, purified DNA that had been immunoprecipitated with the second H3K18me1 antibody from the WTH3 strain was sent for high-throughput DNA sequencing (Wellcome Trust Centre for Human Genetics, Oxford). The resulting sequences were mapped to the *S. cerevisiae* genome. Figure 43B shows the distribution of H3K18me1 across *FMP27* obtained by ChIP-sequencing. This profile is comparable to that obtained by conventional ChIP-qPCR (Figure 43A), thus validating the ChIP-seq data.

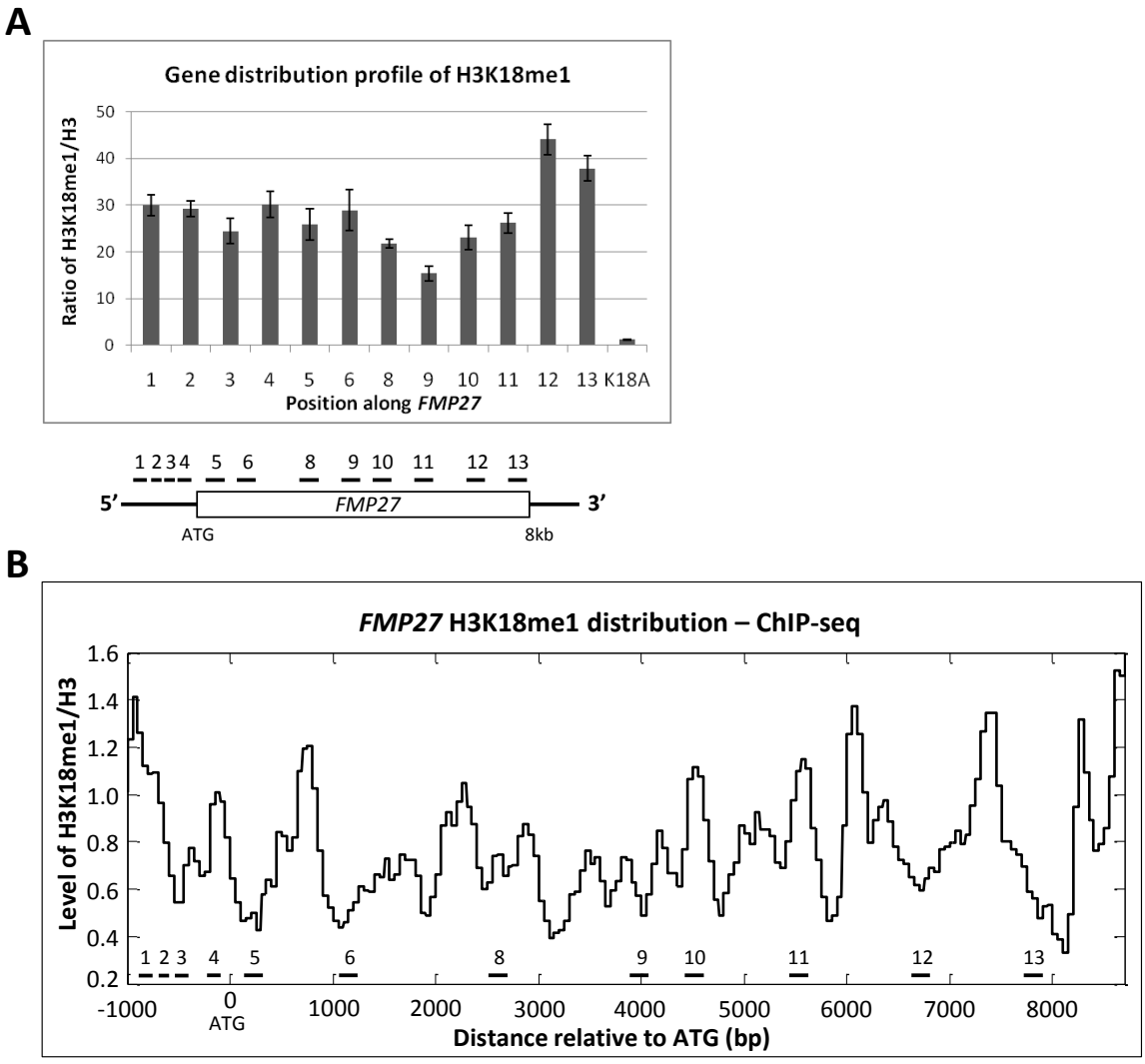


Figure 43. Distribution profile of H3K18me1 across *FMP27*.

(A) Levels of H3K18me1 assessed by ChIP-qPCR across the *FMP27* gene locus with the second H3K18me1 antibody. Primer positions are indicated on the locus map. Signals are normalised to levels of H3 at each site. Results shown are representative of at least two experiments. Error bars show standard errors of the real-time PCR reaction. **(B)** Distribution profile of H3K18me1 (normalised to H3 levels) across *FMP27* in the WTH3 strain extracted from the ChIP-sequencing dataset. The positions of primers used in the ChIP-qPCR experiment (A) are indicated for comparative purposes.

Following on from the distribution of H3K18me1 on *FMP27*, the average profile of H3K18me1 across all *S. cerevisiae* genes was generated (Figure 44A). This again showed that H3K18me1 is evenly distributed over the gene body but with a dip in H3K18me1 just upstream of the transcription start site. This profile of H3K18me1 across genes is very similar to the profile of histone H3 (Figure 44B). Indeed, the genome-wide Spearman rank correlation between H3K18me1 and histone H3 is 0.94, indicating that H3K18me1 is widespread throughout the

genome. As a comparison, the Spearman rank correlation between H3K4me3 and histone H3 is 0.5. However, when the average gene profile for H3K18me1 is normalised to histone H3 levels, this dip at the transcription start site remains, perhaps a result of the high H3K18ac in this region genome-wide (Liu et al. 2005) reducing the available sites of methylation (Figure 44C).

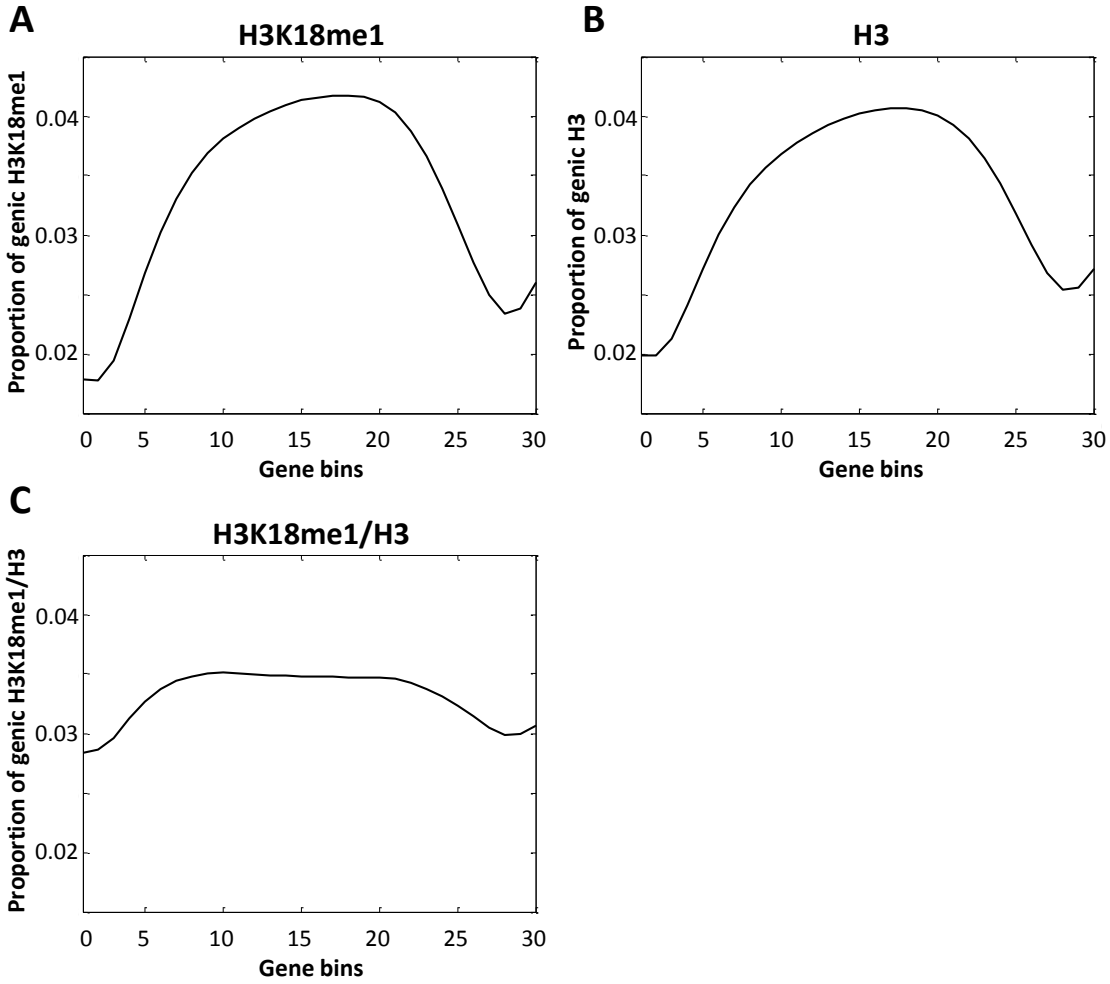


Figure 44. Distribution profile of H3K18me1 across genes.

(A) The average distribution profile of H3K18me1 across all genes in *S. cerevisiae.* Genes were divided into 30 bins and the level of H3K18me1 in each bin is expressed as a proportion of the total H3K18me1 across the gene (S. Murray). **(B)** The average distribution profile of histone H3 across all genes in *S. cerevisiae* (as (A)). **(C)** The average distribution profile of H3K18me1 after normalisation to histone H3 levels across all genes in *S. cerevisiae* (as (A)).

To investigate the link between H3K18me1 and H3 levels further, the correlation between H3K18me1 and H3 was calculated at three different regions of all protein-coding genes (Table 17). The correlation between H3K18me1 and H3 is slightly higher in promoter regions than in the 5'

coding region or over the whole transcription unit. Interestingly, promoter regions have higher rates of nucleosome turnover than the coding regions of genes (Dion et al. 2007) and yet the correlation between H3K18me1 and H3 remains high in this region, suggesting that either newly-incorporated histone H3 is already methylated on lysine 18 or that this methylation occurs very shortly after incorporation. In a strain in which nucleosome assembly is decreased (H3 K56A, (Li et al. 2008)), global H3K18me1 levels are also slightly reduced, potentially arguing against the possibility that histone H3 is methylated on lysine 18 before incorporation into chromatin. In contrast, H3K4me2 and H3K4me3 remain at wildtype levels in the K56A strain (Figure 45). Together these analyses suggest that levels of H3K18me1 are closely linked to the levels of histone H3.

Gene region	Correlation between H3K18me1 and H3
Transcription units	0.8497
Promoters (-300 to 0 bp)	0.8910
5’ coding regions (0 to +300 bp)	0.8186

Table 17. Pearson correlations between average H3K18me1 and H3 levels at three regions of protein-coding genes in the WTH3 strain.

Gene regions are numbered with respect to the transcription start site. H3K18me1 was not normalised to histone H3 levels.

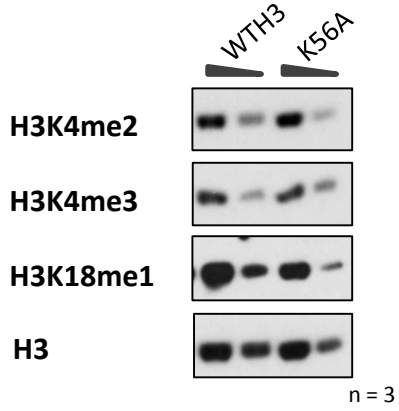


Figure 45. Mutation of H3K56 decreases global levels of H3K18me1.

A Western blot showing reduced levels of H3K18me1 in the wildtype and K56A strains. H3K4me2 and H3K4me3 levels are shown as a comparison and H3 levels are shown as a loading control. Two-fold dilutions of each extract were loaded to ensure the experiment was in the linear range of detection.

6.2.4. Exploring the relationship between H3K18me1 and transcription

Next it was investigated whether the distribution of H3K18me1 changes upon activation of previously repressed genes. To this end, *MET16* and *GAL1* were induced by changes in the growth medium. Monitoring of *MET16* RNA levels by reverse transcription indicated that the conditions used resulted in *MET16* activation (Figure 46A). Upon induction of *MET16*, there was a rapid loss of H3K18me1 at the 5' gene region followed by a slight recovery by 30 min (Figure 46B). Preliminary ChIP data suggested an increase in H3K18ac levels at the 5' gene region upon *MET16* induction (Figure 46C) and so this loss of H3K18me1 could reflect a possible switch with H3K18ac in this region. In contrast, at the 3' end of *MET16* there were only slight fluctuations in H3K18me1 upon gene induction.

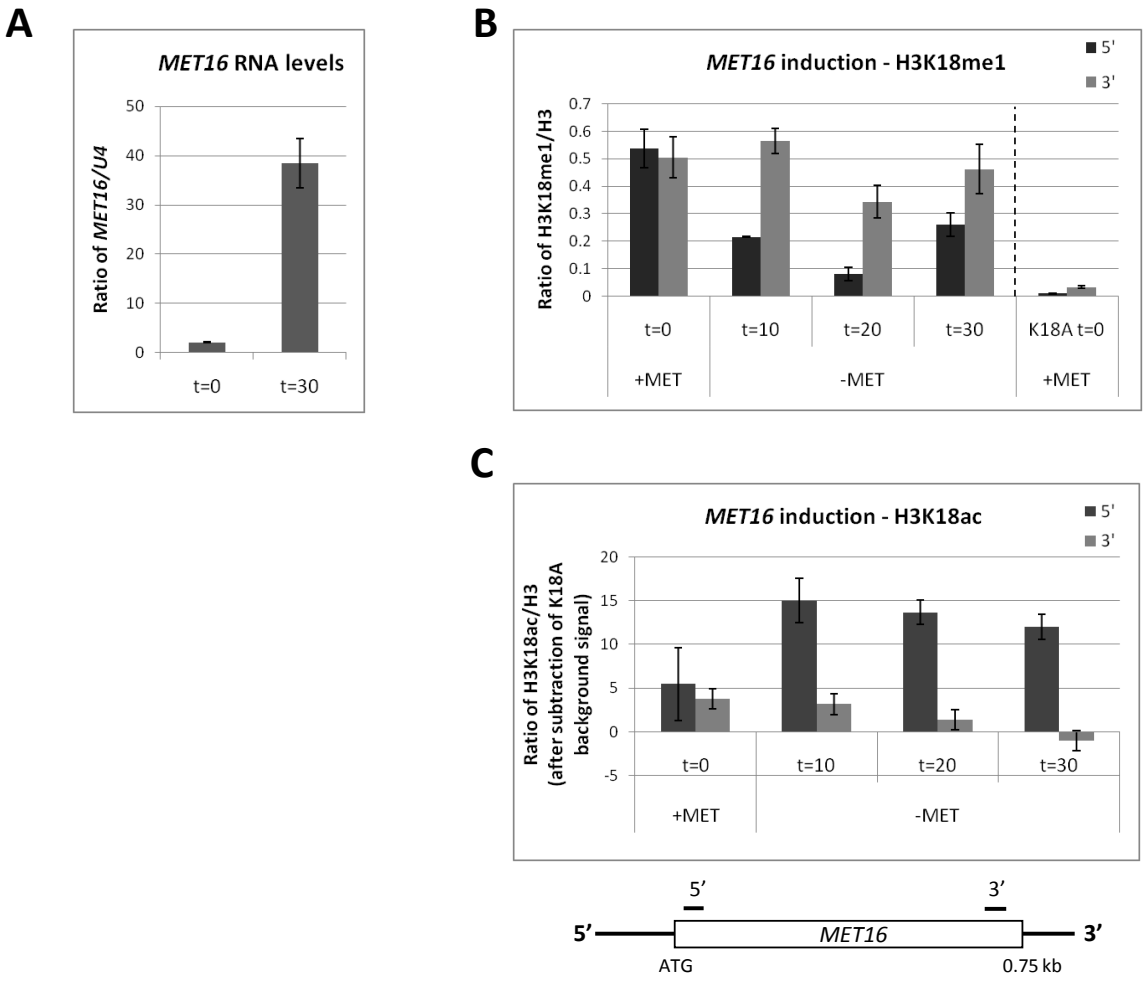


Figure 46. Changes in H3K18me1 levels upon modulation of *MET16* gene expression.

(A) Real-time PCR using the 5' *MET16* primers following reverse transcription of RNA from wildtype yeast grown in CSM (+Met, t=0) and after 30 min in media without methionine (t=30). Signals are normalised to levels of *U4* RNA. Error bars display the standard error of the real-time PCR reactions. **(B)** Wildtype yeast (WTH3) was grown to log phase in minimal media (CSM) followed by removal of methionine from the media (CSM –Met) to induce *MET16* at t=0 min. Samples were taken at 10 min intervals for 30 min. Levels of H3K18me1 (measured using the first antibody final bleed serum) in the K18A strain are shown as a control for antibody specificity. Signals at the gene positions indicated are normalised to H3 levels. Results are representative of three independent experiments. Error bars display the standard error of the real-time PCR reactions. **(C)** As (B) except for H3K18ac. Background signals of H3K18ac in the K18A negative control strain were subtracted from the signals in the WTH3 strain.

To discover whether this loss of H3K18me1 also occurred at other nutrient-regulated genes, *GAL1* was induced by changing the carbon source from glucose to galactose. There was an apparent loss of H3K18me1 across the length of the gene upon activation (Figure 47A). H3K18me1 was then restored when *GAL1* was re-repressed in glucose-containing media. Monitoring of *GAL1* sense transcript levels by Northern blotting confirmed the activation and subsequent repression of *GAL1* by the changing media conditions (Figure 47B, A. Serra Barros). Closer inspection of the *GAL1* histone H3 levels uncovered a marked increase in H3 upon gene activation (Figure 47C), which strongly influenced the normalised H3K18me1 results (before normalisation to H3, the H3K18me1 signals were relatively even throughout the experiment). This could indicate that the newly-deposited histones are not already methylated on lysine 18, in agreement with the results in section 6.2.3, and that the rapid histone turnover caused by the high levels of transcription of activated *GAL1* could prevent the re-stabilisation of H3K18me1 until the gene is re-repressed. Interestingly, no increase in H3 was observed on chromatin following *MET16* induction so 'dilution' of H3K18me1 cannot be the only explanation for the loss of H3K18me1 upon induction of these two nutrient-regulated genes.

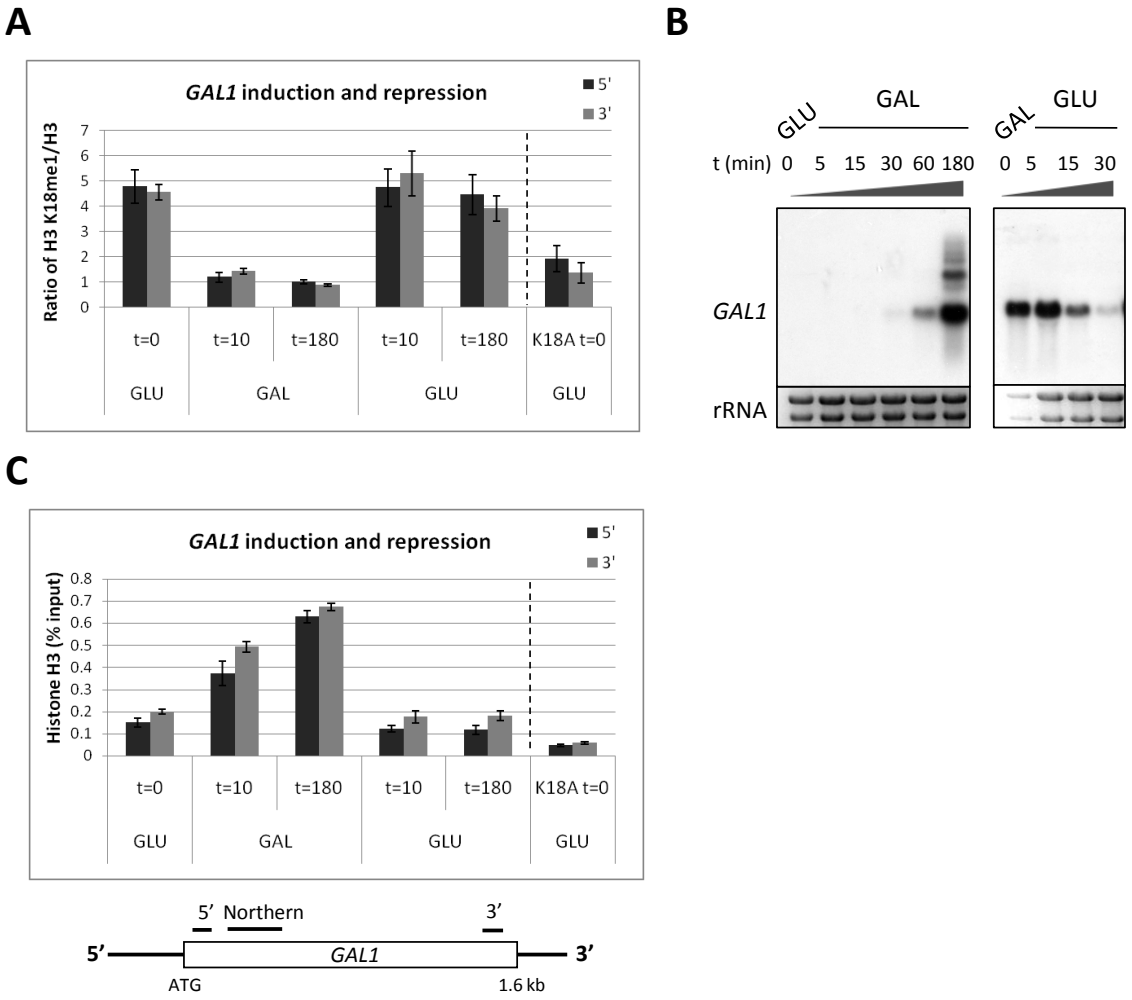


Figure 47. Changes in H3K18me1 levels upon modulation of *GAL1* gene expression.

(A) Wildtype yeast were grown for 4 h in YP glucose (Glu), followed by 3 h in YP galactose (Gal) to induce *GAL1* and then 3 h in YP glucose (Glu) to re-repress *GAL1.* Aliquots were taken at the times indicated (min) and analysed by ChIP-qPCR for H3K18me1 (first antibody) normalised to H3. ChIP primer (5' and 3') and Northern probe positions are indicated on the *GAL1* gene map (C). Results are representative of two independent experiments. Error bars display the standard error of the real-time PCR reactions. **(B)** Northern blot with RNA from wildtype yeast cells grown in glucose and galactose probed for the *GAL1* sense transcript. Samples were taken at the times indicated (min). An ethidium bromide-stained gel showing the levels of rRNA is included as a loading control (A. Serra Barros). **(C)** As (A) except for histone H3.

To identify whether there was a general inverse relationship between gene activity and levels of H3K18me1, the ChIP-sequencing dataset was analysed. There was only a very weak negative correlation (R = -0.080) between gene sense transcription (taken from the NET-seq dataset (Churchman and Weissman 2011)) and H3K18me1 levels (normalised to H3) across the whole transcription unit. However, when looking at the region surrounding the transcription start site, the dip in H3K18me1 in this region was more enhanced for genes falling into the class of high

sense transcription compared to those with medium and low sense transcription (Figure 48) and the correlation between H3K18me1 in this region and sense transcription is -0.2005. Therefore, both the gene-specific and genome-wide studies would suggest that the promoter levels of H3K18me1 correlate negatively with gene transcription. As mentioned before, this could reflect a higher level of H3K18ac at the transcription start sites of genes with greater transcriptional activity, thus providing competition for the site of modification.

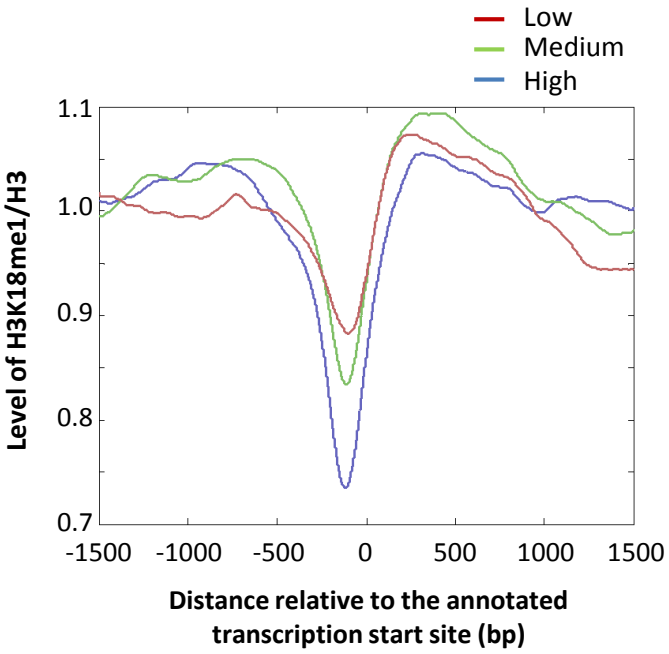


Figure 48. H3K18me1 promoter levels correlate negatively with sense transcription.

The distribution of H3K18me1 (normalised to H3 levels) across the transcription start site from the ChIP-seq dataset for genes with high (blue), medium (green) and low (red) sense transcription (transcription measurements taken from the NET-seq dataset (Churchman and Weisman, 2011). Genes were ranked according to sense transcription levels and divided into three equal groups for analysis.) (S. Murray).

6.2.5. Factors associated with genes with high and low levels of H3K18me1

To gain more insight into the function of H3K18me1 in chromatin, the factors enriched at genes with high and low average H3K18me1 levels were analysed. Genes were ranked according to their average level of H3K18me1 (normalised to histone H3) across the whole transcription unit and those with the top and bottom 200 rankings were used for further analysis (Figure 49). (Since there is such a high correlation between H3K18me1 and histone H3, the normalised

H3K18me1/H3 data was used for this analysis in order to identify the factors whose binding was influenced specifically by H3K18me1 rather than just nucleosome occupancy.)

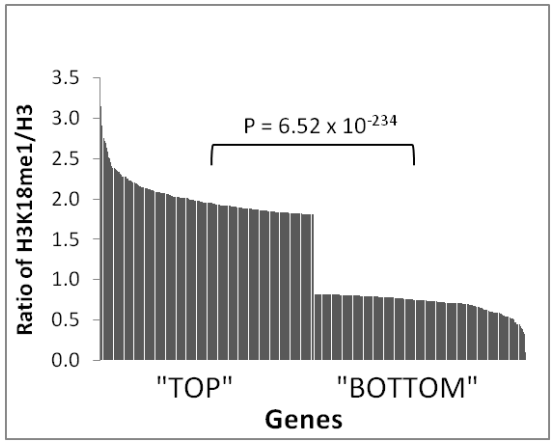


Figure 49. Genes with H3K18me1 enrichment and depletion in the WTH3 strain.

The levels of H3K18me1 in the WTH3 strain at the top and bottom 200 genes after ranking according to the level of the average H3K18me1 across the transcription unit as measured by ChIP-sequencing. Signals are normalised to levels of H3. P values are derived from student's t-test analysis.

This factor binding analysis makes use of the dataset by Venters et al. (2011) containing the genome-wide promoter-bound levels of 200 transcription-related factors. As displayed in Table 18 and Table 19, genes with the highest levels of H3K18me1 across the whole transcription unit are predominantly bound by chromatin remodellers (ISW1, SWR-C, FACT) and high levels of nucleosomes at their promoters. In contrast, genes with the lowest levels of H3K18me1 across the whole transcription unit have promoters that are bound by general transcription factors, TBP, transcriptional regulators (Ifh1, Rap1, Reb1) and enzymes depositing 'activating' chromatin modifications. Of note, Gcn5, which predominantly acetylates H3K18 and H3K23 *in vivo* (Jiang et al. 2007), is enriched at genes with low levels of H3K18me1. These differences in factors bound to genes with different levels of H3K18me1 suggest that H3K18me1 might be associated with a more repressed form of chromatin.

Factor	Complex	P value
Htz1	nucleosome	7.95E-07
Htb2	nucleosome	4.23E-05
Hta2	nucleosome	2.78E-04
Bre1	RAD6/BRE1	5.45E-04
Isw1	ISW1a, ISW1b	8.29E-04
Itc1	ISWI	1.42E-03
Bdf1	TFIID, SWR-C	2.32E-03
Rvb1	SWR-C, INO80	7.51E-03
Hht2	nucleosome	1.08E-02
Tfg2	TFIIF	1.21E-02
Elp3	ELONGATOR	1.21E-02
Rxt2	RPD3-L	1.34E-02
Bre2	COMPASS	1.78E-02
Mft1	THO	2.17E-02
TAF1	TFIID	2.62E-02
Rpd3	RPD3	2.67E-02
Rph1	RPH1	2.72E-02
Fhl1	FHL1, CRF1, FHL1/IFH1	2.79E-02
Pob3	FACT	2.88E-02
Not5	CCR4-NOT	3.03E-02
Chd1	SAGA, SLIK	3.06E-02
Hpr1	THO	3.42E-02
Epl1	NuA4, Piccolo	3.62E-02
Srb4	Mediator	4.00E-02

Table 18. Factors enriched at genes with the top 200 levels of H3K18me1/H3 across the transcription unit in the WTH3 strain.

The most significantly enriched ($p<0.05$) promoter-bound factors at genes with the top versus the bottom 200 levels of H3K18me1/H3 across the transcription unit. Genome-wide factor binding data is taken from Venters et al. (2011). P values are derived from the Wilcoxon rank-sum test (S. Murray).

Factor	Complex	P value
Ifh1	FHL1/IFH1	4.73E-06
Mot1	Mot1	9.87E-05
Rap1	RAP1	1.92E-04
Bur6	NC2	1.21E-03
Rpt6	APIS	2.03E-03
TBP	TFIID, SAGA, SLIK	2.13E-03
Med2	Mediator	2.88E-03
Nhp6a	NHP	7.33E-03
Gcn5	SAGA, SLIK, ADA	8.21E-03
Reb1	REB1	1.06E-02
Esa1	NuA4/Piccolo	1.12E-02
Jhd1	JHDM1	1.69E-02
Rad6	RAD6/BRE1	1.84E-02
Ume6	RPD3-L	2.11E-02
Tfc7	TFIIIC	2.15E-02
Ssl1	TFIIH	2.19E-02
Rfx1	RFX1	2.27E-02
Dst1	TFIIS	2.44E-02
Fkh1	FKH1	2.59E-02
Taf9	TFIID, SAGA, SLIK	2.96E-02
Ste12	STE12	3.11E-02
Ncb2	NC2	3.21E-02
Spt21	SPT10,21	3.66E-02
Rpb2	Pol II	3.82E-02
Taf5	TFIID, SAGA, SLIK	4.21E-02

Table 19. Factors enriched at genes with the bottom 200 levels of H3K18me1/H3 across the transcription unit in the WTH3 strain.

The most significantly enriched ($p<0.05$) promoter-bound factors at genes with the bottom versus the top 200 levels of H3K18me1/H3 across the transcription unit. Genome-wide factor binding data is taken from Venters et al. (2011). P values are derived from the Wilcoxon rank-sum test (S. Murray).

6.2.6. Gene ontology analysis of genes with the highest and lowest levels of H3K18me1

It is possible that genes with particularly high or low levels of H3K18me1 may have distinct functions. To explore this concept further, a gene ontology analysis was performed on the genes with the top and bottom 200 levels of H3K18me1 (normalised to histone H3) across the whole transcription unit in the WTH3 strain. As can be seen from Table 20 and Table 21, genes with the highest levels of H3K18me1/H3 across the whole transcription unit are those involved in translation, vitamin biosynthesis, response to abiotic stimulus and cell wall organisation. Genes with the lowest levels of H3K18me1/H3 are also involved in translation in addition to chromatin assembly, RNA transport and mitochondrion organisation. The presence of the 'translation' GO term on both lists was unexpected. However, closer inspection of the genes from the two classes

listed under this term yielded some differences: the genes with the lowest levels of H3K18me1/H3 are enriched for the ribosomal protein genes whereas the genes with the highest levels of H3K18me1/H3 are enriched for other, non-RPGs involved in translation.

Biological process	Count	%	P value
translational elongation	24	12.1	4.40E-06
translation	29	14.6	2.20E-03
vitamin metabolic process	7	3.5	6.30E-03
response to abiotic stimulus	16	8.1	1.70E-02
water-soluble vitamin metabolic process	6	3	2.00E-02
flocculation	3	1.5	3.20E-02
cell wall organisation	12	6.1	3.80E-02

Table 20. Gene ontology terms associated with the genes with the top 200 levels of H3K18me1/H3 in the WTH3 strain.

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the top 200 levels of H3K18me1 (after normalisation to H3) in the WTH3 strain. GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

Biological process	Count	%	P value
translation	29	14.8	1.30E-06
macromolecular complex assembly	16	8.2	1.90E-03
regulation of transcription, mating-type specific	3	1.5	2.60E-03
ribosome biogenesis	14	7.1	4.30E-03
macromolecular complex subunit organisation	17	8.7	9.10E-03
nucleosome assembly	4	2	9.40E-03
cellular macromolecular complex assembly	12	6.1	1.10E-02
reproductive developmental process	8	4.1	1.20E-02
nuclear export	7	3.6	1.60E-02
chromatin assembly	4	2	1.60E-02
RNA export from nucleus	6	3.1	1.80E-02
protein-DNA complex assembly	5	2.6	2.70E-02
translational initiation	4	2	3.10E-02
mitochondrion organisation	10	5.1	3.50E-02
RNA transport	6	3.1	3.60E-02

Table 21. Gene ontology terms associated with the genes with the bottom 200 levels of H3K18me1/H3 in the WTH3 strain.

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the bottom 200 levels of H3K18me1 (after normalisation to H3) in the WTH3 strain. GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

6.2.7. Searching for the H3K18me1 methyltransferase

The effect of H3K18 methylation cannot be studied genetically using the K18A strain because this mutation removes both the potential to acetylate and methylate this residue. Identification of the methyltransferase for H3K18 would greatly facilitate experiments to discover the role of H3K18me1 in chromatin and may also give insight into the regulation of H3K18me1 deposition. Seventy-two known and putative methyltransferases were identified as candidates for the H3K18me1 methyltransferase based on bioinformatic searches for characterised methyltransferase motifs (Petrossian and Clarke 2009). Chromatin was prepared from the sixty-seven deletion strains available (detailed in Table 29) and H3K18me1 levels were assessed by ChIP. No deletion strain showed a marked loss of H3K18me1 as would be expected if the deleted protein were the sole methyltransferase for H3K18 (Figure 50). This could suggest that there is a degree of redundancy or that the methyltransferase is an essential gene (and so a deletion strain is not available). Alternatively, the methyltransferase domain may not have been previously characterised and so would not appear in the bioinformatic searches performed.

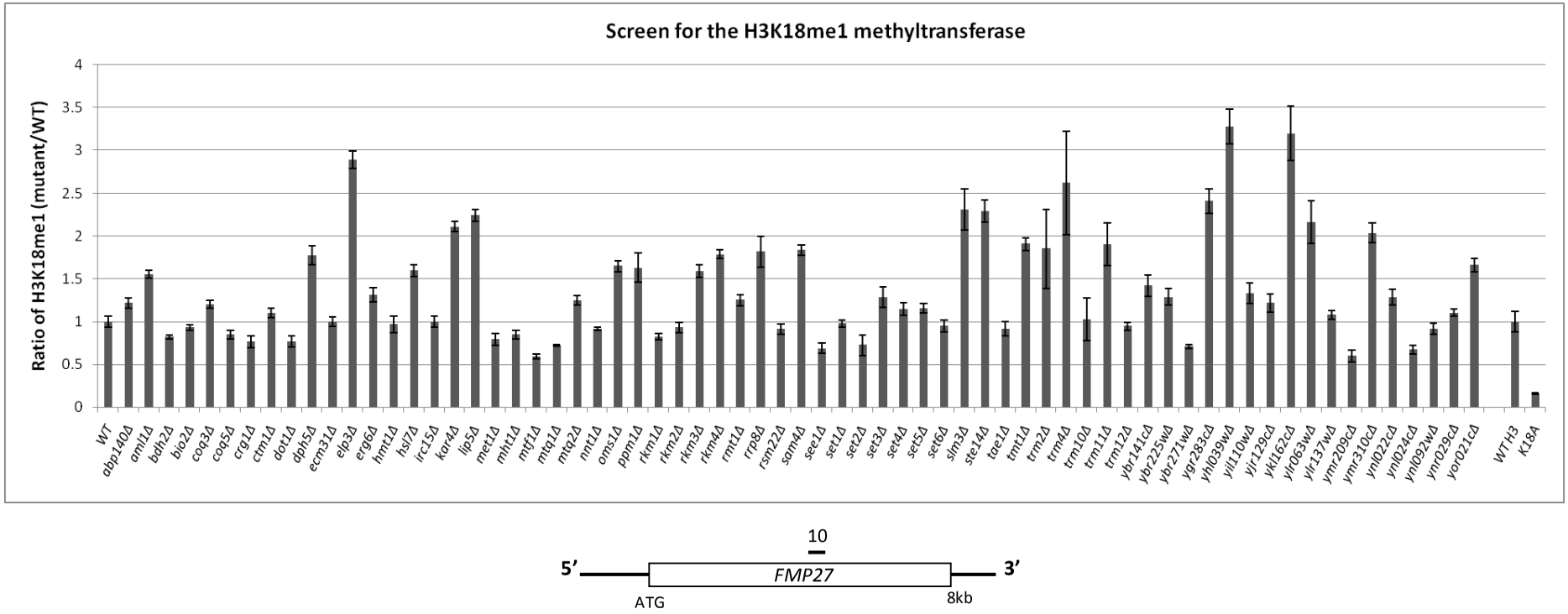


Figure 50. Screen for the H3K18me1 methyltransferase.

ChIP-qPCR experiments were performed using the deletion strains indicated and the final bleed serum preparation of the first H3K18me1 antibody. Real-time PCR primers were located in the coding region of *FMP27* (primer 10, as displayed on the gene map). H3K18me1 signals (without normalisation to H3) are normalised to the wildtype (BY4741) signal obtained with each batch of strains tested. Error bars display the standard error for the real-time PCR reaction.

Whilst no deletions substantially reduced the levels of H3K18me1, several gene deletions resulted in a more than two-fold increase in H3K18me1 compared to the wildtype. These genes are listed in Table 22. It is likely that the proteins encoded by these genes are having indirect effects on H3K18me1 levels. It should be noted that the H3K18me1 levels measured during this screen were not normalised to levels of histone H3. Therefore, given the high positive correlation between H3K18me1 and histone H3 levels, if any of the gene deletions caused increases in H3 in chromatin, this may have raised the absolute levels of H3K18me1 but no effect may have been seen if normalisation to H3 levels had been performed. For example, deletion of *ELP3*, a subunit of the Elongator complex with histone acetyltransferase activity (Wittschieben et al. 1999) could result in defective transcriptional elongation, lower histone turnover and so higher levels of H3K18me1. Alternatively, Elp3 acetyltransferase activity is predominantly directed towards H3K14 and H4K8 (Winkler et al. 2002), which might be involved in crosstalk with H3K18me1 (see section 6.2.9). Future studies would be needed to investigate the mechanisms by which these genes are influencing H3K18me1 levels.

Systematic gene name	Alias	Description
YBL024W	*TRM4*	S-adenosylmethionine-dependent tRNA methyltransferase
YCL055W	*KAR4*	transcription factor required for gene regulation in response to pheromones
YDL033C	*SLM3*	tRNA-specific 2-thiouridylase
YDR410C	*STE14*	farnesyl cysteine-carboxyl methyltransferase
YGR283C	-	protein of unknown function
YHL039W	-	SET-domain putative protein of unknown function
YKL162C	-	putative protein of unknown function
YLR063W	-	putative S-adenosylmethionine-dependent methyltransferase
YMR310C	-	putative protein of unknown function
YOR196C	*LIP5*	protein involved in biosynthesis of the coenzyme lipoic acid
YPL086C	*ELP3*	subunit of Elongator complex with HAT activity

Table 22. Gene deletions resulting in a greater than two-fold increase in H3K18me1 compared to wildtype.

All descriptions are taken from the *Saccharomyces* genome database (SGD: www.yeastgenome.org).

6.2.8. Searching for the H3K18me1 demethylase

The loss of H3K18me1 observed upon gene activation at *MET16* and *GAL1* could be a result of histone demethylation or histone turnover. To identify a possible H3K18 demethylase, a screen of the six yeast JmjC domain-containing demethylases was performed by ChIP (Figure 51A). An increase in H3K18me1 was observed in the strain lacking *RPH1*, a gene encoding a characterised H3K36 demethylase (Kim and Buratowski 2007; Tu, S et al. 2007). To investigate this effect further, a strain in which the endogenous *RPH1* gene is HA-tagged and placed under the control of the *GAL1* promoter was used. This strain was validated by monitoring the expression of Rph1-HA protein and the levels of H3K36me3 by Western blot (Figure 51B): Rph1-HA was only detected in cells grown in the presence of galactose and this Rph1-HA overexpression resulted in decreased H3K36me3. In the absence of Rph1-HA (achieved by growing the cells in glucose), a slight global increase in H3K36me3 was observed when taking into account the lower levels of H3 protein loaded in that lane. This small increase in histone methylation upon removal of a demethylase is in line with other reports in the literature (Tu, S et al. 2007; Kwon and Ahn 2011). The validated *pGAL1-3HA-RPH1* strain was then employed to test the effect of *RPH1* depletion and overexpression on H3K18me1 by ChIP at the 3' end of *FMP27*, where levels of H3K18me1 in the wildtype strain are high (Figure 51C). As noted in the initial screen and Figure 51B, depletion of Rph1 protein resulted in an increased level of H3K18me1 and H3K36me3. However, whilst there is a decrease in H3K18me1 upon overexpression of *RPH1*, this is not as great as the reduction in the characterised Rph1 substrate, H3K36me3.

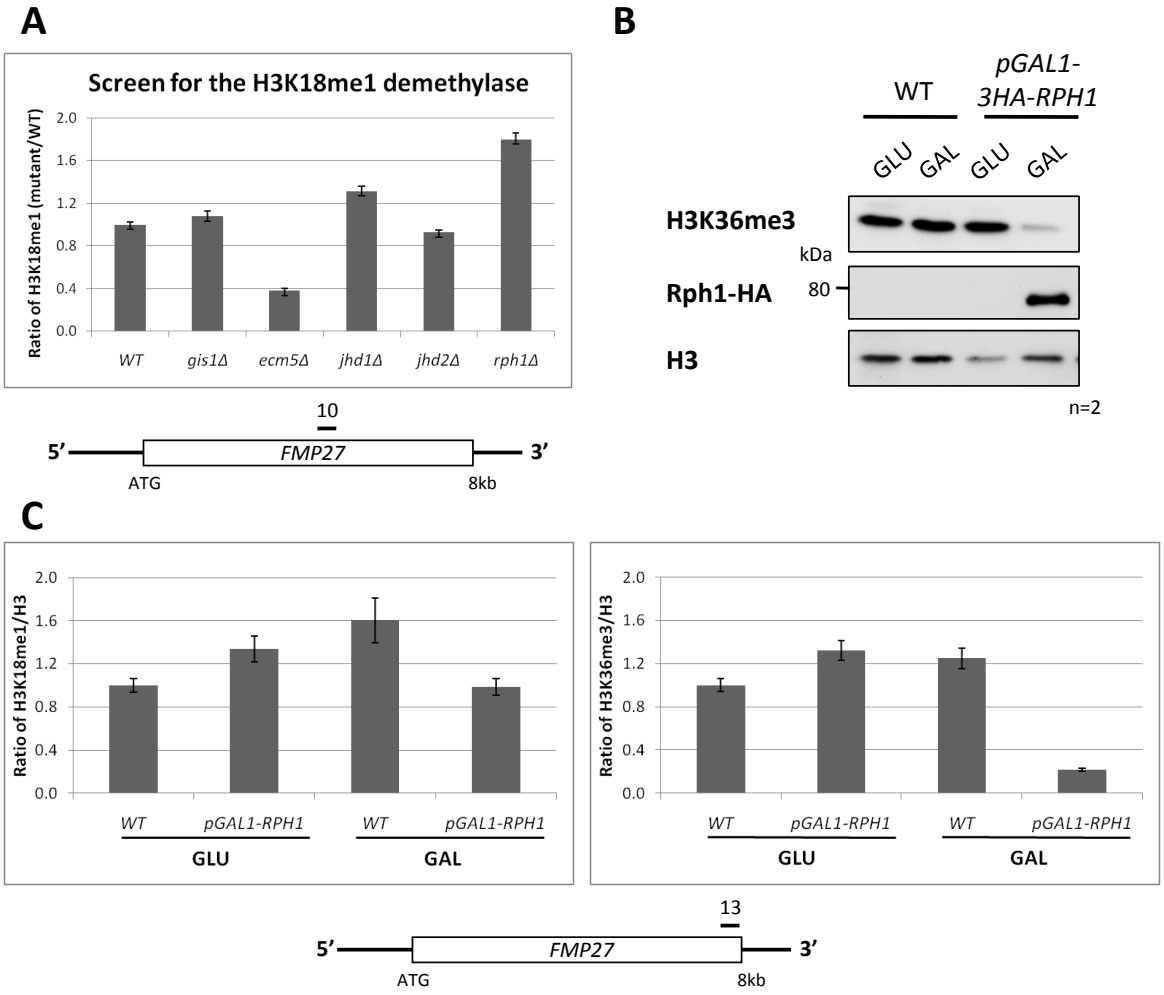


Figure 51. Searching for the H3K18me1 demethylase.

(A) Screen for the H3K18me1 demethylase by ChIP-qPCR with the first H3K18me1 antibody in the indicated deletion strains at *FMP27* primer 10 (indicated on the locus map). H3K18me1 signals (without normalisation to H3) are normalised to the wild-type (BY4741) signal. Error bars display the standard error for the real-time PCR reaction. **(B)** Validation of the *pGAL1-3HA-RPH1* strain by monitoring H3K36me3 and Rph1-HA levels in whole cell extracts from cells grown in YP glucose or YP galactose by Western blot. H3 levels are used as a loading control. **(C)** Testing the effect of *RPH1* depletion and overexpression, as controlled by the carbon source in the media, on H3K18me1 and H3K36me3 by ChIP-qPCR at *FMP27* primer 13 (indicated on the locus map). Signals are normalised to levels of H3. Error bars show standard errors of the real-time PCR reaction.

In order to assess whether Rph1 is acting directly on H3K18me1, an *in vitro* demethylase assay was performed. Wildtype Rph1 and a catalytic mutant of Rph1 (H235A, a mutation that prevents the coordination of the Fe^{2+} ion required for catalysis) were expressed in and purified from *E. coli* using the C-terminal His_6 tag as detailed in the Materials and Methods. The purified protein samples used in the assay are shown in Figure 52A. Wildtype and catalytically mutant Rph1 were

then tested *in vitro* against histones purified from calf thymus and synthetic histone H3 monomethylated on lysine 18 (Figure 52B). EDTA, an inhibitor of JmjC demethylase activity, was included in some reactions as a control. In contrast to H3K36me3, which displayed a small but reproducible decrease in levels in the reaction with wildtype Rph1 but not any of the control reactions, H3K18me1 levels did not decrease with wildtype Rph1. Thus it can be concluded that Rph1 is not directly affecting the levels of H3K18me1 but is acting indirectly, perhaps via H3K36 methylation. Evidence in support of this hypothesis is that H3K18me1 levels decrease slightly in the absence of Set2, the methyltransferase for H3K36 (Figure 50).

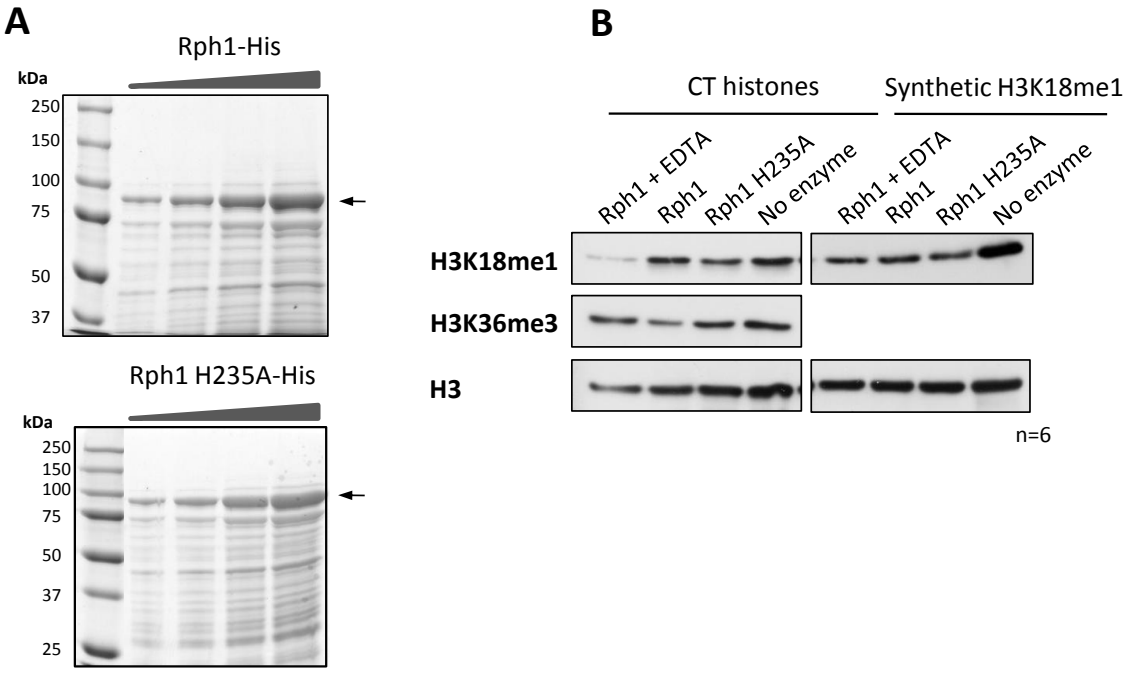


Figure 52. Rph1 does not directly demethylate H3K18.

(A) Coomassie-stained SDS-polyacrylamide gels showing recombinant wildtype and catalytically mutant (H235A) Rph1 protein after Ni^{2+} purification of the His_6-tagged proteins from *E. coli* (indicated by the arrows). **(B)** Histone demethylase assay using calf thymus (CT) histones or synthetic histone H3 with lysine 18 methylated (kindly provided by R. Klose). EDTA, a demethylase inhibitor, was included where indicated at 10 mM final concentration. Western blots were probed for H3K18me1, H3K36me3 and H3 (as a loading control).

6.2.9. Investigating the crosstalk between H3K14 and H3K18me1

As mentioned in the introduction to this chapter, the impetus for studying H3K18me1 came from the observation that there were undetectable levels of H3K18ac in the strains in which H3K14 had been substituted for alanine, arginine and glutamine (Chapter 3, and reproduced in Figure 53A). It would therefore be interesting to establish the effect of H3K14 substitution on the levels of H3K18me1. A Western blot with extracts from the H3K14 substitution strains indicates that mutation of H3K14 did not appear to affect global levels of H3K18me1 (Figure 53B). However, the dot blot in Figure 53C demonstrates that mutation of H3K14 to alanine slightly reduces the ability of the antibody raised against H3K18me1 to recognise its epitope. This would suggest that H3K18me1 levels may be slightly higher in the K14A strain than is apparent on the Western blot in Figure 53B. Interestingly, the K14R substitution does not affect antibody binding to H3K18me1 (Figure 53C).

Next, the levels of H3K18me1 were assessed by ChIP at *FMP27* in the wildtype and K14A strains (Figure 53D). In contrast to the Western blot results, the levels of H3K18me1 were higher in the K14A strain relative to the wildtype at *FMP27,* especially in the promoter and 5' coding region of the gene (primers 3 and 5). In addition to the epitope effects, this relatively localised effect of H3K14 substitution on H3K18me1 may account for the lack of any observable increase in H3K18me1 in the K14A strain at a global level. Following this finding, the levels of H3K18me1 in the K14R and K14Q strains were also assessed by ChIP (Figure 53E). Again in contrast to what was observed globally, the levels of H3K18me1 were higher in the K14A and K14Q strains than the WTH3 strain at the 5' region of *FMP27*. Interestingly, there were near-wildtype levels of H3K18me1 in the K14R strain. Given that H3K18ac was undetectable in all H3K14 mutants, independent of the residue substituted whereas H3K18me1 levels varied with the substituted residue, it seems unlikely that there is a simple acetyl/methyl switch between these two modifications at lysine 18 that relies solely on competition for the modification site.

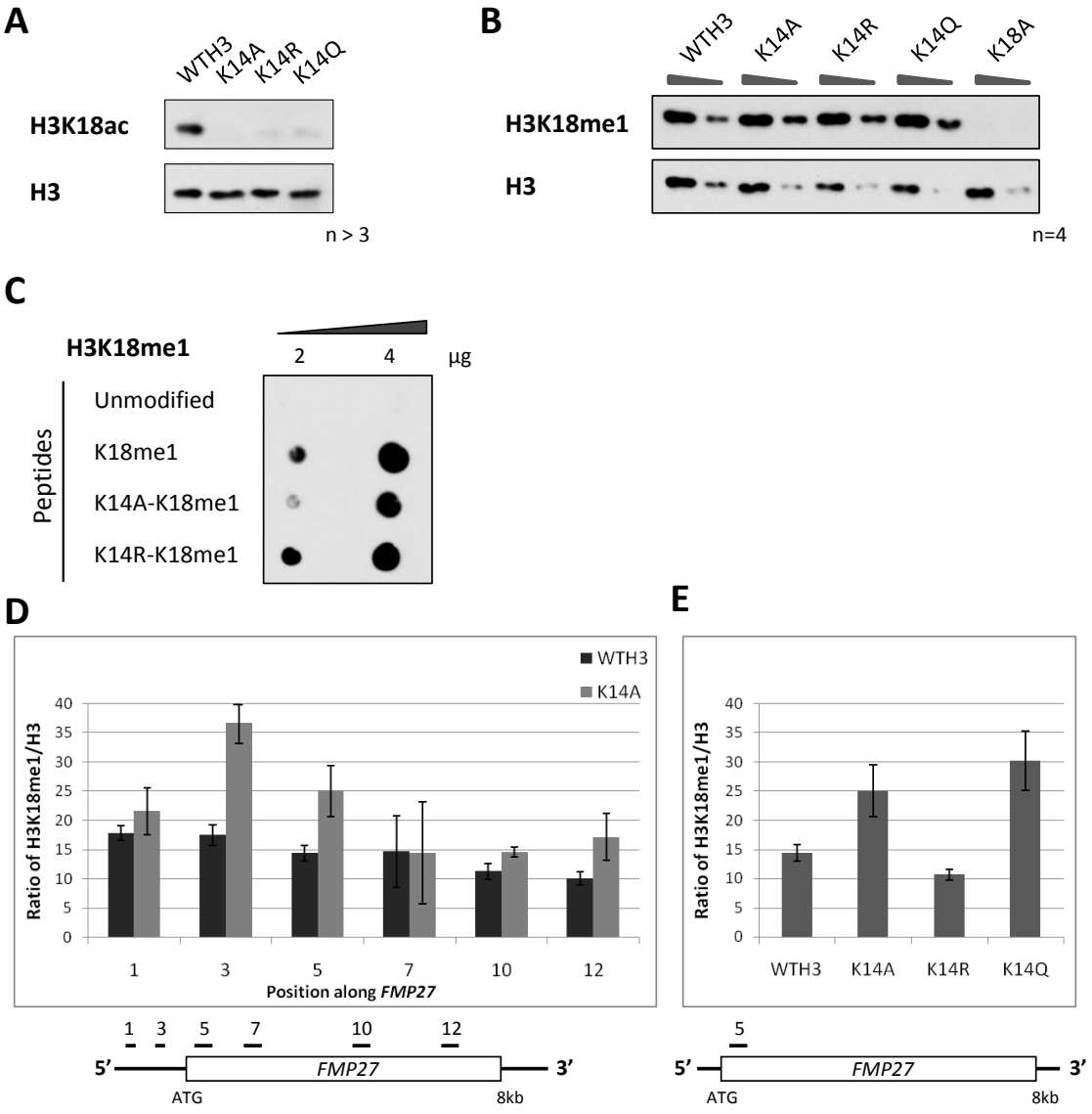


Figure 53. Levels of H3K18me1 are affected by mutation of H3K14.

(A) A Western blot showing levels of H3K18ac in whole cell extracts from the WTH3 and H3K14 substitution strains (reproduced from Figure 13). H3 levels are shown as a loading control. **(B)** As (A) for H3K18me1. An extract from the K18A strain has been included as a negative control. Two-fold dilutions of each extract were loaded to ensure the experiment was in the linear range of detection. **(C)** A dot blot showing slightly diminished binding of the second H3K18me1 antibody to the alanine but not the arginine K14-substituted methylated H3 peptide. **(D)** Levels of H3K18me1 (second antibody) by ChIP-qPCR at *FMP27* in the K14A strain compared to the WTH3 strain. Primer positions are indicated on the locus map. Signals are normalised to levels of H3. Results shown are representative of at least three experiments. Error bars show standard errors of the real-time PCR reaction. **(E)** As (D) except with all three H3K14 substitution strains at primer 5 in *FMP27* (indicated on locus map).

Next it was investigated whether the K14A substitution affected the changes in H3K18me1 at *MET16* upon induction of the gene. As can be seen from Figure 54A, whereas the levels of H3K18me1 decrease upon induction in the WTH3 strain, substitution of H3K14 with alanine caused the levels of H3K18me1 to increase upon *MET16* induction at both the 5' and 3' ends of the gene. Preliminary results suggest that H3K18ac would normally increase in the wildtype strain at both the 5' and 3' of *MET16* during gene induction (Figure 47C). The absence of H3K18ac in the K14A strain may remove a potential constraint on H3K18me1 levels, so that instead of decreasing upon gene induction, the levels of H3K18me1 increase. Interestingly, these changes in histone modification levels in the K14A strain do not appear to affect the levels of *MET16* RNA by 30 min, although it is possible that the initial induction kinetics may be slower (Figure 54B).

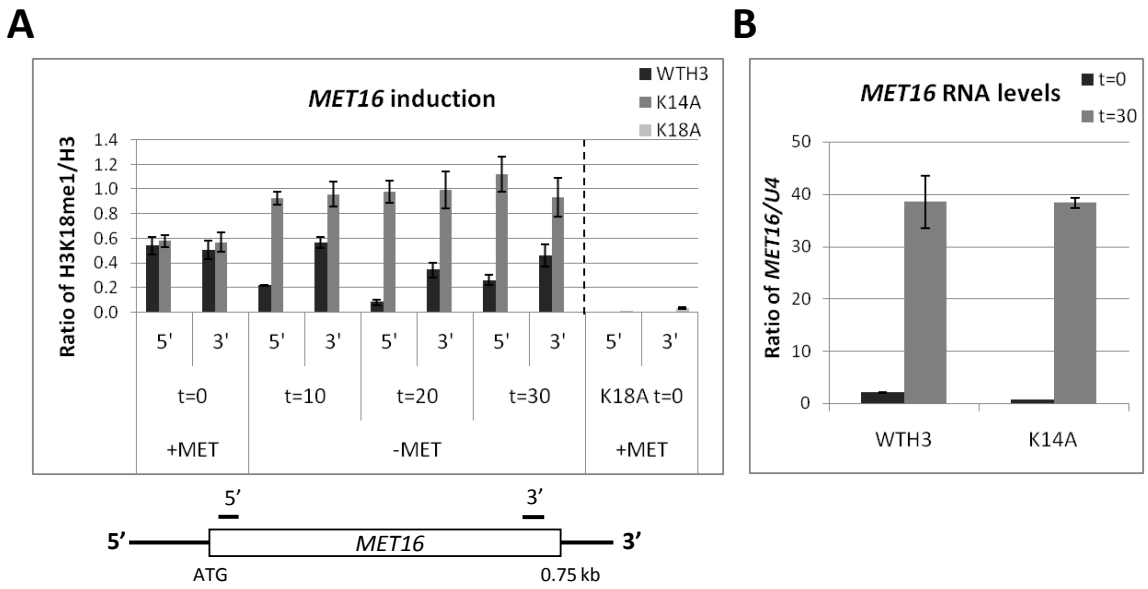


Figure 54. Investigating the effect of H3K14 substitution on levels of H3K18me1 during the induction of *MET16.*

(A) Wildtype (WTH3) and K14A strains were grown to log phase in minimal media (CSM, +Met) followed by removal of methionine from the media (CSM –Met) to induce *MET16* at t=0 min. Samples were taken at 10 min intervals for 30 min. Levels of H3K18me1 (measured using the first antibody final bleed serum) in the K18A strain are shown as a control for antibody specificity. Signals at the gene positions indicated are normalised to H3 levels. Results are representative of two independent experiments. Error bars display the standard error of the real-time PCR reactions. **(B)** Real time PCR using the 5' *MET16* primers following reverse transcription of RNA from the wildtype and K14A strain grown in CSM (+Met) and after 30 min in media without methionine. Signals are normalised to levels of *U4* RNA. Error bars display the standard error of the real-time PCR reactions.

6.2.10. Investigating a possible crosstalk between H3K4me3 and H3K18me1

Comparison of the levels of H3K4me3 and H3K18me1 at the 5' end of *FMP27* in the H3K14 substitution strains indicated that these two modifications may be reciprocal (shown previously in Figure 10C and Figure 53E, but reproduced in Figure 55A): the H3K14 substitutions leading to the highest levels of H3K18me1 (K14A and K14Q) were strains with the lowest levels of H3K4me3 whereas the K14R strain had near-wildtype levels of both H3K18me1 and H3K4me3. However, it is unclear whether there is any causality in this anti-correlation and/or whether there is any directionality in the H3K4-H3K18 interaction. To distinguish between a direct reciprocity between H3K4me3 and H3K18me1 and indirect opposite regulation of these two modifications by H3K14 (as depicted in Figure 55B), H3K18me1 levels were evaluated in stains with varying levels of H3K4me3 (Figure 55C, D). Deletion of *BRE2* and *SPP1* resulted in complete or partial loss of H3K4me3 respectively but did not affect the global levels of H3K18me1. Furthermore, deletion of *JHD2* caused a slight global increase in H3K4me3 but again did not affect H3K18me1. Thus it seems that levels of H3K4me3 alone do not influence global levels of H3K18me1. Unfortunately it is not yet possible to perform the reverse experiment to determine whether specific removal or enhancement of H3K18me1 affects H3K4me3 since the methyltransferase and demethylase for H3K18 have not been identified. However, leucine has been used in another study as a mimic of methylated lysine (Hyland et al. 2011). To this end, H3K18 was substituted with leucine and the levels of H3K4me3 were assessed (Figure 55E). The K18L strain had wildtype levels of H3K4me3, suggesting that H3K18 methylation would not directly influence H3K4me3. Thus, any reciprocity between H3K4me3 and H3K18me1 most likely only occurs as a result of opposite regulation by H3K14.

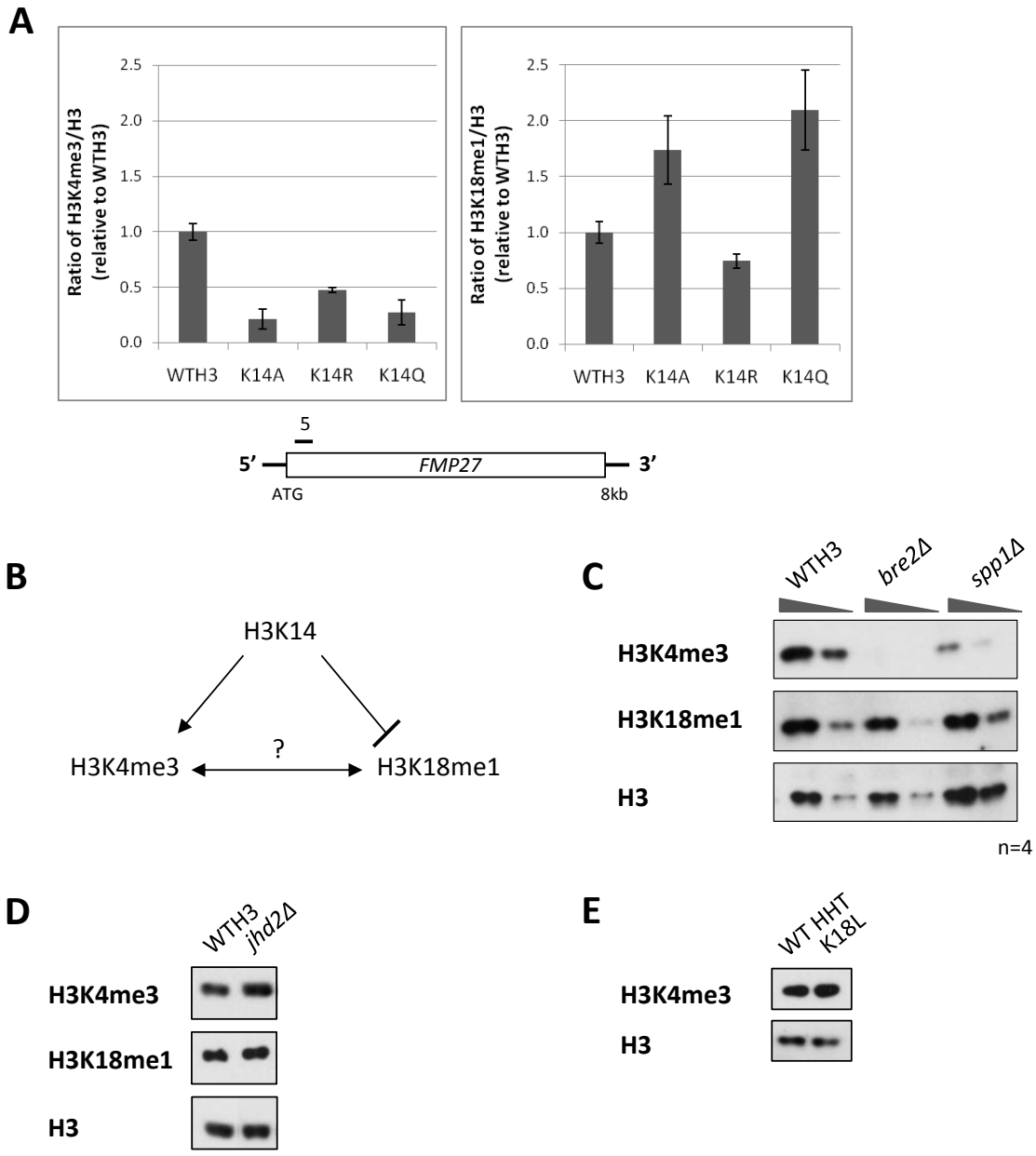


Figure 55. Investigating a potential crosstalk between H3K18me1 and H3K4me3.

(A) ChIP-qPCR results reproduced from Figure 10C and Figure 53E for ease of comparison. See original figure legends for experimental details. **(B)** A schematic illustrating the possible crosstalk between H3K4me3 and H3K18me1 involving H3K14. **(C)** Levels of H3K4me3 and H3K18me1 by Western blot in the wildtype strain (WTH3) and strains lacking *BRE2* and *SPP1*. Two-fold dilutions of each extract were loaded to ensure the experiment was in the linear range of detection. Histone H3 levels are shown as a loading control. **(D)** Levels of H3K4me3 and H3K18me1 by Western blot in the wildtype strain (WTH3) and strain lacking *JHD2*. Histone H3 levels are shown as a loading control. **(E)** Levels of H3K4me3 by Western blot in the wildtype strain (WT HHT) and a strain in which H3K18 had been substituted with leucine. Histone H3 levels are shown as a loading control.

6.2.11. H3P16 may not influence the levels of H3K18me1

Since H3K14 can influence levels of both H3K4me3 and H3K18me1, it was investigated whether H3P16, which is also important for H3K4me3 (Chapter 4), can additionally regulate H3K18me1. By Western blot, there was a slight decrease in the global level of H3K18me1 in whole cell extracts from the P16V strain (Figure 56A), a trend that was also observed by ChIP at *FMP27* (Figure 56B). However, the P16V mutation diminished the ability of the H3K18me1 antibody to recognise its epitope on a peptide, as tested by dot blot (Figure 56C). Therefore, the small decreases in H3K18me1 observed both globally and at *FMP27* are most probably caused by the attenuation of antibody binding when the neighbouring H3P16 has been mutated. Further investigations into this potential crosstalk would have to be carried out using mass spectrometry.

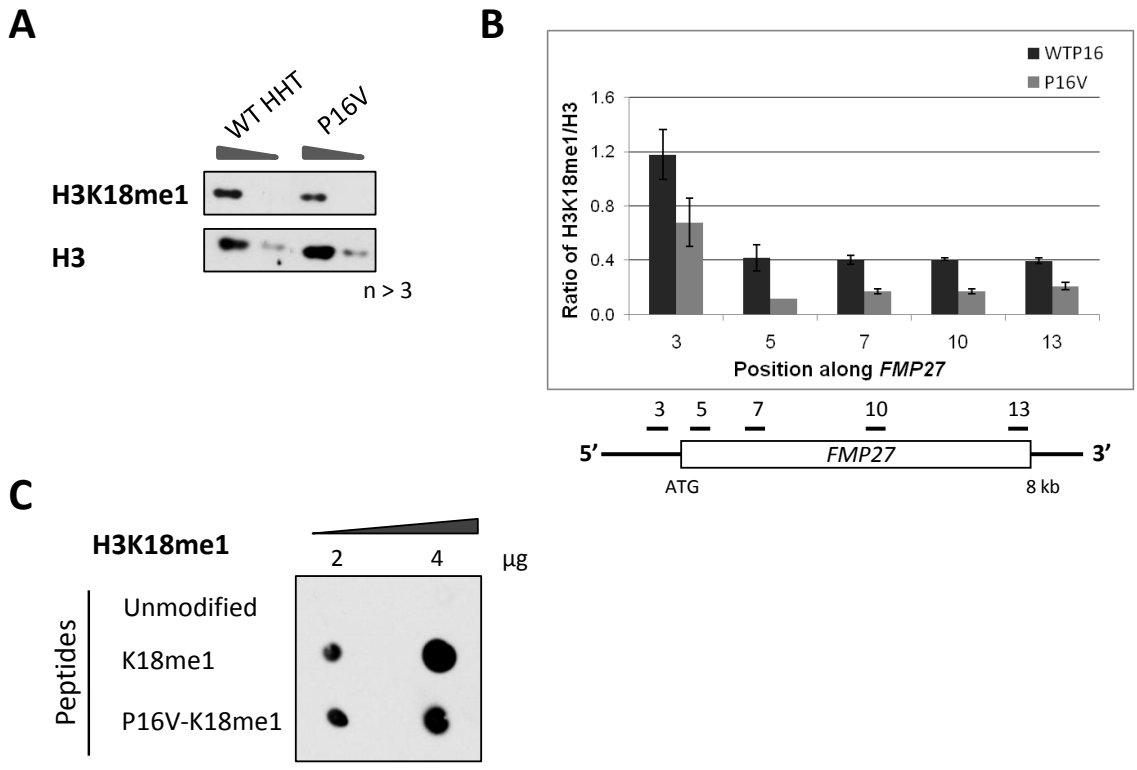


Figure 56. Levels of H3K18me1 are slightly affected by mutation of H3P16.

(A) A Western blot showing levels of H3K18me1 in the wildtype (WT HHT) and P16V strain. Two-fold dilutions of each extract were loaded to ensure the experiment was in the linear range of detection. Histone H3 levels are included as a loading control. **(B)** A ChIP-qPCR experiment at *FMP27* confirming the reduction in H3K18me1 in the P16V strain. Primer positions are indicated on the locus map. Signals are normalised to levels of histone H3. Results are representative of two independent experiments. Error bars display the standard error of the real-time PCR reactions. **(C)** A dot blot showing the reduction in H3K18me1 antibody recognition of a methylated H3 peptide in which proline 16 had been mutated to valine at two amounts of peptide (2 and 4 µg).

6.2.12. Investigating a possible mechanism for the H3K14-H3K4 crosstalk involving H3K18me1

One potential mechanism for the crosstalk between H3K14 and H3K4me3 could involve H3K18me1 and Jhd2. Experiments in Chapter 5 suggested that the H3K14 mutants were more susceptible to Jhd2 action than the wildtype strain. It is possible that the higher levels of H3K18me1 in the K14A/K14Q strains could influence the positioning or activation of Jhd2 on the histone H3 N-terminal tail for optimal H3K4me3 demethylation. This hypothesis is supported by the fact that Jhd2 has a PHD finger, a characterised lysine methyl-binding domain (Shi et al. 2007) that could conceivably bind H3K18me1. Indeed, Jhd2 was a hit in a preliminary pull-down mass spectrometry experiment designed to identify proteins associated with a histone H3 peptide methylated on lysine 18 (D. Clynes). A model for this potential mechanism of crosstalk is shown in Figure 57A. To explore the possibility that H3K18me1 is recruiting/positioning Jhd2 on the H3 tail via its PHD finger to repress H3K4me3, a peptide pull-down experiment was performed with K18-methylated and unmethylated H3 peptides and cell extracts from strains expressing Myc-tagged wildtype Jhd2 or Jhd2 with a mutation in the key residue in the PHD finger important for binding methyl-lysine (H261A) (Figure 57B). However, the wildtype protein displayed no selectivity for binding to the methylated over the unmethylated peptide and this binding was not diminished by mutation of the PHD finger. Therefore it seems unlikely that an H3K18me1-mediated Jhd2 recruitment mechanism is operating in the crosstalk between H3K14 and H3K4me3.

To examine if there was any involvement of H3K18 in the crosstalk between H3K14/H3P16 and H3K4me3, double histone H3 point mutants were created in which H3K18 was mutated to alanine in the H3K14 and H3P16 substitution strains. A Western blot was performed to monitor global levels of H3K4me3 in these strains (Figure 57C). Whilst there was some recovery of H3K4me3 in the double mutant strains, the varying effects of the three H3K14 substitutions were still evident. Thus H3K18, or a modification of that residue, can influence levels of H3K4me3 in the H3K14/H3P16 mutant strains but removal of H3K18 does not restore the levels of H3K4me3 to

wildtype. It is tempting to draw comparisons between the similar effects of H3K18 mutation and *JHD2* depletion on H3K4me3 (Chapter 5, Figure 34A) in the H3K14 mutant strains. Despite the lack of specific Jhd2 interaction with the K18-methylated peptide (Figure 57B), it is still possible that H3K18, or a modification on this residue may be important for positive regulation of Jhd2 positioning/activity. If H3K18 were important for Jhd2 activity *in vivo*, it could be predicted that mutation of H3K18 would, like deletion of *JHD2*, partially rescue the reduction in H3K4me3 in the absence of *SPP1*. Although no such rescue was observed (Figure 57D), making it seem unlikely that Jhd2 activity is influenced by H3K18, one caveat to this experiment is that the effects from the combined loss of H3K18/H3K18ac/H3K18me1 may cancel each other out, resulting in no net effect on Jhd2 activity.

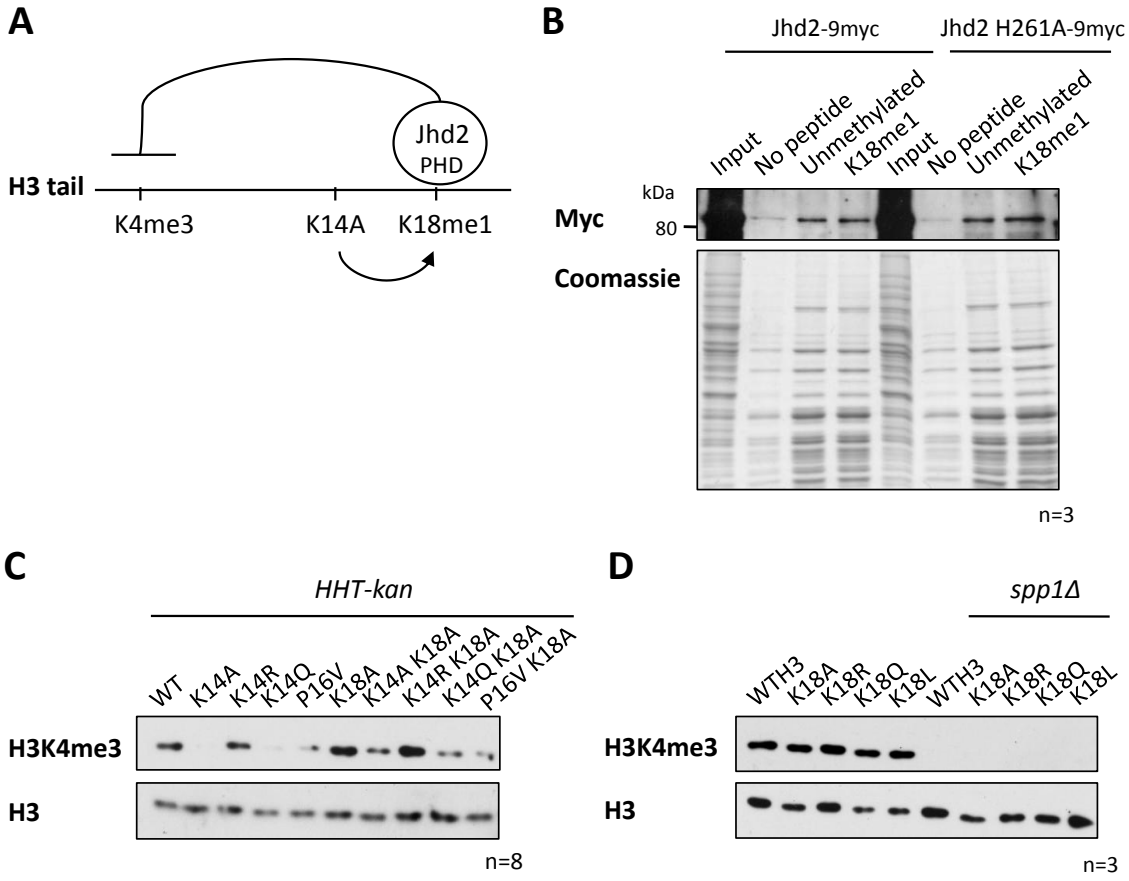


Figure 57. Investigating a possible mechanism for the crosstalk between H3K14 and H3K4me3 via H3K18me1.

(A) A schematic showing a model for the possible involvement of H3K18me1 and Jhd2 in the reduction of H3K4me3 by the K14A substitution. **(B)** A peptide pull-down experiment with the unmodified and K18-methylated H3 peptides (as Figure 42A) and extracts from *jhd2Δ* strains transformed with plasmids encoding wildtype *JHD2* and *JHD2* with a point mutation in the PHD finger (H261A) (kindly provided by M. Chandrasekharan). Interactions were detected by Western blot, probing for the Myc tag on Jhd2. 1/100th of the input sample was loaded. A Coomassie-stained gel is shown as a loading control. **(C)** A Western blot showing the effect of H3K18 substitution on the crosstalk between H3K14/H3P16 and H3K4me3. H3 levels are shown as a loading control. **(D)** As (C) except with the single and double mutant strains indicated.

6.2.13. Genome-wide analysis of the H3K14-H3K18me1 crosstalk

As with H3K4me3 (Chapter 3), substitution of H3K14 affected H3K18me1 to different extents on individual genes (Figure 58A). To investigate the effect of H3K14 on H3K18me1 further and to explore the relationship between H3K4me3 and H3K18me1 in the K14A strain, a ChIP-sequencing experiment was performed to assess H3K18me1 levels on every gene in the K14A strain. Genes were then ranked according to the ratio of average H3K18me1 across the transcription unit in the K14A relative to the WTH3 strain. The top- and bottom-ranked 200 genes were used for further

analysis (Figure 58B). Gene ontology terms were extracted for the genes in each class and are detailed in Table 23 and Table 24. The most common biological processes associated with the genes with the top 200 ratios of H3K18me1 (K14A/WTH3) are those involving mitochondrial function and translation whereas those genes with the bottom 200 ratios of H3K18me1 (K14A/WTH3) are involved in cell wall organisation. Since these are not opposite to the GO terms of the genes in the 'top' and 'bottom' H3K4me3 K14A classes, this analysis may hint that, in the K14A strain, the levels of H3K4me3 and H3K18me1 are not reciprocal on genes functioning in particular biological processes.

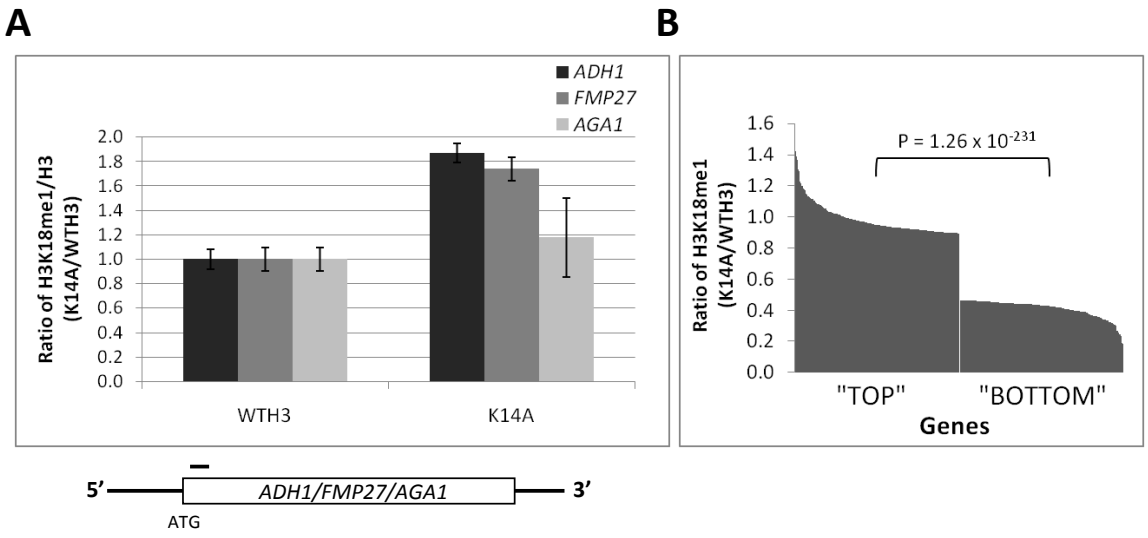


Figure 58. Mutation of H3K14 affects H3K18me1 to different extents on individual genes.

(A) A ChIP-qPCR experiment showing the levels of H3K18me1 at the 5' end of three genes (*ADH1, FMP27* and *AGA1*) in the K14A relative to the WTH3 strain. Primer positions are indicated on the locus map. Signals are normalised to levels of H3. Error bars show standard errors of the real-time PCR reaction. **(B)** The levels of H3K18me1 in the K14A relative to the WTH3 strain at the top and bottom 200 genes after ranking according to the ratio of the average H3K18me1 across the transcription unit (K14A/WTH3) as measured by ChIP-sequencing. Signals are not normalised to levels of H3 due to the high correlation between H3 in the WTH3 and K14A strains. P values are derived from student's t-test analysis.

Biological process	Count	%	P value
mitochondrion organisation	19	9.6	4.40E-05
electron transport chain	9	4.5	1.40E-04
generation of precursor metabolites and energy	17	8.6	3.00E-04
translation	30	15.2	4.80E-04
mitochondrial respiratory chain complex assembly	5	2.5	2.00E-03
oxidative phosphorylation	7	3.5	2.20E-03
cell redox homeostasis	5	2.5	3.50E-03
oxidation reduction	17	8.6	6.40E-03
cofactor metabolic process	12	6.1	6.60E-03
mitochondrial ATP synthesis coupled electron transport	4	2.0	2.60E-02
cell death	5	2.5	2.70E-02
mitochondrial respiratory chain complex IV assembly	3	1.5	3.00E-02
protein targeting to mitochondrion	5	2.5	3.30E-02
positive regulation of macromolecule biosynthetic process	10	5.1	3.40E-02
cofactor biosynthetic process	7	3.5	3.60E-02
respiratory electron transport chain	4	2.0	3.60E-02
positive regulation of cellular biosynthetic process	10	5.1	3.80E-02
response to oxidative stress	6	3.0	4.00E-02
respiratory chain complex IV assembly	3	1.5	4.00E-02

Table 23. Gene ontology terms associated with the genes with the top 200 ratios of H3K18me1 (K14A/WTH3).

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the top 200 ratios of H3K18me1 (K14A/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

Biological process	Count	%	P value
flocculation	5	2.6	1.90E-05
flocculation via cell wall protein-carbohydrate interaction	4	2.1	2.30E-04
cell adhesion	3	1.5	5.50E-03
cell wall organisation	10	5.1	1.00E-02
response to pheromone	6	3.1	1.80E-02

Table 24. Gene ontology terms associated with the genes with the bottom 200 ratios of H3K18me1 (K14A/WTH3).

Shown are the significantly enriched ($p<0.05$) biological processes and corresponding gene counts, percentages and modified Fisher exact P values for the genes with the bottom 200 ratios of H3K18me1 (K14A/WTH3). GO analysis was performed using the DAVID functional annotation tool v. 6.7 (http://david.abcc.ncifcrf.gov/home.jsp).

6.2.14. Genome-wide analysis of the H3K4me3-H3K18me1 crosstalk in the K14A strain

The correlation between the ratios of the average H3K4me3 and H3K18me1 across genes was determined genome-wide in the K14A relative to the WTH3 strain using the ChIP-seq datasets. However, in seeming conflict with the situation at *FMP27* (Figure 55A) and other genes (data not shown), there was essentially no correlation between the ratios of H3K4me3 and H3K18me1

genome-wide (R = -0.068). Potentially this disagreement stems from differences in the transcriptional status of or localisation of other histone modifications at *FMP27* compared to the rest of the genome. Thus it can be concluded that, although at certain genes H3K4me3 and H3K18me1 are present at reciprocal levels, the extent to which these two modifications co-vary genome-wide on individual genes in the K14A compared to the wildtype strain is not related.

6.3. Discussion

The work in this chapter aimed to characterise the recently-discovered methylation of H3K18 and its relationship with other histone H3 modifications.

6.3.1. H3K18me1 is a widespread modification

The genome-wide distribution of H3K18me1 indicates that this modification is widespread throughout the genome and its levels are very closely linked to those of histone H3. The genome-wide correlation between H3K18me1 and histone H3 is 0.94 but this correlation is highest at promoter regions, regions of high nucleosome turnover (Dion et al. 2007), compared to other genic regions. Therefore, methylation of H3K18 must occur before or shortly after incorporation of histone H3 into chromatin to maintain the high correlation between H3K18me1 and histone H3 in this region.

The relationship between H3K18me1 and histone H3 described here agrees with the data from Zee et al. (2010). In that study, the authors measured the half-maximal time of a range of modifications on histone H3. They found a similar $t_{1/2}$ for H3K18me1 and histone H3, suggesting that the turnover of these features are linked. Interestingly, the $t_{1/2}$ for H3K18me1 is more similar to the $t_{1/2}$ of 'repressive' rather than 'activating' modifications and thus it was proposed that H3K18me1 may have a repressive role in chromatin, possibly by opposing the deposition of H3K18ac.

6.3.2. Investigating the relationship between acetylation and methylation of H3K18

H3K18 can be both methylated and acetylated but these two modifications cannot co-exist on the same residue. However, there are two copies of the histone H3 N-terminal tail per nucleosome and so, even with the competition for the modification site, the distribution of acetylation or methylation of H3K18 is not automatically reciprocal genome-wide. Despite this, there is some

evidence to suggest there may be a functional acetyl/methyl switch at H3K18. The average distribution of H3K18me1 across genes after normalisation to histone H3 levels shows a slight dip at the promoter regions. This dip is greater in genes with higher transcriptional activity. In contrast, H3K18ac peaks in promoter regions and correlates positively with transcript levels (Liu et al. 2005), thus there may be a genome-wide anti-correlation between the localisations of H3K18ac and H3K18me1. Additionally, Gcn5, which predominantly acetylates H3K14 and H3K18 *in vivo* (Jiang et al. 2007) is significantly enriched at genes with the 200 lowest levels of H3K18me1/H3 across the transcription unit. However, if there is a simple acetyl/methyl switch at H3K18, it would be expected that strains in which H3K18ac global levels are altered would have oppositely-affected levels of H3K18me1. This was not the case, for example all three H3K14 substitution strains lacked detectable H3K18ac and yet H3K18me1 was differentially increased at several genes in these three strains.

Given the widespread nature of H3K18me1 throughout the genome, it is possible that this modification acts to dampen the levels of H3K18ac and/or restrict H3K18ac spatially or temporally, especially at dynamically-regulated genes. The changes in H3K18 modifications during the induction of *MET16* support this concept. When the gene was repressed, there was a high relative level of H3K18me1 and very low H3K18ac across the gene. However, upon *MET16* induction, H3K18ac levels increased at the 5' end and this was accompanied by a concomitant decrease in H3K18me1 specifically in the 5' region (3' H3K18me1 levels did not change greatly). Importantly, this decrease in H3K18me1 was not observed in the K14A strain, which lacks detectable H3K18ac. This suggests that the decrease in H3K18me1 in the 5' region of *MET16* during induction is a result of the increase in H3K18ac and not vice versa. The presence of H3K18me1 may act to prevent the deposition of H3K18ac until the signal for gene activation has been received and relevant factors have been recruited. The subsequent higher acetylation of

H3K18 may then restrict the sites available for methylation, resulting in the observed H3K18ac-dependent decrease in H3K18me1 upon gene activation.

6.3.3. Examining the relationship between H3K4me3 and H3K18me1

At the 5' end of *FMP27*, the levels of H3K4me3 and H3K18me1 appeared to be oppositely affected in the H3K14 substitution strains, raising the possibility that H3K18me1 may mediate the crosstalk between H3K14 and H3K4me3. However, this reciprocity could not be observed at a global level in the H3K14 substitution strains by Western blot or genome-wide by ChIP, even after focussing on the 5' coding regions. This discrepancy could be due to the specific transcriptional status of *FMP27* since it is constitutively transcribed at a relatively low level compared to the rest of the genome (Churchman and Weissman 2011). Whatever the reason(s) for the difference between the effect at *FMP27* and the rest of the genome, the lack of anti-correlation between H3K4me3 and H3K18me1 genome-wide in the K14A strain argues against a general involvement of H3K18me1 in the H3K14-H3K4me3 crosstalk.

6.3.4. Failure to identify a methyltransferase for H3K18

The candidate-based screen for the H3K18 methyltransferase failed to produce any notable hits. However, difficulties with this approach have been documented in the literature, for example the failure to identify the methyltransferase for H3R2 (Kirmizis et al. 2009). One possibility is that the catalytic domain of the H3K18 methyltransferase may not be similar to any of the known methyltransferases and so may not be included in any of the *in silico* searches: this was the case with Dot1 and its previously uncharacterised catalytic domain. To avoid this problem, the search for the methyltransferase should be widened to include all yeast proteins. Analysis of this type, termed Global Proteomic analysis of *S. cerevisiae* (GPS) has been developed by the Shilatifard laboratory (Schneider et al. 2004). A collaboration was set up with this laboratory and it was hoped that this analysis may have been included as part of this thesis, however, the H3K18me1

antibody was deemed to be of insufficient quality to perform a screen of this magnitude. Another factor to consider is that, although the sites of lysine methylation characterised to date in yeast only have a single methyltransferase acting on them, it is possible that, as in higher eukaryotes, more than one enzyme can act on the same substrate. This would make identification by a genetic method using single methyltransferase gene deletions impossible. Additionally, the methyltransferase may be essential, which would again preclude study by a genetic method unless a form of conditional deletion were used. Alternative methods could include biochemical purification of H3K18 methyltransferase activity, as was performed to identify the H3K36 methyltransferase Set2 (Strahl et al. 2002).

6.3.5. None of the six JmjC-containing proteins in yeast acts on H3K18

The search for the H3K18 demethylase was also unsuccessful. However, Rph1, the characterised H3K36 demethylase (Kim and Buratowski 2007; Tu, S et al. 2007) was found to indirectly influence H3K18me1 levels. A mechanism for this indirect interaction could involve H3K36 methylation since H3K18me1 levels were slightly lowered in the absence of this modification (*set2Δ*). An alternative explanation is that Rph1 is a repressor of cell growth, a function that does not rely on its catalytic activity (Klose et al. 2007). Since deletion and overexpression of *RPH1* promotes and inhibits cell growth respectively, this could potentially be responsible for the changes in H3K18me1 in these strains.

Many of the same caveats discussed above for the candidate-based genetic screen for the methyltransferase can also be applied to the demethylase screen. However, another possibility is that there is no demethylase for H3K18. Given the high correlation between H3K18me1 and histone H3, it is feasible that the main mechanism for H3K18me1 removal from chromatin is nucleosome turnover. Of note, another widely-distributed modification, H3K79 methylation, does

not have a characterised demethylase (Katan-Khaykovich and Struhl 2005; Liang et al. 2007; Tu, S et al. 2007) and is proposed to be a mark of old nucleosomes (De Vos et al. 2011).

6.3.6. Existence of H3K18me2/3

Until recently, all three methylation states have been identified on all the methylated lysines characterised. The exceptions are the Set5-catalysed lysines H4K5, H4K8 and H4K12, which have only been shown to be monomethylated (Green et al. 2012). H3K18me2 and H3K18me3 were not identified in any of the organisms tested in the original mass spectrometry study (Garcia et al. 2007). When antibodies raised against these modifications (kindly provided by R. Klose) were tested in wildtype yeast by both Western blot and ChIP, there was a very low signal and no specificity over the K18A negative control strain (data not shown). The failure to detect H3K18me2/3 over background may suggest that these modifications are not present in *S. cerevisiae*.

Chapter 7.

Discussion

7. Discussion

The work in this thesis has aimed to increase the knowledge about the crosstalk between histone modifications, defined as how one modification or residue can influence the deposition or removal of another modification on a nearby or distant residue. In particular, the crosstalk modulating trimethylation of H3K4 has been investigated because this modification attracts much attention due to its correlation with transcript levels and distribution at the 5' end of active genes. It was hoped that by increasing the understanding about how H3K4me3 is regulated, this may shed some light on the function of this modification in chromatin, which has yet to be satisfactorily defined.

The work in Chapter 3 characterised the crosstalk between lysine 14 on histone H3 and H3K4me3. Building on this, in Chapter 4 the function of the H3K14-adjacent proline 16 on histone H3 was investigated and a novel crosstalk between this residue and H3K4me3 was identified. Furthermore, the levels of H3K4me3 on individual genes encoding proteins with distinct cellular functions were differentially reduced by mutation of H3K14 or H3P16. The high correlation between the extent to which H3K4me3 was decreased on genes in the K14A and P16V strains relative to the wildtype strain suggests that these two residues may operate by an overlapping mechanism to influence H3K4me3. In Chapter 5, the mechanism for the H3K14/P16-H3K4me3 crosstalk was explored, focussing on the idea that the balance between H3K4me3 deposition and removal had been shifted in the mutant strains more in favour of demethylation over methylation. Lastly, in Chapter 6, a recently discovered modification, monomethylation of H3K18 was characterised and its involvement in the H3K14-H3K4me3 crosstalk was probed.

7.1. What is it about H3K14 that promotes H3K4me3?

As discussed in Chapter 3, during their initial characterisation of the H3K14-H3K4me3 crosstalk, Nakanishi et al. (2008) did not detect H3K4me3 in extracts from strains in which H3K14 had been mutated to alanine, arginine or glutamine. Based on this result and a previously described requirement of Gcn5 for proper H3K4me3 (Govind et al. 2007; Jiang et al. 2007), the authors proposed that acetylation of H3K14 is required for H3K4me3. For unknown reasons, possibly stemming from different growth conditions or thresholds of H3K4me3 detection by Western blot, the results with the three H3K14 substitution strains in the current study did not replicate the published findings. This discrepancy warrants the re-evaluation of the conclusions drawn by Nakanishi et al.

Wildtype levels of H3K4me3 are determined by the relative actions of the Set1 methyltransferase and the Jhd2 demethylase. The hypothesis put forward in Chapter 5 of this thesis is that this balance is tipped towards demethylation over methylation in the H3K14 mutant strains. Continuing with this logic, it could be said that, in the K14R strain, with the highest residual H3K4me3, this balance has been shifted less than for the K14Q and K14A strains, with the least residual H3K4me3 (Figure 59). It is not known what causes the shift in the balance between methylation and demethylation in the H3K14 mutant strains. This shift could be due to a decrease in methylation or an increase in demethylation or both. Although the investigation into the tri-methyltransferase function of Set1 in the H3K14 mutant strains has been inconclusive, it is evident that the H3K4me3 in these strains is more susceptible to demethylation by Jhd2.

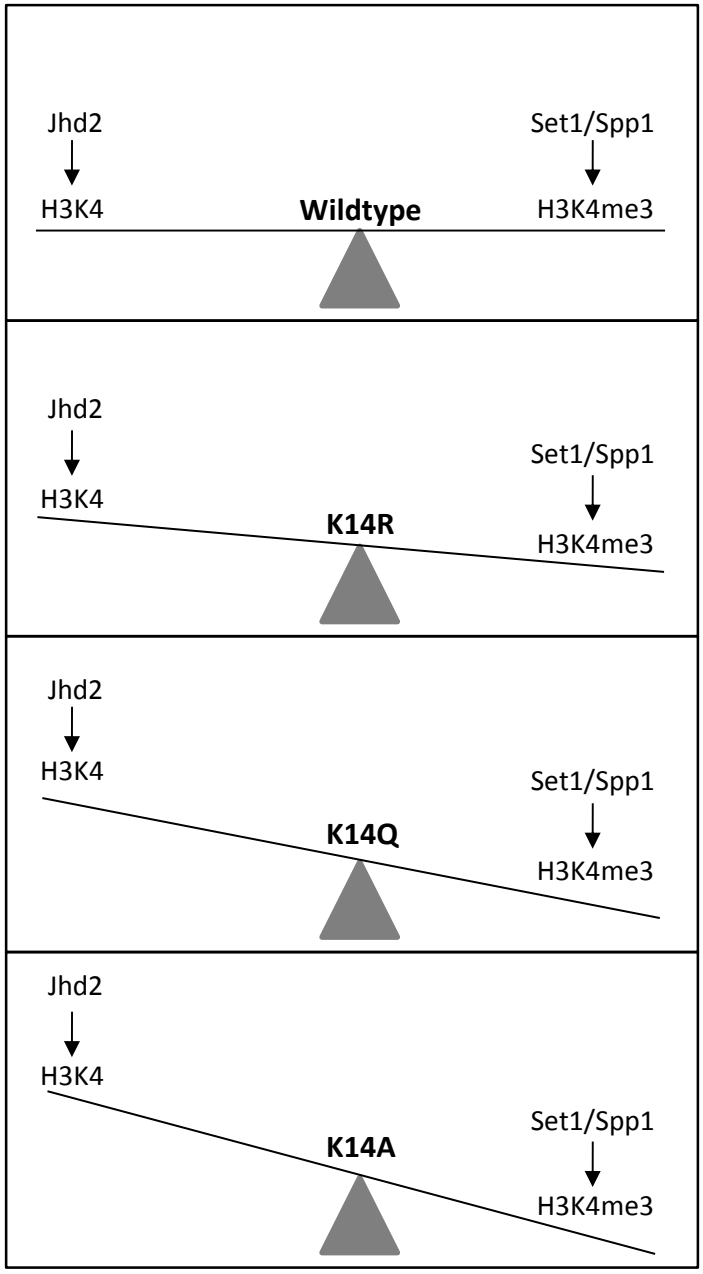


Figure 59. The balance between H3K4 methylation and demethylation is shifted in the H3K14 mutant strains.

A diagram illustrating the idea that H3K14 may be required to maintain the correct balance between H3K4 methylation by Set1/Spp1 and demethylation by Jhd2 and that in the H3K14 mutants, H3K4 demethylation is favoured over methylation. The dissimilar levels of H3K4me3 in the three substitution strains could be due to a greater shift in the balance in the K14A and K14Q strains than the K14R strain.

As discussed, the arginine substitution roughly mimics unmodified lysine and the glutamine substitution can recapitulate some of the features of acetylated lysine. This allows the results from the K14R and K14Q strains to be used to infer putative functions of the unmodified and acetylated H3K14 residues in the modulation of H3K4me3 levels. Because there are different global amounts of H3K4me3 in these two strains, it can be predicted that unmodified and acetylated H3K14 would also have different effects on the regulation of H3K4me3. Therefore the increased susceptibility of H3K4me3 in the K14Q strain to demethylation by Jhd2 shown in Chapter 5 would suggest that H3K14ac could also confer this sensitivity. (It is also possible that, instead of directly increasing the sensitivity to Jhd2, H3K14ac may have this effect by virtue of decreasing the extent of unmodified H3K14, which may be required to protect H3K4me3 from demethylation. It should be noted though that the K14R mutation, mimicking the unmodified H3K14 state also increases the sensitivity to demethylation.)

From the high correlation between the ratios of H3K4me3 in the K14A and P16V strains compared to the wildtype strain and the results from the H3K14/P16 double mutant strains, it can be predicted that H3K14 and H3P16 affect H3K4me3 by an overlapping mechanism. Another factor to consider therefore is how this overlap might be achieved. It is unknown whether the conformation of H3P16 can affect the modification state of H3K14; however, two lines of evidence suggest that H3K14 may affect the isomerisation state of the peptidyl-prolyl bond between H3A15 and H3P16. Firstly, the abrogation and then recovery of the binding of the Spt7 bromodomain to the K14R-K18ac and K14R-P16V-K18ac peptides respectively suggests that H3P16*trans* may be more permissive for Spt7 bromodomain binding and that mutation of H3K14 to arginine could prevent this interaction by increasing the proportion of H3P16*cis*. Secondly, an *in vitro* assay indicates that acetylation of H3K14 increases the proportion of H3P16 in *trans*. Whilst each of these results alone may not argue strongly for an effect of H3K14 on H3P16, these combined findings allow the proposal that H3K14 in its unmodified state may promote the *cis*

conformation of the H3A15-P16 bond whereas H3K14 in its acetylated state may increase the proportion of H3P16*trans*. It is assumed that the findings using H3 peptides *in vitro* can be translated into predictions for the behaviour of the histone H3 tail *in vivo*.

The above conclusion can form the basis of a model for this crosstalk mechanism (Figure 60). As discussed above, the H3K14 mutation confers sensitivity on H3K4me3 to Jhd2 action. Furthermore, a decrease in the amount of Spp1 associated with chromatin has been observed in the P16V strain, implying that H3P16 in *cis* may be more preferential for Spp1 binding. These effects of the two mutations can be used to explain the seemingly conflicting results in the H3K14-H3P16 double mutant strains. In the K14R strain, there is an increased susceptibility to Jhd2 action. However, through the predicted ability of the K14R mutation to increase the proportion of H3P16*cis*, the increased recruitment of Spp1 by this isomerisation state may be able to compete with the increased demethylation, so that overall there are near-wildtype levels of H3K4me3 in the K14R strain. When the K14R and P16V mutations are combined, the decreased amount of Spp1 recruitment (caused by the P16V mutation) can no longer compensate for the increased Jhd2 action (caused by the K14R mutation) and so the levels of H3K4me3 are lower in this double mutant strain compared to both the K14R and P16V strains.

In the K14Q strain, there is also an increased sensitivity to Jhd2. However, the glutamine substitution may mimic acetylated lysine in its ability to increase the proportion of H3P16*trans* and so this substitution could have the added effect of decreasing the recruitment of Spp1 to chromatin. These reasons can explain why the K14Q strain has lower levels of H3K4me3 than the K14R strain. Since the P16V substitution may mimic the H3P16*trans* state, the state that is expected to be predominant in the K14Q strain, the combination of the K14Q and P16V mutations has only a small additional negative effect on H3K4me3 beyond the reduction caused by the K14Q mutation alone. Because alanine is not comparable to any of the wildtype states of H3K14, it is

unclear what effect this substitution would have on H3P16 isomerisation. However, from the high positive correlation between the H3K4me3 ratios (mutant/wildtype) from the K14A and P16V ChIP-seq experiments, it can be predicted that the K14A mutation would also increase the proportion of H3P16*trans*. This may explain the similar levels of H3K4me3 in the K14A and K14Q strains and the epistasis observed for both of these substitutions in combination with the P16V mutation.

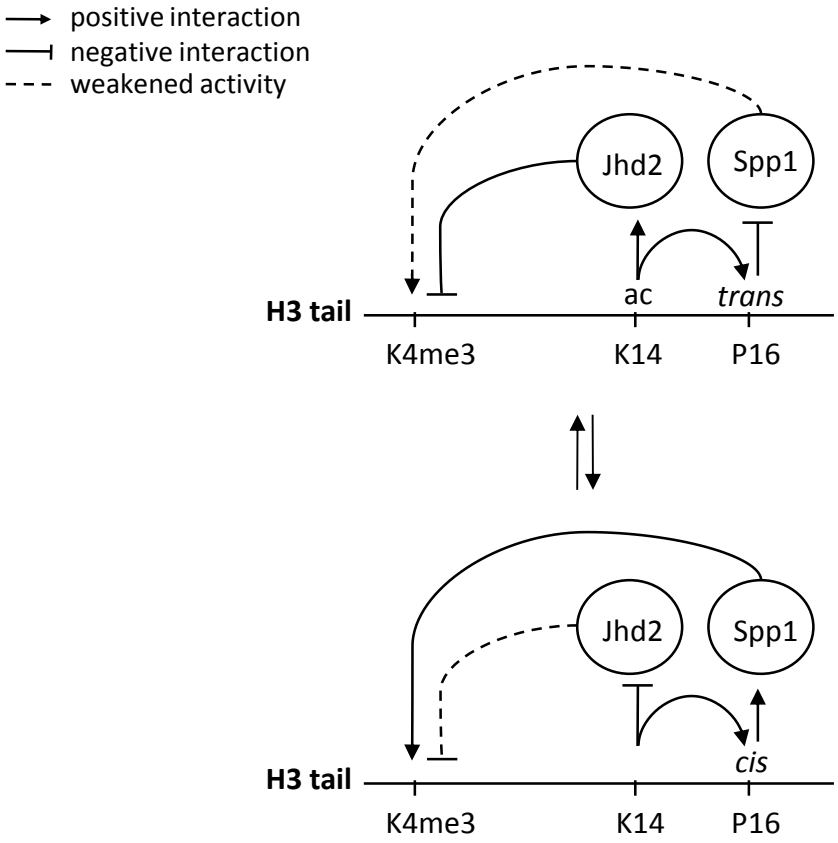


Figure 60. Working model for the regulation of H3K4me3 by H3K14 and H3P16.

H3K14 can both affect the activity of Jhd2 and the isomerisation of the H3A15-H3P16 peptidyl-prolyl bond, which can influence Spp1 chromatin association, to regulate H3K4me3.

It may therefore be that H3K14, when it is acetylated may actually inhibit rather than promote H3K4me3. Because the evidence in the literature suggests that H3K4me3 promotes the acetylation of H3K14ac, this sets up a feedback loop that allows dynamic regulation of the levels of both H3K4me3 and H3K14ac (Figure 61A). Dynamic acetylation associated with K4-trimethylated H3 has previously been described in mouse (Hazzalin and Mahadevan 2005). From

this model, it would be predicted that H3K4me3 and H3K14ac would not occur simultaneously on the same histone tails, which contradicts published findings. The genome-wide co-localisation of these modifications (Liu et al. 2005; Pokholok et al. 2005) can be explained by the population averaging that occurs in standard ChIP experiments. However, the co-localisation of these modifications on the same histone tails by mass spectrometry cannot be explained by this effect (Jiang et al. 2007). Another factor to consider is that H3K4me3 can also recruit the HDAC complex Rpd3L (Pinskaya et al. 2009; Wang et al. 2011). The decision to recruit a HAT or HDAC complex to H3K4me3 may allow a switch between the transcriptional 'on' and 'off' states of a gene (Figure 61A, B), an idea proposed by Buratowski and Kim (2011). The recruitment of an HDAC complex to H3K4me3 would presumably lead to a reduction in H3K14ac. From the model, it would be anticipated that this decrease in H3K14ac could result in an increase in H3K4me3, allowing more HDAC recruitment and thus further reinforcing the transition to the transcriptional 'off' state. This could explain the seemingly contrary finding that, at some genes, H3K4me3 levels increase during gene repression in response to stress (Weiner et al. 2012). Thus this dynamic form of regulation may be advantageous in the control of transcription because it could allow the rapid switching between transcriptional states. It would be interesting to determine how the choice is made between the recruitment of a HAT or HDAC complex to H3K4me3 to control this switch.

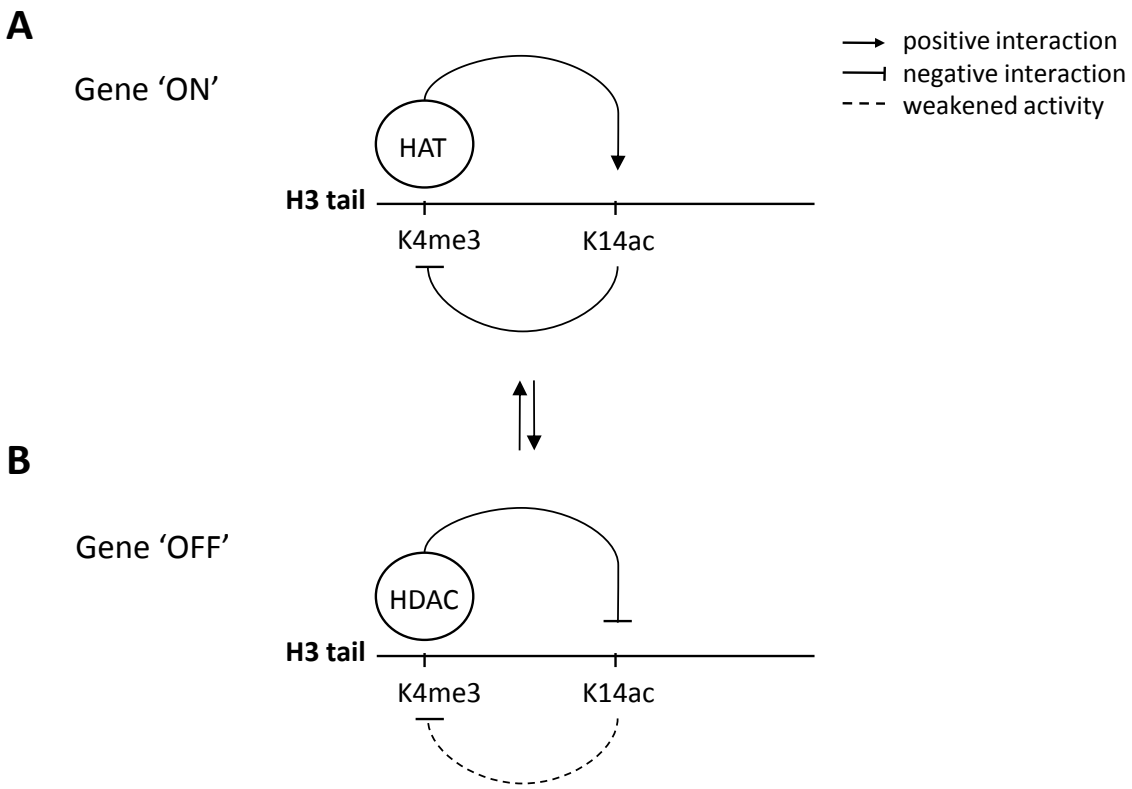


Figure 61. Dynamic H3K4me3 and H3K14ac may allow rapid changes in transcriptional state.

The ability of H3K4me3 to recruit **(A)** a histone acetyltransferase complex or **(B)** a histone deacetylase complex allows a switch between two chromatin states that are differentially permissive to gene transcription. The dynamic nature of these modifications may be advantageous in responding rapidly to changing cellular conditions.

The mutation of H3K14 or H3P16 did not affect H3K4me3 on all genes to the same extent. In the wildtype strain, the 'top' class genes with H3K4me3 that was decreased the least by the H3K14/H3P16 mutations are highly transcribed during optimal growth conditions whereas the 'bottom' class genes with H3K4me3 that was most affected are induced at times of environmental stress and slow growth. There was therefore a significant overlap between the genes in the 'top' and 'bottom' classes and those expressed in the oxidative and reductive phases of the yeast metabolic cycle respectively.

There are two proposed ideas that could explain the observed non-uniform effects of H3K14 mutation on H3K4me3. Firstly, the different ratios of H3K4me3 on genes in these two classes in the K14A relative to the wildtype strain could be due to K14A-induced lengthening of the

oxidative phase relative to the reductive phases, leading to the increased co-transcriptional deposition of H3K4me3 on the oxidative relative to the reductive phase genes. Potentially, the P16V substitution could also have this effect, although the overlap with the YMC-regulated genes for the 'top' and 'bottom' class genes in this strain was less significant. Secondly, it is clear from the analysis of the genes in the 'top' and 'bottom' classes that these K14A- and P16V-affected genes undergo distinct modes of transcriptional regulation in the wildtype strain. For example, transcription of the growth genes in the oxidative phase is dependent on histone acetylation whereas, although a significant proportion of genes in the 'bottom' class are regulated by SAGA, it is known that upon transcription of these SAGA-regulated genes during the reductive charging phase, there is no peak of histone acetylation (Cai et al. 2011). The promoter architectures of the genes in the two classes are also likely to be dissimilar: ribosomal protein genes in the 'top' class have a wide nucleosome-depleted region whereas reductive phase genes in the 'bottom' class have a more 'fuzzy' nucleosome occupancy around their transcription start sites (Machné and Murray 2012). Accompanying these differences in chromatin regulation are alterations in sense and antisense transcription. Genes in the 'top' class have high levels of sense and similar amounts of antisense transcription compared to the average of all genes in the genome. In contrast, genes in the 'bottom' class are depleted for sense but enriched in antisense transcription. Additionally, based on the altered positioning of the Nrd1 5' peak compared to the genomic average, genes in the 'top' class are predicted to have short sense transcripts at their 5' ends, which may imply that these genes are regulated by transcriptional attenuation. All these transcripts have both the potential to be controlled by H3K4me3 via the enhancement of Nrd1-dependent transcription termination and to regulate H3K4me3 by co-transcriptional deposition. Whilst it is not known how the H3K14 mutation would affect all these modes of regulation, the underlying variations between distinct classes of genes in the wildtype strain could result in non-uniform sensitivity of H3K4me3 on individual genes to the loss of H3K14.

Therefore, at least three factors may affect H3K4me3 in the H3K14 mutant strains. Firstly, H3K14 is required to protect H3K4me3 from demethylation by Jhd2 by an as yet unknown mechanism. Secondly, the modification state of H3K14 may influence the isomerisation state of the peptidyl-prolyl bond between H3A15 and H3P16. Since the *cis* or *trans* conformation may influence Spp1 association with chromatin, this could allow H3K14 to additionally indirectly influence H3K4me3 deposition. Thirdly, the H3K14 mutation may cause an increase in the duration of the oxidative phase of the yeast metabolic cycle so that the genes expressed during this phase experience more co-transcriptional H3K4me3 deposition. This could account for the differential H3K4me3 losses on the genes expressed in the oxidative and reductive phases of the yeast metabolic cycle. Alternatively, but not necessarily mutually exclusively, the distinct modes of transcriptional regulation of the genes in the 'top' and 'bottom' classes may also contribute to these differential gene-specific effects.

7.2. H3K14 also participates in crosstalk with H3K18

In all three H3K14 substitution strains, there were undetectable levels of H3K18ac. This reduction in H3K18ac could result from the less efficient binding of Gcn5 via its bromodomain (or another H3K18 HAT) in the absence of H3K14ac. Alternatively, the proximity between H3K14 and H3K18 allows the proposal of a model whereby the substrate selection and/or binding of the acetyltransferase active site to the histone H3 tail is inhibited by mutation of H3K14. A third intriguing possibility is that the conformational state of H3P16 may be altered by mutation of H3K14. H3P16 participates in important interactions with the active site of Gcn5 when acetylating H3K14 and it is plausible that similar interactions may be required for the recognition of H3K18 as a substrate. The fact that there was undetectable H3K18ac in all three H3K14 substitution strains but that the K14R and K14Q substitutions are envisaged to oppositely affect the conformation of H3P16 may argue against this hypothesis. It would be interesting to determine whether mutation of H3P16 to valine in the H3K14 mutant could rescue H3K18ac.

H3K18ac is not the only modification on H3K18 that is affected by mutation of H3K14. Monomethylation of H3K18, a modification whose levels and distribution correlate very strongly with histone H3, is increased by mutation of H3K14 to alanine and glutamine but not in the K14R strain at the 5' coding regions of several genes. The reciprocal levels of H3K4me3 and H3K18me1 in the H3K14 mutant strains at certain genes by ChIP were interesting and it is still unclear whether there is a functional crosstalk between H3K18me1 and H3K4me3 genome-wide. Also unknown is the exact relationship between H3K18ac and H3K18me1. Because H3K18me1 was not uniformly raised in the three H3K14 mutant strains with undetectable H3K18ac it seems that there is not a simple competition-based acetyl/methyl switch at this residue. However, genomic regions that are enriched for H3K18ac, such as promoter regions, appear to be more depleted for H3K18me1 than other regions of the genome. The widespread distribution of H3K18me1 may function to restrict H3K18ac deposition until a signal for gene activation has been received. Similar mechanisms may operate at other histone lysine residues that can both be methylated and acetylated, for example H3K4 and H3K36.

7.3. Future work

The hypotheses proposed in this work could be strengthened by further experiments. The published evidence suggesting a requirement of H3K4me3 for H3K14ac is reasonably compelling but yet no effect of the loss of integrity or methylation of H3K4 on H3K14ac was detected in this work. The first step to resolve these different results would be to perform Western blots with dilution series of protein extracts to ensure that the antibody/protein ratio is optimal for detection of potentially small changes in H3K14ac. Furthermore, this crosstalk may be gene-specific and therefore may be masked at a global level. A ChIP-qPCR experiment could be performed, assessing the levels of H3K14ac at a range of genes in the *spp1Δ* strain. If this preliminary experiment showed differences in the H3K4me3-H3K14ac crosstalk at different genes, then it could be extended to a ChIP-seq experiment.

It is unknown how mutation of H3K14 sensitises H3K4me3 to Jhd2. The ChIP experiment assessing Jhd2 chromatin association in the wildtype and H3K14 mutant strains would need to be repeated before any conclusions could be drawn about increased recruitment of Jhd2 to chromatin in the absence of H3K14. However, increased Jhd2 recruitment in the H3K14 mutant strains is not critical for the model proposed above as there are other mechanisms by which H3K14 could influence Jhd2 activity. The effect of the P16V mutation on the sensitivity of H3K4me3 to Jhd2 has yet to be determined; the same experiments performed with the H3K14 mutant strains will be repeated with the P16V strain. More evidence is also needed to support the hypothesis for the differential effects of H3K14/P16 mutation on H3K4me3 at individual genes. The proposed K14A-induced extension of the YMC oxidative phase could be assessed by measuring the proportion of $DiOC_6$-positive cells undergoing oxidative metabolism in the K14A relative to the wildtype strain. Additionally, from the proposed decrease in the duration of the reductive phases in the K14A strain, it can be predicted that this population would be more susceptible to heat shock (Slavov et al. 2012). An assay measuring the viability of wildtype and H3K14-mutant yeast cultures after increasing severity of heat shock could be used to test this hypothesis. As well as potentially adding weight to the proposed increase in the oxidative phase duration by measuring the levels of oxidative phase transcripts in the K14A strain, a strand-specific microarray could also be used to investigate whether the dissimilar modes of transcriptional regulation are responsible for the gene-specific effects of the K14A mutation on H3K4me3. It is possible that the K14A mutation may alter the ratio of sense/antisense transcription, which could influence H3K4me3.

So far, there is only indirect evidence to indicate that the modification state of H3K14 may influence the conformation of H3P16. It may be possible to use NMR to assess the effects of H3K14 acetylation on the structure of a long histone H3 peptide. Although proline residues cannot be detected by 2D ^{1}H-^{15}N HSQC experiments due to the lack of a backbone amide group, interactions between proximal residues in space can be observed and the difference between

these interactions in the presence and absence of acetylation on H3K14 can be used to infer structural changes. Conversely, the potential of the P16V mutation to influence the level of H3K14ac could be further probed by mass spectrometry, although as discussed in Chapter 4, there could be some difficulties in the normalisation of modification levels on tryptic peptides that harbour a mutation.

The mechanism for the H3K14-H3K18ac crosstalk is not known. Peptide pull-down experiments could be used to determine whether loss of acetylation of H3K14 affects the binding of Gcn5 to an H3 peptide. Furthermore, *in vitro* acetyltransferase assays could be performed with a range of differently modified/mutated histone H3 peptides to assess the ability of Gcn5 (or other HATs) to acetylate H3K18. To determine whether H3P16 is involved in this crosstalk, mass spectrometry could be used to measure the relative amounts of H3K18ac in the wildtype, P16V and K14A P16V strains.

Further progress on the characterisation of H3K18me1 would be difficult without identifying the methyltransferase that deposits this modification. The search could not be expanded to test all the non-essential yeast genes unless a new, more sensitive antibody was raised. It could be possible to analyse the conditional yeast deletion library but this may also require a higher quality antibody. A biochemical rather than genetic approach may therefore be required to purify the methyltransferase activity from yeast cell extracts.

7.4. Concluding remarks

In summary, the work in this project has characterised several known and novel relationships between histone modifications and, in doing so, has increased the understanding about the regulation of these modifications. Since the H3K14/P16-H3K4me3 crosstalk is the first to be analysed genome-wide, it is unknown whether other examples of histone crosstalk would also

exhibit gene-specific effects. If they do, this would add another layer of complexity onto the already complicated regulation of histone modifications.

Chapter 8.

Appendix

8. Appendix

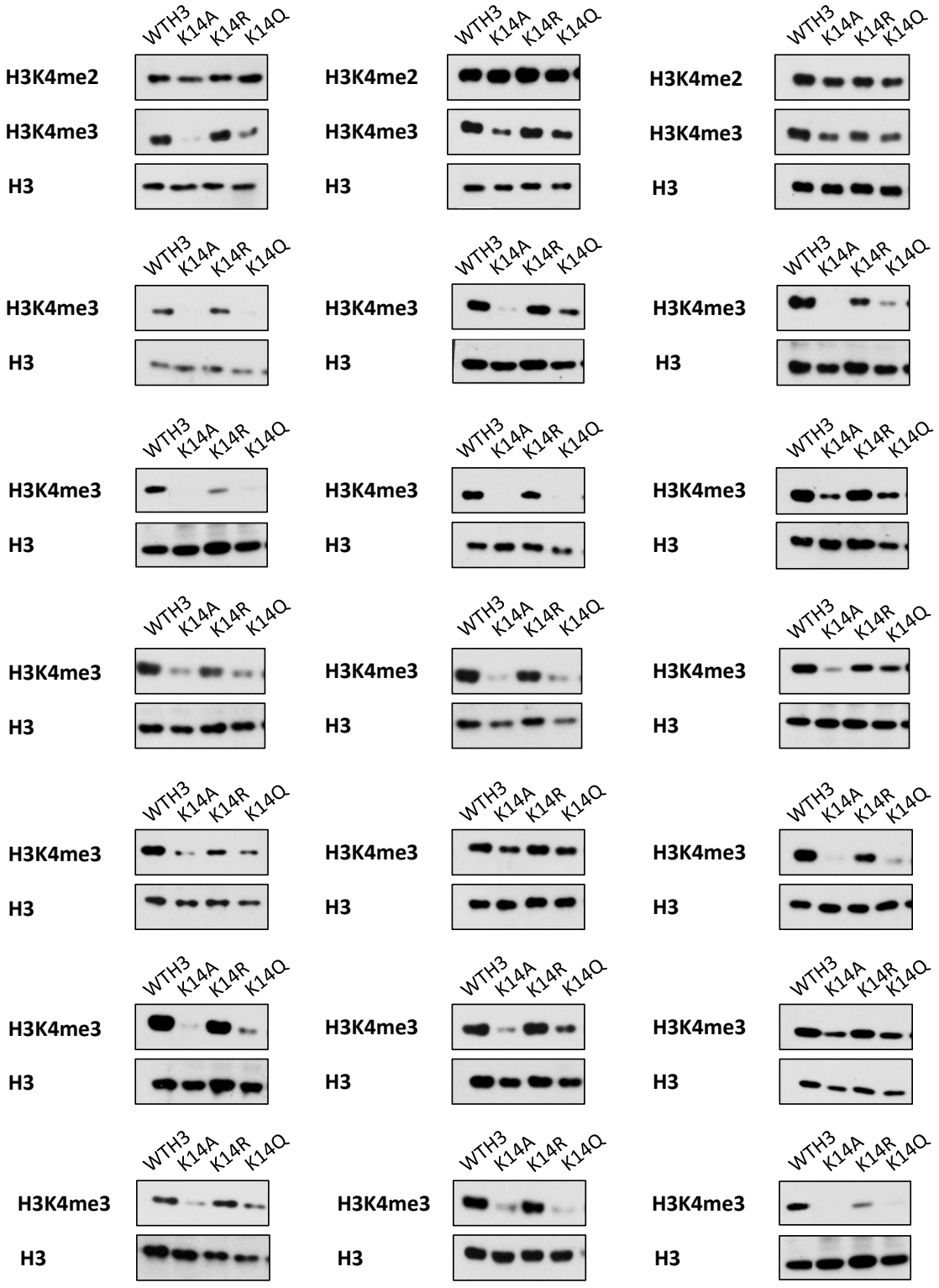


Figure 62. Three substitutions of H3K14 differentially affect levels of H3K4me3 but not H3K4me2.

As Figure 10B, showing the replicates of the H3K4me2 and H3K4me3 Western blot results. Histone H3 levels are shown as a loading control.

Genes with top 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YKL225W		1
YOL166W-A		2
YPL278C		3
YAR060C		4
YNR075C-A		5
YGL260W		6
YCR025C		7
YFR057W		8
YOL166C		9
YDR008C		10
YDR007W	*TRP1*	11
YOL161C	*PAU20*	12
YHR212W-A		13
YOR192C-C		14
YAL068W-A		15
YLR279W		16
YJR162C		17
YBR296C-A		18
YJR005C-A		19
YLL065W		20
YCL076W		21
YBR099C		22
YDR544C		23
YBL108W		24
YLR112W		25
YIL054W		26
YNL269W	*BSC4*	27
YDR024W	*FYV1*	28
YCR102W-A		29
YOR108C-A		30
YAL069W		31
YMR325W	*PAU19*	32
YJL075C	*APQ13*	33
YNL067W-B		34
YFR009W-A		35
YDR194W-A		36
YMR107W	*SPG4*	37
YOL083C-A		38
YKL006W	*RPL14A*	39
YOL110W	*SHR5*	40
YCL065W		41
YGL188C-A		42
YDR500C	*RPL37B*	43
YNL337W		44
YLR013W	*GAT3*	45
YOR312C	*RPL20B*	46
YKL117W	*SBA1*	47
YKL202W		48
YGR118W	*RPS23A*	49
YPR136C		50
YPL205C		51
YDL159W-A		52
YLR444C		53
YFL051C		54
YNL235C		55
YIL148W	*RPL40A*	56
YKL009W	*MRT4*	57
YGL007C-A		58
YPR108W-A		59
YDL075W	*RPL31A*	60
YOR096W	*RPS7A*	61
YDL083C	*RPS16B*	62
YAR035C-A		63
YLR075W	*RPL10*	64

Genes with top 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YMR193C-A		65
YML063W	*RPS1B*	66
YMR096W	*SNZ1*	67
YML026C	*RPS18B*	68
YLR076C		69
YKL156W	*RPS27A*	70
YBL071C-B		71
YHR010W	*RPL27A*	72
YLR406C-A		73
YMR230W	*RPS10B*	74
YMR141W-A		75
YLR149C-A		76
YDL061C	*RPS29B*	77
YER117W	*RPL23B*	78
YGL054C	*ERV14*	79
YOR366W		80
YDL114W		81
YPL079W	*RPL21B*	82
YLR322W	*VPS65*	83
YOR170W		84
YLL044W		85
YKL156C-A		86
YNL067W	*RPL9B*	87
YDL136W	*RPL35B*	88
YJL189W	*RPL39*	89
YAL067W-A		90
YFL038C	*YPT1*	91
YPR002C-A		92
YBR191W	*RPL21A*	93
YEL053W-A		94
YGL135W	*RPL1B*	95
YDR468C	*TLG1*	96
YPR077C		97
YOR182C	*RPS30B*	98
YJR123W	*RPS5*	99
YDR450W	*RPS18A*	100
YOR293W	*RPS10A*	101
YLR441C	*RPS1A*	102
YLL045C	*RPL8B*	103
YNL162W	*RPL42A*	104
YHR021C	*RPS27B*	105
YLR448W	*RPL6B*	106
YOR210W	*RPB10*	107
YDR447C	*RPS17B*	108
YIL029W-A		109
YGL031C	*RPL24A*	110
YBR084C-A	*RPL19A*	111
YHL015W	*RPS20*	112
YDR427W	*RPN9*	113
YJL148W	*RPA34*	114
YKR057W	*RPS21A*	115
YHR126C	*ANS1*	116
YMR242W-A		117
YHR180C-B		118
YMR143W	*RPS16A*	119
YDR199W		120
YDL020C	*RPN4*	121
YEL054C	*RPL12A*	122
YGR085C	*RPL11B*	123
YBR181C	*RPS6B*	124
YHR050W-A		125
YKL068W-A		126
YLR461W	*PAU4*	127
YOR050C		128

Genes with top 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YNL110C	*NOP15*	129
YOL121C	*RPS19A*	130
YIL133C	*RPL16A*	131
YOL127W	*RPL25*	132
YLR029C	*RPL15A*	133
YNR046W	*TRM112*	134
YLR264C-A		135
YOR394C-A		136
YPL143W	*RPL33A*	137
YDR417C		138
YGL123C-A		139
YEL009C	*GCN4*	140
YNL114C		141
YOR194C	*TOA1*	142
YNL096C	*RPS7B*	143
YJL188C	*BUD19*	144
YGR148C	*RPL24B*	145
YLR264W	*RPS28B*	146
YOR261C	*RPN8*	147
YGR160W		148
YGL011C	*SCL1*	149
YOR281C	*PLP2*	150
YLR163W-A		151
YPR102C	*RPL11A*	152
YBL072C	*RPS8A*	153
YMR254C		154
YOL040C	*RPS15*	155
YLR395C	*COX8*	156
YOR282W		157
YLR033W	*RSC58*	158
YMR242C	*RPL20A*	159
YJL136C	*RPS21B*	160
YLR333C	*RPS25B*	161
YDR071C	*PAA1*	162
YGL189C	*RPS26A*	163
YER034W		164
YCL031C	*RRP7*	165

Genes with top 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YLR388W	*RPS29A*	166
YDR418W	*RPL12B*	167
YHR141C	*RPL42B*	168
YMR194W	*RPL36A*	169
YLR287C-A	*RPS30A*	170
YJR086W	*STE18*	171
YBR190W		172
YBL092W	*RPL32*	173
YLR367W	*RPS22B*	174
YPL044C		175
YJL177W	*RPL17B*	176
YOL109W	*ZEO1*	177
YGR214W	*RPS0A*	178
YDL147W	*RPN5*	179
YML058W-A	*HUG1*	180
YHR203C	*RPS4B*	181
YER102W	*RPS8B*	182
YDR431W		183
YNL216W	*RAP1*	184
YPL282C	*PAU22*	185
YMR142C	*RPL13B*	186
YIL144W	*TID3*	187
YMR307C-A		188
YNL069C	*RPL16B*	189
YER047C	*SAP1*	190
YBR048W	*RPS11B*	191
YGR128C	*UTP8*	192
YGL222C	*EDC1*	193
YCL064C	*CHA1*	194
YDL192W	*ARF1*	195
YAL016C-A		196
YHL001W	*RPL14B*	197
YLR308W	*CDA2*	198
YDR220C		199
YBR179C	*FZO1*	200

Table 25. Genes with the top 200 ratios of H3K4me3 (K14A/WTH3).

Shown are the systematic and common gene names and the corresponding ranks after ordering all the genes in the genome according to (from largest to smallest) the ratio of the average H3K4me3 across the transcription unit in the K14A strain relative to the WTH3 strain.

Genes with bottom 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YHR015W	*MIP6*	6373
YLR307W	*CDA1*	6374
YMR151W	*YIM2*	6375
YCR083W	*TRX3*	6376
YHR210C		6377
YDR317W	*HIM1*	6378
YFL030W	*AGX1*	6379
YEL075W-A		6380
YDL024C	*DIA3*	6381
YEL004W	*YEA4*	6382
YLR142W	*PUT1*	6383
YDR540C	*IRC4*	6384
YOL033W	*MSE1*	6385
YFR007W	*YFH7*	6386
YLL067W-A		6387
YBR280C	*SAF1*	6388
YLR286C	*CTS1*	6389
YOR186C-A		6390
YFR017C		6391
YDL247W	*MPH2*	6392
YFL059W	*SNZ3*	6393
YDR453C	*TSA2*	6394
YPR006C	*ICL2*	6395
YKL036C		6396
YGR236C	*SPG1*	6397
YFL040W		6398
YOR237W	*HES1*	6399
YAR023C		6400
YHR124W	*NDT80*	6401
YBR183W	*YPC1*	6402
YDR085C	*AFR1*	6403
YOL055C	*THI20*	6404
YJL045W		6405
YGL229C	*SAP4*	6406
YFR026C	*ULI1*	6407
YJR153W	*PGU1*	6408
YCL021W-A		6409
YPR127W		6410
YER097W		6411
YHL012W		6412
YDL244W	*THI13*	6413
YKL037W	*AIM26*	6414
YGR290W		6415
YBR013C		6416
YOL159C		6417
YLR070C	*XYL2*	6418
YMR175W-A		6419
YOR382W	*FIT2*	6420
YKL177W		6421
YJR128W		6422
YAR033W	*MST28*	6423
YNL289W	*PCL1*	6424
YJL132W		6425
YKL018C-A		6426
YLL047W		6427
YIL164C	*NIT1*	6428
YKL219W	*COS9*	6429
YCR001W		6430
YMR025W	*CSI1*	6431
YLR453C	*RIF2*	6432
YGR174C	*CBP4*	6433
YGR045C		6434
YDL025W-A		6435
YMR155W		6436

Genes with bottom 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YPL060W		6437
YJR160C	*MPH3*	6438
YAL008W	*FUN14*	6439
YOR381W-A		6440
YER150W	*SPI1*	6441
YPL187W	*MFα1*	6442
YBR072W	*HSP26*	6443
YOL048C	*RRT8*	6444
YBR056C-B		6445
YPR036W-A		6446
YJR010W	*MET3*	6447
YJR085C		6448
YCL001W-B		6449
YBR116C		6450
YHR001W-A	*QCR10*	6451
YAR029W		6452
YPL061W	*ALD6*	6453
YMR175W	*SIP18*	6454
YJL213W		6455
YHR177W		6456
YPR193C	*HPA2*	6457
YKL178C	*STE3*	6458
YJL116C	*NCA3*	6459
YGL184C	*STR3*	6460
YJR094C	*IME1*	6461
YGR029W	*ERV1*	6462
YLR092W	*SUL2*	6463
YML118W	*NGL3*	6464
YDR059C	*UBC5*	6465
YHR209W	*CRG1*	6466
YPL186C	*UIP4*	6467
YNL012W	*SPO1*	6468
YGL052W		6469
YJL037W	*IRC18*	6470
YPL185W		6471
YBR054W	*YRO2*	6472
YBL098W	*BNA4*	6473
YBR056W-A		6474
YIR028W	*DAL4*	6475
YOR012W		6476
YJL169W		6477
YDR106W	*ARP10*	6478
YIR032C	*DAL3*	6479
YDR126W	*SWF1*	6480
YGL258W	*VEL1*	6481
YPR201W	*ARR3*	6482
YJR151W-A		6483
YNL019C		6484
YOL114C		6485
YOL154W	*ZPS1*	6486
YCR098C	*GIT1*	6487
YOR013W	*IRC11*	6488
YEL076C-A		6489
YDL242W		6490
YJR115W		6491
YLR307C-A		6492
YCL048W-A		6493
YLR121C	*YPS3*	6494
YJL214W	*HXT8*	6495
YMR169C	*ALD3*	6496
YLR040C		6497
YEL076C		6498
YKL065W-A		6499
YOR190W	*SPR1*	6500

Genes with bottom 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YEL009C-A		6501
YHR054C		6502
YNL128W	*TEP1*	6503
YJL057C	*IKS1*	6504
YDL049C	*KNH1*	6505
YIR030C	*DCG1*	6506
YOL159C-A		6507
YBR250W	*SPO23*	6508
YOR071C	*NRT1*	6509
YBR117C	*TKL2*	6510
YLR081W	*GAL2*	6511
YFR012W		6512
YBL029W		6513
YEL010W		6514
YPL171C	*OYE3*	6515
YCR101C		6516
YOR339C	*UBC11*	6517
YLR122C		6518
YAL062W	*GDH3*	6519
YFL060C	*SNO3*	6520
YGL088W		6521
YCL049C		6522
YNL009W	*IDP3*	6523
YGR254W	*ENO1*	6524
YAL015C	*NTG1*	6525
YDR516C	*EMI2*	6526
YNL013C		6527
YPL088W		6528
YNL031C	*HHT2*	6529
YLL049W	*LDB18*	6530
YIL120W	*QDR1*	6531
YLR041W		6532
YLR232W		6533
YHR041C	*SRB2*	6534
YGR122C-A		6535
YLR123C		6536
YNL333W	*SNZ2*	6537

Genes with bottom 200 H3K4me3 ratios (K14A v WTH3)	Alias	Rank
YBL100W-C		6538
YNL033W		6539
YOR003W	*YSP3*	6540
YOR391C	*HSP33*	6541
YHR044C	*DOG1*	6542
YMR174C	*PAI3*	6543
YGL194C-A		6544
YBR009C	*HHF1*	6545
YKR106W		6546
YHR052W-A		6547
YNL042W-B		6548
YML087C	*AIM33*	6549
YHR053C	*CUP1-1*	6550
YDR461C-A		6551
YBR240C	*THI2*	6552
YNL195C		6553
YDL022C-A		6554
YMR324C		6555
YMR170C	*ALD2*	6556
YGR240C-A		6557
YCL026C-A	*FRM2*	6558
YGL258W-A		6559
YJL205C	*NCE101*	6560
YGR121W-A		6561
YHR054W-A		6562
YLR042C		6563
YHR055C	*CUP1-2*	6564
YLL057C	*JLP1*	6565
YPL054W	*LEE1*	6566
YBR196C-B		6567
YLR120W-A		6568
YML054C-A		6569
YJL222W-B		6570
YHR180W		6571
YLR111W		6572

Table 26. Genes with the bottom 200 ratios of H3K4me3 (K14A/WTH3).

Shown are the systematic and common gene names and the corresponding ranks after ordering all the genes in the genome according to (from largest to smallest) the ratio of the average H3K4me3 across the transcription unit in the K14A strain relative to the WTH3 strain.

Genes with top 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YKL225W		1
YOL166W-A		2
YLR111W		3
YPL278C		4
YDR544C		5
YNR075C-A		6
YMR325W	*PAU19*	7
YCL076W		8
YLL065W		9
YFR057W		10
YBR296C-A		11
YNL337W		12
YOR108C-A		13
YJR005C-A		14
YOL161C	*PAU20*	15
YJL075C	*APQ13*	16
YLR279W		17
YCR025C		18
YJR162C		19
YLR112W		20
YBR099C		21
YDR008C		22
YDR194W-A		23
YAL068W-A		24
YBL108W		25
YDL159W-A		26
YAR060C		27
YMR272W-B		28
YAL066W		29
YOR394C-A		30
YPR159C-A		31
YDR468C	*TLG1*	32
YOL110W	*SHR5*	33
YNR046W	*TRM112*	34
YMR107W	*SPG4*	35
YFR009W-A		36
YHR212W-A		37
YGL260W		38
YNL067W	*RPL9B*	39
YLR444C		40
YMR158C-A		41
YDR500C	*RPL37B*	42
YAL034C-B		43
YCR102W-A		44
YNL170W		45
YLR322W	*VPS65*	46
YKL117W	*SBA1*	47
YGL054C	*ERV14*	48
YPL079W	*RPL21B*	49
YGL135W	*RPL1B*	50
YMR230W	*RPS10B*	51
YFL038C	*YPT1*	52
YPR108W-A		53
YML058W-A	*HUG1*	54
YER117W	*RPL23B*	55
YAR047C		56
YJL027C		57
YDR024W	*FYV1*	58
YBR174C		59
YOL127W	*RPL25*	60
YOR192C-C		61
YMR141W-A		62
YBL071C-B		63

Genes with top 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YNL216W	*RAP1*	64
YLR163W-A		65
YOR312C	*RPL20B*	66
YDR450W	*RPS18A*	67
YHR068W	*DYS1*	68
YKL123W		69
YKL009W	*MRT4*	70
YOL062C	*APM4*	71
YHL015W-A		72
YHR063W-A		73
YDL212W	*SHR3*	74
YER092W	*IES5*	75
YIL029W-A		76
YIL148W	*RPL40A*	77
YNL100W	*AIM37*	78
YPL220W	*RPL1A*	79
YDR118W-A		80
YDL020C	*RPN4*	81
YKL006W	*RPL14A*	82
YML063W	*RPS1B*	83
YFL051C		84
YOR050C		85
YBR191W	*RPL21A*	86
YJR123W	*RPS5*	87
YIL054W		88
YJL190C	*RPS22A*	89
YPR136C		90
YGL024W		91
YML026C	*RPS18B*	92
YOR272W	*YTM1*	93
YKL196C	*YKT6*	94
YLR230W		95
YMR182W-A		96
YKR025W	*RPC37*	97
YDR394W	*RPT3*	98
YOR282W		99
YLR395C	*COX8*	100
YOR210W	*RPB10*	101
YPR166C	*MRP2*	102
YDR220C		103
YOR281C	*PLP2*	104
YOR293W	*RPS10A*	105
YDR344C		106
YOR194C	*TOA1*	107
YIL134C-A		108
YLR388W	*RPS29A*	109
YDR177W	*UBC1*	110
YGL188C-A		111
YJL189W	*RPL39*	112
YCL065W		113
YHR010W	*RPL27A*	114
YDL092W	*SRP14*	115
YPR074W-A		116
YDL083C	*RPS16B*	117
YGL011C	*SCL1*	118
YAL069W		119
YAL042C-A		120
YFL067W		121
YOR331C		122
YLR013W	*GAT3*	123
YBR190W		124
YBR109W-A		125
YER072W	*VTC1*	126

Genes with top 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YNL067W-B		127
YKL156C-A		128
YPR101W	*SNT309*	129
YMR071C	*TVP18*	130
YGR265W		131
YHR203C	*RPS4B*	132
YOR096W	*RPS7A*	133
YJL096W	*MRPL49*	134
YDL162C		135
YGR128C	*UTP8*	136
YGR275W	*RTT102*	137
YBR106W	*PHO88*	138
YGR160W		139
YMR242W-A		140
YOL121C	*RPS19A*	141
YDR300C	*PRO1*	142
YDL136W	*RPL35B*	143
YLR075W	*RPL10*	144
YDR184C	*ATC1*	145
YKL213C	*DOA1*	146
YOR122C	*PFY1*	147
YEL053W-A		148
YIL144W	*TID3*	149
YMR254C		150
YKL156W	*RPS27A*	151
YGR118W	*RPS23A*	152
YNL162W	*RPL42A*	153
YLR076C		154
YMR168C	*CEP3*	155
YER112W	*LSM4*	156
YLR029C	*RPL15A*	157
YLL056C		158
YHR141C	*RPL42B*	159
YJR063W	*RPA12*	160
YDR469W	*SDC1*	161
YIL021C-A		162
YLR170C	*APS1*	163
YLR019W	*PSR2*	164

Genes with top 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YHR050W-A		165
YFL012W-A		166
YGR214W	*RPS0A*	167
YLR441C	*RPS1A*	168
YDL055C	*PSA1*	169
YML067C	*ERV41*	170
YLL044W		171
YGL106W	*MLC1*	172
YBL065W		173
YOR276W	*CAF20*	174
YLR466C-B		175
YOR082C		176
YJL188C	*BUD19*	177
YBR223W-A		178
YGL147C	*RPL9A*	179
YER050C	*RSM18*	180
YER048C	*CAJ1*	181
YBR109C	*CMD1*	182
YOR078W	*BUD21*	183
YLR185W	*RPL37A*	184
YOR146W		185
YIL012W		186
YCL063W	*VAC17*	187
YER087C-B	*SBH1*	188
YGL149W		189
YKL068W-A		190
YLR052W	*IES3*	191
YDR331W	*GPI8*	192
YPL267W	*ACM1*	193
YMR307W	*GAS1*	194
YAL016C-A		195
YGR063C	*SPT4*	196
YOR261C	*RPN8*	197
YBR084C-A	*RPL19A*	198
YDL211C		199
YJL168C	*SET2*	200

Table 27. Genes with the top 200 ratios of H3K4me3 (P16V/WTH3).

Shown are the systematic and common gene names and the corresponding ranks after ordering all the genes in the genome according to (from largest to smallest) the ratio of the average H3K4me3 across the transcription unit in the P16V strain relative to the WTH3 strain.

Genes with bottom 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YDL024C	*DIA3*	6373
YDL222C	*FMP45*	6374
YFL065C		6375
YPL088W		6376
YHR180W		6377
YFR023W	*PES4*	6378
YCR101C		6379
YLR198C		6380
YMR105W-A		6381
YPR200C	*ARR2*	6382
YJL161W	*FMP33*	6383
YHL046C	*PAU13*	6384
YBR116C		6385
YKR009C	*FOX2*	6386
YHR053C	*CUP1-1*	6387
YBR072W	*HSP26*	6388
YOR338W		6389
YFR020W		6390
YGR131W	*FHN1*	6391
YBL100W-C		6392
YIL177W-A		6393
YPL186C	*UIP4*	6394
YDR059C	*UBC5*	6395
YPR002W	*PDH1*	6396
YHL049C		6397
YDR215C		6398
YDL210W	*UGA4*	6399
YJR115W		6400
YLR164W		6401
YDR248C		6402
YJL219W	*HXT9*	6403
YGL033W	*HOP2*	6404
YJL052W	*TDH1*	6405
YHR140W		6406
YGL088W		6407
YLR081W	*GAL2*	6408
YBR201C-A		6409
YBL039W-B		6410
YNL274C	*GOR1*	6411
YIL120W	*QDR1*	6412
YJL170C	*ASG7*	6413
YOR339C	*UBC11*	6414
YGL096W	*TOS8*	6415
YBL029W		6416
YGR294W	*PAU12*	6417
YCL048W-A		6418
YJR094C	*IME1*	6419
YNL200C		6420
YDR285W	*ZIP1*	6421
YLR297W		6422
YJL043W		6423
YOR268C		6424
YHR184W	*SSP1*	6425
YHR001W-A	*QCR10*	6426
YCL026C-A	*FRM2*	6427
YIL072W	*HOP1*	6428
YNL259C	*ATX1*	6429
YKR097W	*PCK1*	6430
YNL012W	*SPO1*	6431
YLR466C-A		6432
YNL042W-B		6433
YGR197C	*SNG1*	6434
YBR183W	*YPC1*	6435
YDR014W-A	*HED1*	6436

Genes with bottom 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YMR170C	*ALD2*	6437
YHL032C	*GUT1*	6438
YPR001W	*CIT3*	6439
YJR151W-A		6440
YLR464W		6441
YER015W	*FAA2*	6442
YPR036W-A		6443
YAL026C-A		6444
YLR070C	*XYL2*	6445
YDR042C		6446
YLR302C		6447
YLR463C		6448
YDR533C	*HSP31*	6449
YLR121C	*YPS3*	6450
YNL194C		6451
YLL064C	*PAU18*	6452
YER053C-A		6453
YML131W		6454
YPL264C		6455
YOL052C-A	*DDR2*	6456
YHR044C	*DOG1*	6457
YML118W	*NGL3*	6458
YOR003W	*YSP3*	6459
YOR032W-A		6460
YEL076C		6461
YDR085C	*AFR1*	6462
YEL057C		6463
YHR054C		6464
YFR017C		6465
YJL038C	*LOH1*	6466
YPR007C	*REC8*	6467
YBR255C-A		6468
YIL014C-A		6469
YLL057C	*JLP1*	6470
YFR026C	*ULI1*	6471
YIR030C	*DCG1*	6472
YEL076C-A		6473
YOR177C	*MPC54*	6474
YNR076W	*PAU6*	6475
YOL156W	*HXT11*	6476
YGL006W-A		6477
YDR317W	*HIM1*	6478
YBR300C		6479
YLR040C		6480
YDR401W		6481
YGL007W	*BRP1*	6482
YPR064W		6483
YDL244W	*THI13*	6484
YJL132W		6485
YHR041C	*SRB2*	6486
YDR278C		6487
YLR041W		6488
YBR117C	*TKL2*	6489
YKR106W		6490
YJR157W		6491
YKL178C	*STE3*	6492
YHL012W		6493
YBR045C	*GIP1*	6494
YMR174C	*PAI3*	6495
YOR237W	*HES1*	6496
YOL164W-A		6497
YOL132W	*GAS4*	6498
YLR365W		6499
YLR227C	*ADY4*	6500

Genes with bottom 200 H3K4me3 ratios (P16V v WTH3)	Alias	Rank
YBL098W	*BNA4*	6501
YCL049C		6502
YJL045W		6503
YJL223C	*PAU1*	6504
YDR015C		6505
YDR079C-A	*TFB5*	6506
YFL040W		6507
YAL062W	*GDH3*	6508
YNL140C		6509
YOR190W	*SPR1*	6510
YHR015W	*MIP6*	6511
YNL033W		6512
YJR159W	*SOR1*	6513
YDR516C	*EMI2*	6514
YEL049W	*PAU2*	6515
YHR209W	*CRG1*	6516
YOR381W-A		6517
YFL060C	*SNO3*	6518
YFR012W		6519
YAL068C	*PAU8*	6520
YLR120W-A		6521
YOR071C	*NRT1*	6522
YEL010W		6523
YLR307W	*CDA1*	6524
YNL195C		6525
YHR014W	*SPO13*	6526
YDL247W	*MPH2*	6527
YIR028W	*DAL4*	6528
YJR156C	*THI11*	6529
YFL055W	*AGP3*	6530
YJL037W	*IRC18*	6531
YPR201W	*ARR3*	6532
YGL184C	*STR3*	6533
YDR119W-A		6534
YHR214C-D		6535
YNL334C	*SNO2*	6536
YKL177W		6537
YEL009C-A		6538
YOL155W-A		6539
YHL048C-A		6540
YPL171C	*OYE3*	6541
YOL160W		6542
YER044C-A	*MEI4*	6543
YJR160C	*MPH3*	6544
YKL026C	*GPX1*	6545
YGL194C-A		6546
YNL128W	*TEP1*	6547
YDR540C	*IRC4*	6548
YGR256W	*GND2*	6549
YHR177W		6550
YOL013W-B		6551
YOR072W-A		6552
YDR403W	*DIT1*	6553
YOL152W	*FRE7*	6554
YOR235W	*IRC13*	6555
YJL222W-B		6556
YNL019C		6557
YOL163W		6558
YGL258W	*VEL1*	6559
YNL333W	*SNZ2*	6560
YBL107W-A		6561
YGR122C-A		6562
YDR504C	*SPG3*	6563
YDR225W	*HTA1*	6564
YBR250W	*SPO23*	6565
YCR001W		6566
YNL030W	*HHF2*	6567
YNR034W-A		6568
YOR072W		6569
YLL046C	*RNP1*	6570
YAR053W		6571
YLL047W		6572

Table 28. Genes with the bottom 200 ratios of H3K4me3 (P16V/WTH3).

Shown are the systematic and common gene names and the corresponding ranks after ordering all the genes in the genome according to (from largest to smallest) the ratio of the average H3K4me3 across the transcription unit in the P16V strain relative to the WTH3 strain.

Systematic gene name	Alias	Description
YAL061W	*BDH2*	putative medium-chain alcohol dehydrogenase
YBL024W	*TRM4*	S-adenosylmethionine-dependent tRNA methyltransferase
YBR030W	*RKM3*	SET-domain ribosomal lysine methyltransferase
YBR034C	*HMT1/ RMT1*	nuclear SAM-dependent mono- and asymmetric arginine dimethylating methyltransferase; also methylates ribosomal protein Rps2p
YBR133C	*HSL7*	protein arginine N-methyltransferase
YBR141C	-	putative S-adenosylmethionine-dependent methyltransferase
YBR225W	-	putative protein of unknown function
YBR261C	*TAE1*	putative S-adenosylmethionine-dependent methyltransferase
YBR271W	-	putative S-adenosylmethionine-dependent methyltransferase
YCL055W	*KAR4*	transcription factor required for gene regulation in response to pheromones
YDL033C	*SLM3*	tRNA-specific 2-thiouridylase
YDR083W	*RRP8*	rRNA processing enzyme
YDR140W	*MTQ2*	S-adenosylmethionine-dependent methyltransferase
YDR198C	*RKM2*	SET-domain ribosomal lysine methyltransferase
YDR257C	*RKM4/ SET7*	SET-domain ribosomal lysine methyltransferase
YDR316W	*OMS1*	putative mitochondrial methyltransferase
YDR410C	*STE14*	farnesyl cysteine-carboxyl methyltransferase
YDR435C	*PPM1*	carboxyl methyltransferase
YDR440W	*DOT1*	histone methyltransferase for H3K79
YDR465C	*RMT2*	ribosomal arginine methyltransferase
YER175C	*TMT1*	trans-aconitate methyltransferase
YGR001C	*AML1*	putative methyltransferase
YGR283C	-	protein of unknown function
YGR286C	*BIO2*	biotin synthase
YHL039W	-	SET-domain putative protein of unknown function
YHR109W	*CTM1*	cytochrome c lysine methyltransferase
YHR119W	*SET1*	SET-domain histone methyltransferase for H3K4
YHR207C	*SET5*	SET-domain protein of unknown function
YHR209W	*CRG1*	putative S-adenosylmethionine-dependent methyltransferase
YIL064W	*SEE1*	putative S-adenosylmethionine-dependent methyltransferase
YIL110W	*MNI1*	putative S-adenosylmethionine-dependent methyltransferase
YJL105W	*SET4*	SET-domain protein of unknown function
YJL168C	*SET2*	SET-domain histone methyltransferase for H3K36
YJR129C	-	putative S-adenosylmethionine-dependent methyltransferase
YKL155C	*RSM22*	putative mitochondrial ribosomal S-adenosylmethionine-dependent methyltransferase
YKL162C	-	putative protein of unknown function
YKR029C	*SET3*	SET-domain component of Set3 deacetylase complex
YKR056W	*TRM2*	tRNA methyltransferase
YKR069W	*MET1*	S-adenosylmethionine uroporphyrinogen III transmethylase
YLL062C	*MHT1*	S-methylmethionine-homocysteine methyltransferase
YLR063W	-	putative S-adenosylmethionine-dependent methyltransferase
YLR137W	-	putative S-adenosylmethionine-dependent methyltransferase
YLR172C	*DPH5*	methyltransferase required for diphthamide synthesis
YLR285W	*NNT1*	putative nicotinamide N-methyltransferase
YML005W	*TRM12*	S-adenosylmethionine-dependent tRNA methyltransferase
YML008C	*ERG6*	delta(24)-sterol C-methyltransferase
YML110C	*COQ5*	2-hexaprenyl-6-methoxy-1,4-benzoquinone methyltransferase (ubiquinone biosynthesis)
YMR209C	-	putative S-adenosylmethionine-dependent methyltransferase
YMR228W	*MTF1*	putative S-adenosylmethionine-dependent methyltransferase
YMR310C	-	putative protein of unknown function
YNL022C	-	putative cytosine 5-methyltransferase

Systematic gene name	Alias	Description
YNL024C	-	putative protein of unknown function with methyltransferase motif
YNL063W	*MTQ1*	S-adenosylmethionine-dependent methyltransferase
YNL092W	-	putative S-adenosylmethionine-dependent methyltransferase
YNR029C	-	putative protein of unknown function
YOL093W	*TRM10*	tRNA methyltransferase
YOL096C	*COQ3*	O-methyltransferase (ubiquinone biosynthesis)
YOL124C	*TRM11*	S-adenosylmethionine-dependent tRNA methyltransferase
YOR021C	-	putative protein of unknown function
YOR196C	*LIP5*	protein involved in biosynthesis of the coenzyme lipoic acid
YOR239W	*ABP140*	putative S-adenosylmethionine-dependent methyltransferase
YPL017C	*IRC15*	putative S-adenosylmethionine-dependent methyltransferase
YPL086C	*ELP3*	subunit of Elongator complex with HAT activity
YPL165C	*SET6*	SET-domain protein of unknown function
YPL208W	*RKM1*	SET-domain ribosomal lysine methyltransferase
YPL273W	*SAM4*	S-adenosylmethionine-homocysteine methyltransferase

Table 29. List of known and putative methyltransferases tested for loss of H3K18me1 (Chapter 6).

All descriptions are taken from the *Saccharomyces* genome database (SGD: www.yeastgenome.org).

9. Bibliography

Ahmad M, Hamid A, Hussain A, Majeed R, Qurishi Y, Bhat JA, Najar RA, Qazi AK, Zargar MA, Singh SK, et al. 2012. Understanding histone deacetylases in the cancer development and treatment: an epigenetic perspective of cancer chemotherapy. *DNA and cell biology* **00**.

Ahn SH, Kim M, and Buratowski S. 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Molecular cell* **13**: 67–76.

Akhtar MS, Heidemann M, Tietjen JR, Zhang DW, Chapman RD, Eick D, and Ansari AZ. 2009. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Molecular cell* **34**: 387–93.

Albig AR, and Decker CJ. 2001. The target of rapamycin signaling pathway regulates mRNA turnover in the yeast Saccharomyces cerevisiae. *Molecular biology of the cell* **12**: 3428–38.

Allard S, Utley RT, Savard J, Clarke AS, Grant PA, Brandl CJ, Pillus L, Workman JL, and Côté J. 1999. NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *The EMBO Journal* **18**: 5108–5119.

Allen C, Büttner S, Aragon AD, Thomas JA, Meirelles O, Jaetao JE, Benn D, Ruby SW, Veenhuis M, Madeo F, et al. 2006. Isolation of quiescent and nonquiescent cells from yeast stationary-phase cultures. *The journal of cell biology* **174**: 89–100.

Allison LA, Moyle M, Shales M, and Ingles CJ. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**: 599–610.

Anderson JD, Lowary PT, and Widom J. 2001. Effects of histone acetylation on the equilibrium accessibility of nucleosomal DNA target sites. *Journal of molecular biology* **307**: 977–85.

Arigo JT, Carroll KL, Ames JM, and Corden JL. 2006a. Regulation of yeast NRD1 expression by premature transcription termination. *Molecular cell* **21**: 641–51.

Arigo JT, Eyler DE, Carroll KL, and Corden JL. 2006b. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Molecular cell* **23**: 841–51.

Armache K-J, Mitterweger S, Meinhart A, and Cramer P. 2005. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *The journal of biological chemistry* **280**: 7131–4.

Arévalo-Rodríguez M, Cardenas ME, Wu X, Hanes SD, and Heitman J. 2000. Cyclophilin A and Ess1 interact with and regulate silencing by the Sin3-Rpd3 histone deacetylase. *The EMBO journal* **19**: 3739–49.

Banerjee T, and Chakravarti D. 2011. A peek into the complex realm of histone phosphorylation. *Molecular and cellular biology* **31**: 4858–73.

Bannister AJ, and Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell research* **21**: 381–95.

Banères JL, Martin A, and Parello J. 1997. The N tails of histones H3 and H4 adopt a highly structured conformation in the nucleosome. *Journal of molecular biology* **273**: 503–8.

Bataille AR, Jeronimo C, Jacques P-É, Laramée L, Fortin M-È, Forest A, Bergeron M, Hanes SD, and Robert F. 2012. A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Molecular cell* **45**: 158–70.

Batta K, Zhang Z, Yen K, Goffman DB, and Pugh BF. 2011. Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes & development* **25**: 2254–2265.

Beck C, and Von Meyenburg HK. 1968. Enzyme pattern and aerobic growth of Saccharomyces cerevisiae under various degrees of glucose limitation. *Journal of Bacteriology* **96**: 479.

Berger SL. 2007. The complex language of chromatin regulation during transcription. *Nature* **447**: 407–12.

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides Tony, and Schreiber SL. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *PNAS* **99**: 8695–700.

Berretta J, and Morillon A. 2009. Pervasive transcription constitutes a new level of eukaryotic genome regulation. *EMBO reports* **10**: 973–82.

Berretta J, Pinskaya M, and Morillon A. 2008. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. *Genes & development* **22**: 615–26.

Bhaumik SR, and Green MR. 2001. SAGA is an essential in vivo target of the yeast acidic activator Gal4p. *Genes & development* **15**: 1935–45.

Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Barsyte D, Li J, Wu Minhao, Zhou BO, et al. 2011. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *The EMBO journal* **30**: 2829–2842.

Bitoun E, Oliver PL, and Davies KE. 2007. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Human molecular genetics* **16**: 92–106.

Bolden JE, Peart MJ, and Johnstone RW. 2006. Anticancer activities of histone deacetylase inhibitors. *Nature reviews. Drug discovery* **5**: 769–84.

Brauer MJ, Huttenhower C, Airoldi EM, Rosenstein R, Matese JC, Gresham D, Boer VM, Troyanskaya OG, and Botstein D. 2008. Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Molecular biology of the cell* **19**: 352–367.

Breitkreutz A, Choi H, Sharom JR, Boucher L, Neduva V, Larsen B, Lin Z-Y, Breitkreutz B-J, Stark C, Liu G, et al. 2010. A global protein kinase and phosphatase interaction network in yeast. *Science* **328**: 1043–6.

Briggs S, Bryk M, Strahl BD, Cheung WL, Davie JK, Dent SYR, Winston F, and Allis CD. 2001. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. *Genes & development* **15**: 3286–95.

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, and Allis CD. 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**: 843–51.
Buratowski S. 2009. Progression through the RNA polymerase II CTD cycle. *Molecular cell* **36**: 541–6.
Buratowski S, and Kim T. 2011. The role of co-transcriptional histone methylations. *Cold Spring Harb Symp Quant Biol* **75**: 95–102.
Burgess RJ, Zhou H, Han J, and Zhang Z. 2010. A role for Gcn5 in replication-coupled nucleosome assembly. *Molecular cell* **37**: 469–80.
Cai L, Sutter BM, Li B, and Tu BP. 2011. Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Molecular cell* **42**: 426–37.
Cai L, and Tu BP. 2012. Driving the Cell Cycle Through Metabolism. *Annual review of cell and developmental biology* 1–29.
Carroll KL, Ghirlando R, Ames JM, and Corden JL. 2007. Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA* **13**: 361–73.
Carroll KL, Pradhan DA, Granek JA, Clarke ND, and Corden JL. 2004. Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. *Molecular and cellular biology* **24**: 6241–6252.
Carrozza MJ, Florens L, Swanson SK, Shia W-J, Anderson S, Yates JRI, Washburn MP, and Workman JL. 2005a. Stable incorporation of sequence specific repressors Ash1 and Ume6 into the Rpd3L complex. *Biochimica et biophysica acta* **1731**: 77–87.
Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia W-J, Anderson S, Yates JRI, Washburn MP, et al. 2005b. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581–92.
Carvin CD, and Kladde MP. 2004. Effectors of lysine 4 methylation of histone H3 in Saccharomyces cerevisiae are negative regulators of PHO5 and GAL1-10. *The journal of biological chemistry* **279**: 33057–62.
Causton HC, Ren B, Koh SS, Harbison CT, Kanin E, Jennings EG, Lee TI, True HL, Lander ES, and Young RA. 2001. Remodeling of yeast genome expression in response to environmental changes. *Molecular biology of the cell* **12**: 323–37.
Chance B, Estabrook RW, and Ghosh AK. 1964. Dampened sinusoidal oscillations of cytoplasmic reduced pyridine nucleotide in yeast cells. *PNAS* **51**: 1244–1251.
Chandrasekharan MB, Huang F, Chen Y-C, and Sun Z-W. 2010. Histone H2B C-terminal helix mediates trans-histone H3K4 methylation independent of H2B ubiquitination. *Molecular and cellular biology* **30**: 3216–32.
Chandrasekharan MB, Huang F, and Sun Z-W. 2009. Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *PNAS* **106**: 16686–91.
Chang B, Chen Y, Zhao Y, and Bruick RK. 2007. JMJD6 is a histone arginine demethylase. *Science (New York, N.Y.)* **318**: 444–7.
Charles GM, Chen C, Shih SC, Collins SR, Beltrao P, Zhang Xin, and Sharma T. 2011. Site-specific acetylation mark on an essential chromatin-remodeling complex promotes resistance to replication stress. *PNAS* **108**: 10620–10625.
Cheng H, He X, and Moore CL. 2004. The essential WD repeat protein Swd2 has dual functions in RNA polymerase II transcription termination and lysine 4 methylation of histone H3. *Molecular and cellular biology* **24**: 2932–2943.
Chern M-K, Chang K-N, Liu L-F, Tam T-CS, Liu Y-C, Liang Y-L, and Tam MF. 2002. Yeast ribosomal protein L12 is a substrate of protein-arginine methyltransferase 2. *The journal of biological chemistry* **277**: 15345–53.
Cho E-J, Kobor MS, Kim M, Greenblatt JF, and Buratowski S. 2001. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & development* **15**: 3319–29.
Cho RJ, Campbell MJ, Winzeler EA, Steinmetz LM, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ, et al. 1998. A genome-wide transcriptional analysis of the mitotic cell cycle. *Molecular cell* **2**: 65–73.
Churchman LS, and Weissman JS. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**: 368–73.
Clarke AS, Lowell JE, and Jacobson SJ. 1999. Esa1p is an essential histone acetyltransferase required for cell cycle progression. *Molecular and cellular biology* **19**: 2515–2526.
Clements A, Poux AN, Lo W-S, Pillus Lorraine, Berger SL, and Marmorstein R. 2003. Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. *Molecular cell* **12**: 461–73.
Cramer P, Bushnell DA, and Kornberg RD. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863–76.
Creamer TJ, Darby MM, Jamonnak N, Schaughency P, Hao H, Wheelan SJ, and Corden JL. 2011. Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS genetics* **7**: e1002329.
Dai J, Hyland EM, Yuan DS, Huang H, Bader JS, and Boeke JD. 2008. Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. *Cell* **134**: 1066–78.
Daniel JA, Torok MS, Sun Z-W, Schieltz D, Allis CD, Yates JRI, and Grant PA. 2004. Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *The journal of biological chemistry* **279**: 1867–71.
Darby MM, Serebreni L, Pan X, Boeke JD, and Corden JL. 2012. The Saccharomyces cerevisiae Nrd1-Nab3 transcription termination pathway acts in opposition to Ras signaling and mediates response to nutrient depletion. *Molecular and cellular biology* **32**: 1762–1775.
David L, Huber W, Granovskaia MV, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, and Steinmetz LM. 2006. A high-resolution map of transcription in the yeast genome. *PNAS* **103**: 5320–5.
Dehé P-M, Dichtl B, Schaft D, Roguev A, Pamblanco M, Lebrun R, Rodríguez-Gil A, Mkandawire M, Landsberg K, Shevchenko A, et al. 2006. Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *The journal of biological chemistry* **281**: 35404–12.
Dehé P-M, Pamblanco M, Luciano P, Lebrun R, Moinier D, Sendra R, Verreault A, Tordera V, and Géli V. 2005. Histone H3 lysine 4 mono-methylation does not require ubiquitination of histone H2B. *Journal of molecular biology* **353**: 477–84.

Dichtl B, Aasland R, and Keller W. 2004. Functions for S. cerevisiae Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. *RNA* **10**: 965–977.
Van Dijk EL, Chen CL, D'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legoix-Né P, Loeillet S, et al. 2011. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475**: 114–117.
Dillon SC, Zhang X, Trievel RC, and Cheng X. 2005. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome biology* **6**: 227.
Dilova I, Chen C-Y, and Powers T. 2002. Mks1 in concert with TOR signaling negatively regulates RTG target gene expression in S. cerevisiae. *Current biology* **12**: 389–95.
Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, and Rando OJ. 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**: 1405–8.
Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD, and Roeder RG. 2006. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nature structural & molecular biology* **13**: 713–9.
Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, and Shilatifard A. 2002. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *The journal of biological chemistry* **277**: 28368–71.
Du H-N, and Briggs S. 2010. A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. *The journal of biological chemistry* **285**: 11704–13.
Du H-N, Fingerman IM, and Briggs S. 2008. Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. *Genes & development* **22**: 2786–98.
Duina AA, Marsh JA, and Gaber RF. 1996. Identification of two CyP-40-like cyclophilins in Saccharomyces cerevisiae, one of which is required for normal growth. *Yeast* **12**: 943–952.
Eberharter A, Sterner DE, Schieltz D, Hassan AH, Yates JRI, Berger SL, and Workman JL. 1999. The ADA complex is a distinct histone acetyltransferase complex in Saccharomyces cerevisiae. *Molecular and cellular biology* **19**: 6621–6631.
Edmondson DG, Smith M M, and Roth SY. 1996. Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes & development* **10**: 1247–1259.
Emre NCT, Ingvarsdottir K, Wyce A, Wood A, Krogan NJ, Henry KW, Li K, Marmorstein R, Greenblatt JF, Shilatifard A, et al. 2005. Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing. *Molecular cell* **17**: 585–94.
Endo H, Nakabayashi Y, Kawashima S, Enomoto T, Seki M, and Horikoshi M. 2012. Nucleosome surface containing nucleosomal DNA entry/exit site regulates H3-K36me3 via association with RNA polymerase II and Set2. *Genes to cells* **17**: 65–81.
Espinosa MC, Rehman MA, Chisamore-Robert P, Jeffery D, and Yankulov K. 2010. GCN5 is a positive regulator of origins of DNA replication in Saccharomyces cerevisiae. *PloS one* **5**: e8964.
Faravelli F. 2005. NSD1 mutations in Sotos syndrome. *American journal of medical genetics* **137C**: 24–31.
Feaver WJ, Svejstrup JQ, Henry NL, and Kornberg RD. 1994. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell* **79**: 1103–9.
Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y, Hill C, and Carolina N. 2002. Methylation of H3-Lysine 79 is mediated by a new family of HMTases without a SET domain. *Current Biology* **12**: 1052–1058.
Ferrando AA, Armstrong SA, Neuberg DS, Sallan SE, Silverman LB, Korsmeyer SJ, and Look AT. 2003. Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. *Blood* **102**: 262–8.
Fierz B, Chatterjee C, McGinty RK, Bar-Dagan M, Raleigh DP, and Muir TW. 2011. Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nature chemical biology* **7**: 113–9.
Fingerman IM, Li H-C, and Briggs S. 2007. A charge-based interaction between histone H4 and Dot1 is required for H3K79 methylation and telomere silencing: identification of a new trans-histone pathway. *Genes & development* **21**: 2018–29.
Fingerman IM, Wu C-L, Wilson BD, and Briggs S. 2005. Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in Saccharomyces cerevisiae. *The journal of biological chemistry* **280**: 28761–5.
Fischer G, Bang H, Berger E, and Schellenberger A. 1984. Conformational specificity of chymotrypsin toward proline-containing substrates. *Biochimica et biophysica acta* **791**: 87–97.
Fleming AB, Kao C-F, Hillyer C, Pikaart M, and Osley MA. 2008. H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Molecular cell* **31**: 57–66.
Frederiks F, Tzouros M, Oudgenoeg G, Van Welsem T, Fornerod M, Krijgsveld J, and Van Leeuwen F. 2008. Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. *Nature structural & molecular biology* **15**: 550–7.
Gansheroff LJ, Dollard C, Tan P, and Winston F. 1995. The Saccharomyces cerevisiae SPT7 gene encodes a very acidic protein important for transcription in vivo. *Genetics* **139**: 523–536.
Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, Shabanowitz J, Mishra N, Strahl BD, Allis CD, et al. 2007. Organismal differences in post-translational modifications in histones H3 and H4. *The journal of biological chemistry* **282**: 7641–55.
Gardner KE, Allis CD, and Strahl BD. 2011. Operating on chromatin, a colorful language where context matters. *Journal of molecular biology* **409**: 36–46.
Gardner RG, Nelson ZW, and Gottschling DE. 2005. Ubp10 / Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin. *Molecular and cellular biology* **25**: 6123–6139.
Garkavtsev I, Kozin S V, Chernova O, Xu L, Winkler F, Brown E, Barnett GH, and Jain RK. 2004. The candidate tumour suppressor protein ING4 regulates brain tumour growth and angiogenesis. *Nature* **428**: 328–32.

Gary JD, Lin WJ, Yang MC, Herschman HR, and Clarke S. 1996. The predominant protein-arginine methyltransferase from Saccharomyces cerevisiae. *The journal of biological chemistry* **271**: 12585–94.
Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, and Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular biology of the cell* **11**: 4241–57.
Ghosh AK, Chance B, and Pye EK. 1971. Metabolic coupling and synchronization of NADH oscillations in yeast cell populations. *Archives of biochemistry and biophysics* **145**: 319–331.
Glaser S, Schaft J, Lubitz S, Vintersten K, Van der Hoeven F, Tufteland KR, Aasland R, Anastassiadis K, Ang S-L, and Stewart AF. 2006. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**: 1423–32.
Glaves R, Baer M, Schreiner E, Stoll R, and Marx D. 2008. Conformational dynamics of minimal elastin-like polypeptides: the role of proline revealed by molecular dynamics and nuclear magnetic resonance. *Chemphyschem* **9**: 2759–65.
Govind CK, Zhang F, Qiu H, Hofmeyer K, and Hinnebusch AG. 2007. Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Molecular cell* **25**: 31–42.
Granovskaia M V, Jensen LJ, Ritchie ME, Toedling J, Ning Y, Bork P, Huber W, and Steinmetz LM. 2010. High-resolution transcription atlas of the mitotic cell cycle in budding yeast. *Genome biology* **11**: R24.
Grant PA, Duggan LJ, Côté J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, et al. 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & development* **11**: 1640–50.
Grant PA, Eberharter A, John S, Cook RG, Turner BM, and Workman JL. 1999. Expanded lysine acetylation specificity of Gcn5 in native complexes. *The journal of biological chemistry* **274**: 5895–900.
Green EM, Mas G, Young NL, Garcia BA, and Gozani O. 2012. Methylation of H4 lysines 5, 8 and 12 by yeast Set5 calibrates chromatin stress responses. *Nature structural & molecular biology* **19**: 361–3.
Guccione E, Bassi C, Casadio F, Martinato F, Cesaroni M, Schuchlautz H, Lüscher B, and Amati B. 2007. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature* **449**: 933–7.
Gudipati RK, Villa T, Boulay J, and Libri D. 2008. Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nature structural & molecular biology* **15**: 786–94.
Guillemette B, Drogaris P, Lin H-HS, Armstrong H, Hiragami-Hamada K, Imhof A, Bonneil E, Thibault P, Verreault A, and Festenstein RJ. 2011. H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. *PLoS genetics* **7**: e1001354.
Gunasekaran K, Nagarajaram HA, Ramakrishnan C, and Balaram P. 1998. Stereochemical punctuation marks in protein structures: glycine and proline containing helix stop signals. *Journal of molecular biology* **275**: 917–32.
Hahn S, and Young ET. 2011. Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* **189**: 705–36.
Hainer SJ, and Martens JA. 2011. Identification of histone mutants that are defective for transcription-coupled nucleosome occupancy. *Molecular and cellular biology* **31**: 3557–68.
Halbach A, Zhang H, Wengi A, Jablonska Z, Gruber IML, Halbeisen RE, Dehé P-M, Kemmeren P, Holstege FCP, Géli V, et al. 2009. Cotranslational assembly of the yeast SET1C histone methyltransferase complex. *The EMBO journal* **28**: 2959–70.
Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li Meihua, Yagyu R, and Nakamura Y. 2004. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nature cell biology* **6**: 731–40.
Hanes SD, Shank PR, and Bostian KA. 1989. Sequence and mutational analysis of ESS1, a gene essential for growth in Saccharomyces cerevisiae. *Yeast* **5**: 55–72.
Hani J, Stumpf G, and Domdey H. 1995. PTF1 encodes an essential protein in Saccharomyces cerevisiae, which shows strong homology with a new putative family of PPIases. *FEBS letters* **365**: 198–202.
Hassan AH, Awad S, Al-Natour Z, Othman S, Mustafa F, and Rizvi TA. 2007. Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. *The Biochemical journal* **402**: 125–33.
Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, and Workman JL. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111**: 369–79.
Hazzalin CA, and Mahadevan LC. 2005. Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun. *PLoS biology* **3**: e393.
He C, Xu J, Zhang J, Xie D, Ye H, Xiao Z, Cai M, Xu K, Zeng Y, Li Haigang, et al. 2012. High expression of trimethylated histone H3 lysine 4 is associated with poor prognosis in hepatocellular carcinoma. *Human pathology* 1–11.
Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics* **39**: 311–8.
Henry KW, Wyce A, Lo W-S, Duggan LJ, Emre NCT, Kao C-F, Pillus L, Shilatifard A, Osley MA, and Berger SL. 2003. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes & development* **17**: 2648–63.
Henry MF, and Silver PA. 1996. A novel methyltransferase (Hmt1p) modifies poly (A)+-RNA-binding proteins. *Molecular and cellular biology* **16**: 3668–3678.
Van Heusden GPH. 2009. 14-3-3 Proteins: insights from genome-wide studies in yeast. *Genomics* **94**: 287–93.
Hong X, Zang J, White J, Wang C, Pan C, Zhao R, Murphy RC, Dai S, Henson P, Kappler JW, et al. 2010. Interaction of JMJD6 with single-stranded RNA. *PNAS* **107**: 14568–14572.
Houseley J, Rubbi L, Grunstein M, Tollervey D, and Vogelauer M. 2008. A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Molecular cell* **32**: 685–95.
Howe L, Auston D, Grant PA, John S, Cook RG, Workman JL, and Pillus L. 2001. Histone H3 specific acetyltransferases are essential for cell cycle progression. *Genes & development* **15**: 3144–54.

Howe L, Kusch T, Muster N, Chaterji R, Yates JRI, and Workman JL. 2002. Yng1p modulates the activity of Sas3p as a component of the yeast NuA3 histone acetyltransferase complex. *Molecular and cellular biology* **22**: 5047–5053.
Huang F, Chandrasekharan MB, Chen Y-C, Bhaskara S, Hiebert SW, and Sun Z-W. 2010. The JmjN domain of Jhd2 is important for its protein stability, and the plant homeodomain (PHD) finger mediates its chromatin association independent of H3K4 methylation. *The journal of biological chemistry* **285**: 24548–61.
Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC, Hayes DN, Shanmugam KS, Bhattacharjee A, Biondi CA, et al. 2004. Menin associates with a trithorax family histone methyltransferase complex and with the Hoxc8 locus. *Cell* **13**: 587–597.
Huisinga KL, and Pugh BF. 2004. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. *Molecular cell* **13**: 573–85.
Hwang WW, Venkatasubrahmanyam S, Ianculescu AG, Tong A, Boone C, and Madhani HD. 2003. A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Molecular cell* **11**: 261–266.
Hyland EM, Molina H, Poorey K, Jie C, Xie Z, Dai J, Qian J, Bekiranov S, Auble DT, Pandey A, et al. 2011. An evolutionarily "young" lysine residue in histone H3 attenuates transcriptional output in Saccharomyces cerevisiae. *Genes & development* **25**: 1306–1319.
Hyllus D, Stein C, Schnabel K, Schiltz E, Imhof A, Dou Y, Hsieh J, and Bauer U-M. 2007. PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes & development* **21**: 3369–80.
Iberg AN, Espejo A, Cheng D, Kim D, Michaud-Levesque J, Richard S, and Bedford MT. 2008. Arginine methylation of the histone H3 tail impedes effector binding. *The journal of biological chemistry* **283**: 3006–10.
Imai S, Armstrong CM, Kaeberlein M, and Guarente L. 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403**: 795–800.
Ingvarsdottir K, Edwards C, Lee MG, Lee J-S, Schultz DC, Shilatifard A, Shiekhattar R, and Berger SL. 2007. Histone H3 K4 demethylation during activation and attenuation of GAL1 transcription in Saccharomyces cerevisiae. *Molecular and cellular biology* **27**: 7856–64.
Iwase S, Lan F, Bayliss P, De la Torre-Ubieta L, Huarte M, Qi HH, Whetstine JR, Bonni A, Roberts TM, and Shi Y. 2007. The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* **128**: 1077–88.
Iyer NG, Ozdag H, and Caldas C. 2004. p300/CBP and cancer. *Oncogene* **23**: 4225–31.
Jansen A, and Verstrepen KJ. 2011. Nucleosome positioning in Saccharomyces cerevisiae. *Microbiology and molecular biology reviews* **75**: 301–20.
Jelinsky SA, and Samson LD. 1999. Global response of Saccharomyces cerevisiae to an alkylating agent. *PNAS* **96**: 1486–91.
Jenuwein T, and Allis CD. 2001. Translating the histone code. *Science* **293**: 1074–80.
Jenuwein T, Laible G, Dorn R, and Reuter G. 1998. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cellular and molecular life sciences* **54**: 80–93.
Jia Y, Rothermel B, Thornton J, and Butow RA. 1997. A basic helix-loop-helix-leucine zipper transcription complex in yeast functions in a signaling pathway from mitochondria to the nucleus. *Molecular and cellular biology* **17**: 1110–7.
Jiang L, Smith JN, Anderson SL, Ma P, Mizzen CA, and Kelleher NL. 2007. Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry. LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. *The journal of biological chemistry* **282**: 27923–34.
John S, Howe L, Tafrov ST, Grant PA, Sternglanz R, and Workman JL. 2000. The Something About Silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF II 30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)–FACT complex. *Genes & development* **14**: 1196–1208.
Johnson ES. 2004. Protein modification by SUMO. *Annual review of biochemistry* **73**: 355–82.
Jorgensen P, and Tyers M. 2004. How cells coordinate growth and division. *Current biology* **14**: R1014–27.
Joshi AA, and Struhl K. 2005. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Molecular cell* **20**: 971–8.
Juan L-J, Walter PP, Taylor ICA, Kingston RE, and Workman JL. 1993. Transcription of DNA and chromatin: nucleosome cores and histone H1 in the binding of GAL4 derivatives and the reactivation of transcription from nucleosome templates in vitro. *Cold Spring Harb Symp Quant Biol* **58**: 213–223.
Kao C-F, Hillyer C, Tsukuda T, Henry KW, Berger SL, and Osley MA. 2004. Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. *Genes & development* **18**: 184–95.
Kasten MM, Dorland S, and Stillman DJ. 1997. A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators. *Molecular and cellular biology* **17**: 4852–4858.
Kasten MM, Szerlong H, Erdjument-Bromage H, Tempst P, Werner M, and Cairns BR. 2004. Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. *The EMBO journal* **23**: 1348–59.
Katan-Khaykovich Y, and Struhl K. 2005. Heterochromatin formation involves changes in histone modifications over multiple cell generations. *The EMBO journal* **24**: 2138–49.
Keogh M-C, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, et al. 2005. Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**: 593–605.
Kettenbach AN, Rush J, and Gerber SA. 2011. Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. *Nature protocols* **6**: 175–86.
Kim J, Guermah M, McGinty RK, Lee J-S, Tang Z, Milne TA, Shilatifard A, Muir TW, and Roeder RG. 2009. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* **137**: 459–71.
Kim J, and Roeder RG. 2009. Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. *The journal of biological chemistry* **284**: 20582–92.
Kim K-Y, and Levin DE. 2011. Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. *Cell* **144**: 745–56.

Kim M, Suh H, Cho E-J, and Buratowski S. 2009. Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. *The journal of biological chemistry* **284**: 26421–6.
Kim M, Vasiljeva L, Rando OJ, Zhelkovsky A, Moore CL, and Buratowski S. 2006. Distinct pathways for snoRNA and mRNA termination. *Molecular cell* **24**: 723–34.
Kim T, and Buratowski S. 2009. Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell* **137**: 259–72.
Kim T, and Buratowski S. 2007. Two Saccharomyces cerevisiae JmjC domain proteins demethylate histone H3 Lys36 in transcribed regions to promote elongation. *The journal of biological chemistry* **282**: 20827–35.
Kim T, Liu CL, Yassour M, Holik J, Friedman N, Buratowski S, and Rando OJ. 2010. RNA polymerase mapping during stress responses reveals widespread nonproductive transcription in yeast. *Genome biology* **11**: R75.
Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, Green RD, and Kouzarides T. 2009. Distinct transcriptional outputs associated with mono- and dimethylated histone H3 arginine 2. *Nature structural & molecular biology* **16**: 449–51.
Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, Vermeulen M, Mann M, Bähler J, Green RD, and Kouzarides T 2007. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* **449**: 928–32.
Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, and Strahl BD. 2005. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular and cellular biology* **25**: 3305–3316.
Klevecz RR, Bolen J, Forrest G, and Murray DB. 2004. A genomewide oscillation in transcription gates DNA replication and cell cycle. *PNAS* **101**: 1200–5.
Klose RJ, Gardner KE, Liang G, Erdjument-Bromage H, Tempst P, and Zhang Y. 2007. Demethylation of histone H3K36 and H3K9 by Rph1: a vestige of an H3K9 methylation system in Saccharomyces cerevisiae? *Molecular and cellular biology* **27**: 3951–61.
Klose RJ, Kallin EM, and Zhang Y. 2006. JmjC-domain-containing proteins and histone demethylation. *Nature reviews. Genetics* **7**: 715–27.
Komarnitsky P. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & development* **14**: 2452–2460.
Komeili A, Wedaman KP, O'Shea EK, and Powers T. 2000. Mechanism of metabolic control. Target of rapamycin signaling links nitrogen quality to the activity of the Rtg1 and Rtg3 transcription factors. *The journal of cell biology* **151**: 863–78.
Koutelou E, Hirsch CL, and Dent SYR. 2010. Multiple faces of the SAGA complex. *Current opinion in cell biology* **22**: 374–82.
Krishnamurthy S, Ghazy MA, Moore CL, and Hampsey M. 2009. Functional interaction of the Ess1 prolyl isomerase with components of the RNA polymerase II initiation and termination machineries. *Molecular and cellular biology* **29**: 2925–34.
Krishnamurthy S, He X, Reyes-Reyes M, Moore CL, and Hampsey M. 2004. Ssu72 is an RNA Polymerase II CTD phosphatase. *Molecular cell* **14**: 387–394.
Kristjuhan A, Walker J, Suka N, Grunstein M, Roberts D, Cairns BR, and Svejstrup JQ. 2002. Transcriptional inhibition of genes with severe histone H3 hypoacetylation in the coding region. *Molecular cell* **10**: 925–33.
Krivtsov A V, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, Xia Xiaobo, Jesneck J, Bracken AP, Silverman LB, et al. 2008. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer cell* **14**: 355–68.
Krogan NJ, Dover J, Khorrami S, Greenblatt JF, Schneider J, Johnston M, and Shilatifard A. 2002. COMPASS, a histone H3 (lysine 4) methyltransferase required for telomeric silencing of gene expression. *The journal of biological chemistry* **277**: 10753–5.
Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, et al. 2003a. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Molecular cell* **11**: 721–729.
Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, et al. 2003b. Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Molecular and cellular biology* **23**: 4207–4218.
Kuehner JN, and Brow DA. 2008. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Molecular cell* **31**: 201–11.
Kuo M-H, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, Roth SY, and Allis CD. 1996. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* **383**: 269–272.
Kurdistani SK, Tavazoie S, and Grunstein M. 2004. Mapping global histone acetylation patterns to gene expression. *Cell* **117**: 721–33.
Kwon D-W, and Ahn SH. 2011. Role of yeast JmjC-domain containing histone demethylases in actively transcribed regions. *Biochemical and biophysical research communications* **410**: 614–619.
LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, and Tollervey D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**: 713–24.
Lakowski TM, and Frankel A. 2008. A kinetic study of human protein arginine N-methyltransferase 6 reveals a distributive mechanism. *The journal of biological chemistry* **283**: 10015–25.
Laribee RN, Krogan NJ, Xiao T, Shibata Y, Hughes TR, Greenblatt JF, and Strahl BD. 2005. BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. *Current Biology* **15**: 1487–93.
Laribee RN, Shibata Y, Mersman DP, Collins SR, Kemmeren P, Roguev A, Weissman JS, Briggs S, Krogan NJ, and Strahl BD. 2007. CCR4/NOT complex associates with the proteasome and regulates histone methylation. *PNAS* **104**: 5836–41.
Larschan E, and Winston F. 2001. The S. cerevisiae SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. *Genes & development* **15**: 1946–56.

Laxman S, Sutter BM, and Tu BP. 2010. Behavior of a metabolic cycling population at the single cell level as visualized by fluorescent gene expression reporters ed. M. Polymenis. *PLoS ONE* **5**: e12595.
Lee DY, Hayes JJ, Pruss D, and Wolffe AP. 1993. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**: 73–84.
Lee J-H, Cook JR, Pollack BP, Kinzy TG, Norris D, and Pestka S. 2000. Hsl7p, the yeast homologue of human JBP1, is a protein methyltransferase. *Biochemical and biophysical research communications* **274**: 105–111.
Lee J-S, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, and Shilatifard A. 2007. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* **131**: 1084–96.
Lee TI, Causton HC, Holstege FCP, Shen WC, Hannett NM, Jennings EG, Winston F, Green MR, and Young RA. 2000. Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* **405**: 701–4.
Van Leeuwen F, Gafken PR, and Gottschling DE. 2002. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* **109**: 745–56.
Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NACH, Margaritis T, Van de Pasch LAL, Van Heesch SAAC, Brok MO, Groot Koerkamp MJA, et al. 2011. The specificity and topology of chromatin interaction pathways in yeast. *Molecular cell* **42**: 536–49.
Li B, Howe L, Anderson S, Yates JRI, and Workman JL. 2003. The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *The journal of biological chemistry* **278**: 8897–903.
Li M, Luo J, Brooks CL, and Gu W. 2002. Acetylation of p53 inhibits its ubiquitination by Mdm2. *The journal of biological chemistry* **277**: 50607–11.
Li Q, Zhou H, Wurtele H, Davies B, Horazdovsky B, Verreault A, and Zhang Z. 2008. Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* **134**: 244–55.
Li S, and Shogren-Knaak MA. 2009. The Gcn5 bromodomain of the SAGA complex facilitates cooperative and cross-tail acetylation of nucleosomes. *The journal of biological chemistry* **284**: 9411–7.
Liang G, Klose RJ, Gardner KE, and Zhang Y. 2007. Yeast Jhd2p is a histone H3 Lys4 trimethyl demethylase. *Nature structural & molecular biology* **14**: 243–5.
Liao X, and Butow RA. 1993. RTG1 and RTG2: two yeast genes required for a novel path of communication from mitochondria to the nucleus. *Cell* **72**: 61–71.
Licatalosi DD, Geiger G, Minet M, Schroeder SC, Cilli K, McNeil JB, and Bentley DL. 2002. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Molecular cell* **9**: 1101–11.
Lin C, Smith E, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, and Shilatifard A. 2010. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Molecular cell* **37**: 429–37.
Lindquist S. 1981. Regulation of protein synthesis during heat shock. *Nature* **293**: 311–314.
Liu CL, Kaplan T, Kim M, Buratowski S, Schreiber SL, Friedman N, and Rando OJ. 2005. Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS biology* **3**: e328.
Liu Yang, Xu X, Singh-Rodriguez S, Zhao Y, and Kuo M-H. 2005. Phosphorylation-independent function of Snf1 and Reg1 proteins rescues a gcn5− mutant in HIS3 expression. *Molecular and cellular biology* **25**: 10566–10579.
Liu Ying, Warfield L, Zhang C, Luo J, Allen J, Lang WH, Ranish J, Shokat KM, and Hahn S. 2009. Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Molecular and cellular biology* **29**: 4852–63.
Liu Z, Sekito T, Spírek M, Thornton J, and Butow RA. 2003. Retrograde signaling is regulated by the dynamic interaction between Rtg2p and Mks1p. *Molecular cell* **12**: 401–11.
Lo W-S, Duggan LJ, Emre NCT, Belotserkovskaya R, Lane WS, Shiekhattar R, and Berger SL. 2001. Snf1- a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* **293**: 1142–6.
Lo W-S, Trievel RC, Rojas JR, Duggan LJ, Hsu JY, Allis CD, Marmorstein R, and Berger SL. 2000. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Molecular cell* **5**: 917–26.
Loewith R, and Hall MN. 2011. Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* **189**: 1177–201.
Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, and Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**: 953–61.
Di Lorenzo A, and Bedford MT. 2011. Histone arginine methylation. *FEBS letters* **585**: 2024–31.
Lu C, Brauer MJ, and Botstein D. 2009. Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. *Molecular biology of the cell* **20**: 891–903.
Lu KP, Finn G, Lee TH, and Nicholson LK. 2007. Prolyl cis-trans isomerization as a molecular timer. *Nature chemical biology* **3**: 619–29.
Luger K, Mäder AW, Richmond RK, Sargent DF, and Richmond TJ. 1997a. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260.
Luger K, Rechsteiner TJ, Flaus AJ, Waye MMY, and Richmond TJ. 1997b. Characterization of nucleosome core particles containing histone proteins made in bacteria. *Journal of molecular biology* **272**: 301–11.
Luger K, and Richmond TJ. 1998. The histone tails of the nucleosome. *Current opinion in genetics & development* **8**: 140–146.
Ma Z, Atencio D, Barnes C, Defiglio H, and Hanes SD. 2012. Multiple roles for the Ess1 prolyl isomerase in the RNA polymerase II transcription cycle. *Molecular and cellular biology* **32**: 3594–3607.
Machné R, and Murray DB. 2012. The yin and yang of yeast transcription: elements of a global feedback system between metabolism and chromatin. *PloS one* **7**: e37906.

Mantri M, Krojer T, Bagg EA, Webby CJ, Butler DS, Kochan G, Kavanagh KL, Oppermann U, McDonough MA, and Schofield CJ. 2010. Crystal structure of the 2-oxoglutarate- and Fe(II)-dependent lysyl hydroxylase JMJD6. *Journal of molecular biology* **401**: 211–222.

Marquardt S, Hazelbaker DZ, and Buratowski S. 2011. Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. *Transcription* **2**: 145–154.

Martin DGE, Baetz K, Shi X, Walter KL, MacDonald VE, Wlodarski MJ, Gozani O, Hieter P, and Howe L. 2006b. The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Molecular and cellular biology* **26**: 7871–9.

Martin DGE, Grimes DE, Baetz K, and Howe L. 2006a. Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. *Molecular and cellular biology* **26**: 3018–3028.

Martínez-Pastor MT, Marchler G, Schüller C, Marchler-Bauer A, Ruis H, and Estruch F. 1996. The Saccharomyces cerevisiae zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *The EMBO journal* **15**: 2227–35.

Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, and Cramer P. 2012. CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science* **336**: 1723–1725.

Meinhart A, Kamenski T, Hoeppner S, Baumli S, and Cramer P. 2005. A structural perspective of CTD function. *Genes & development* **19**: 1401–15.

Mersman DP, Du H-N, Fingerman IM, South PF, and Briggs S. 2011. A charge-based interaction conserved within the H3K4 methyltransferase complexes is needed for protein stability, histone methylation, and gene expression. *The journal of biological chemistry* **287**: 2652–2665.

Mersman DP, Du H-N, Fingerman IM, South PF, and Briggs S. 2009. Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. *Genes & development* **23**: 951–62.

Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J, Ben Abdelali R, Macintyre E, De Braekeleer E, De Braekeleer M, et al. 2009. New insights to the MLL recombinome of acute leukemias. *Leukemia* **23**: 1490–9.

Millar CB, and Grunstein M. 2006. Genome-wide patterns of histone modifications in yeast. *Nature reviews. Molecular cell biology* **7**: 657–66.

Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, and Shilatifard A. 2001. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *PNAS* **98**: 12902–7.

Min J, Feng Q, Li Z, Zhang Y, and Xu R. 2003. Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* **112**: 711–723.

Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, et al. 1996. The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* **87**: 1261–70.

Mochan E, and Pye EK. 1973. Respiratory oscillations in adapting yeast cultures. *Nature: New Biology* **242**: 177–179.

Morillon A, Karabetsou N, Nair A, and Mellor J. 2005. Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. *Molecular cell* **18**: 723–34.

Morris SA, Rao B, Garcia BA, Hake SB, Diaz RL, Shabanowitz J, Hunt DF, Allis CD, Lieb JD, and Strahl BD. 2007. Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *The journal of biological chemistry* **282**: 7632–40.

Mosley AL, Pattenden SG, Carey M, Venkatesh S, Gilmore JM, Florens L, Workman JL, and Washburn MP. 2009. Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Molecular cell* **34**: 168–78.

Mueller D, Bach C, Zeisig D, García-Cuéllar M-P, Monroe S, Sreekumar A, Zhou R, Nesvizhskii AI, Chinnaiyan A, Hess JL, et al. 2007. A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**: 4445–54.

Mueller D, García-Cuéllar M-P, Bach C, Buhl S, Maethner E, and Slany RK. 2009. Misguided transcriptional elongation causes mixed lineage leukemia. *PLoS biology* **7**: e1000249.

Murton BL, Chin WL, Ponting CP, and Itzhaki LS. 2010. Characterising the binding specificities of the subunits associated with the KMT2/Set1 histone lysine methyltransferase. *Journal of molecular biology* **398**: 481–8.

De Nadal E, Zapater M, Alepuz PM, Sumoy L, Mas G, and Posas F. 2004. The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. *Nature* **427**: 370–4.

Nagy PL, Griesenbeck J, Kornberg RD, and Cleary ML. 2002. A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. *PNAS* **99**: 90–4.

Nakanishi S, Sanderson BW, Delventhal KM, Bradford WD, Staehling-Hampton K, and Shilatifard A. 2008. A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. *Nature structural & molecular biology* **15**: 881–8.

Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, et al. 2006. Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. *Genes & development* **20**: 966–76.

Neil H, Malabat C, D'Aubenton-Carafa Y, Xu Z, Steinmetz LM, and Jacquier A. 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**: 1038–42.

Nelson CJ, Santos-Rosa H, and Kouzarides T. 2006. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126**: 905–16.

Ng HH, Dole S, and Struhl K. 2003b. The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *The journal of biological chemistry* **278**: 33625–8.

Ng HH, Feng Q, Wang H, Erdjument-Bromage H, Tempst P, Zhang Y, and Struhl K. 2002. Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes & development* **16**: 1518–27.

Ng HH, Robert F, Young RA, and Struhl K. 2003a. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular cell* **11**: 709–19.

Nicolas E, Roumillac C, and Trouche D. 2003. Balance between acetylation and methylation of histone H3 lysine 9 on the E2F-responsive dihydrofolate reductase promoter. *Molecular and cellular biology* **23**: 1614–1622.
Niewmierzycka A, and Clarke S. 1999. S-adenosylmethionine-dependent methylation in Saccharomyces cerevisiae. *Biochemistry* **274**: 814–824.
Oh S, Jeong K, Kim H, Kwon CS, and Lee D. 2010. A lysine-rich region in Dot1p is crucial for direct interaction with H2B ubiquitylation and high level methylation of H3K79. *Biochemical and biophysical research communications* **399**: 512–7.
Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, and Zhang Y. 2005. hDOT1L links histone methylation to leukemogenesis. *Cell* **121**: 167–78.
Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P, and Travers AA. 2000. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *The EMBO journal* **19**: 6141–9.
Palancade B, Marshall NF, Tremeau-Bravard A, Bensaude O, Dahmus ME, and Dubois M-F. 2004. Dephosphorylation of RNA polymerase II by CTD-phosphatase FCP1 is inhibited by phospho-CTD associating proteins. *Journal of Molecular Biology* **335**: 415–424.
Parra MA, and Wyrick JJ. 2007. Regulation of gene transcription by the histone H2A N-terminal domain. *Molecular and cellular biology* **27**: 7641–8.
Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, and Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**: 703–17.
Petrossian T, and Clarke S. 2009. Bioinformatic identification of novel methyltransferases. *Epigenomics* **1**: 163–175.
Pijnappel WW, Schaft D, Roguev A, Shevchenko A, Tekotte H, Wilm M, Rigaut G, Séraphin B, Aasland R, and Stewart AF. 2001. The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes & development* **15**: 2991–3004.
Pinskaya M, Gourvennec S, and Morillon A. 2009. H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. *The EMBO journal* **28**: 1697–707.
Pinskaya M, and Morillon A. 2009. Histone H3 lysine 4 di-methylation: a novel mark for transcriptional fidelity? *Epigenetics* **4**: 302–306.
Pokholok DK, Harbison CT, Levine SS, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, et al. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**: 517–27.
Pray-Grant MG, Schieltz D, Mcmahon SJ, Wood JM, Kennedy EL, Cook RG, Workman JL, Yates JRI, and Grant PA. 2002. The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Molecular and cellular biology* **22**: 8774–8786.
Psathas JN, Zheng S, Tan S, and Reese JC. 2009. Set2-dependent K36 methylation is regulated by novel intratail interactions within H3. *Molecular and cellular biology* **29**: 6413–26.
Qiu H, Hu C, and Hinnebusch AG. 2009. Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Molecular cell* **33**: 752–62.
Ramachandran GN, and Mitra AK. 1976. An explanation for the rare occurrence of cis peptide units in proteins and polypeptides. *Journal of molecular biology* **107**: 85–92.
Ramachandran GN, and Sasisekharan V. 1968. Conformation of polypeptides and proteins. *Adv protein chem* **23**: 283–437.
Regenberg B, Grotkjaer T, Winther O, Fausbøll A, Akesson M, Bro C, Hansen LK, Brunak S, and Nielsen J. 2006. Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. *Genome biology* **7**: R107.
Reinke H, and Gatfield D. 2006. Genome-wide oscillation of transcription in yeast. *Trends in biochemical sciences* **31**: 189–91.
Robert F, Pokholok DK, Hannett NM, Rinaldi NJ, Chandy M, Rolfe PA, Workman JL, Gifford DK, and Young RA. 2004. Global position and recruitment of HATs and HDACs in the yeast genome. *Molecular cell* **16**: 199–209.
Robyr D, Suka Y, Xenarios I, Kurdistani SK, Wang A, Suka N, and Grunstein M. 2002. Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. *Cell* **109**: 437–46.
Robzyk K, Recht J, and Osley MA. 2000. Rad6-dependent ubiquitination of histone H2B in yeast. *Science* **287**: 501–504.
Rodriguez CR, Cho E-J, Keogh M-C, Moore CL, Greenleaf AL, and Buratowski S. 2000. Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA Polymerase II. *Molecular and cellular biology* **20**: 104–112.
Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, and Stewart AF. 2001. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *The EMBO journal* **20**: 7137–48.
Rojas JR, Trievel RC, Zhou J, Mo Y, Li X, Berger SL, Allis CD, and Marmorstein R. 1999. Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* **401**: 93–8.
Romagnoli G, Cundari E, Negri R, Crescenzi M, Farina L, Giuliani A, and Bianchi MM. 2011. Synchronous protein cycling in batch cultures of the yeast Saccharomyces cerevisiae at log growth phase. *Experimental cell research* **317**: 2958–68.
Ruiz-García AB, Sendra R, Pamblanco M, and Tordera V. 1997. Gcn5p is involved in the acetylation of histone H3 in nucleosomes. *FEBS letters* **403**: 186–90.
Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, and Grunstein M. 1996. HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *PNAS* **93**: 14503–8.
Sabet N, Tong F, Madigan JP, Volo S, Smith M Mitchell, and Morse RH. 2003. Global and specific transcriptional repression by the histone H3 amino terminus in yeast. *PNAS* **100**: 4084–9.
Sabet N, Volo S, Yu C, Madigan JP, and Morse RH. 2004. Genome-wide analysis of the relationship between transcriptional regulation by Rpd3p and the histone H3 and H4 amino termini in budding yeast genome. *Molecular and cellular biology* **24**: 8823–8833.

Sanli D, Keskin O, Gursoy A, and Erman B. 2011. Structural cooperativity in histone H3 tail modifications. *Protein science* **20**: 1982–1990.
Santos-Rosa H, Schneider R, Bannister Andrew J, Sherriff J, Bernstein BE, Emre NCT, Schreiber SL, Mellor J, and Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419**: 407–411.
Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, Schreiber SL, Mellor J, and Kouzarides T. 2003. Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Molecular cell* **12**: 1325–32.
Sawada K, Yang Z, Horton JR, Collins RE, Zhang X, and Cheng X. 2004. Structure of the conserved core of the yeast Dot1p, a nucleosomal histone H3 lysine 79 methyltransferase. *The journal of biological chemistry* **279**: 43296–306.
Schaft D. 2003. The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Research* **31**: 2475–2482.
Schlichter A, and Cairns BR. 2005. Histone trimethylation by Set1 is coordinated by the RRM, autoinhibitory, and catalytic domains. *The EMBO journal* **24**: 1222–31.
Schmid FX. 1995. Prolyl isomerases join the fold. *Current Biology* **5**: 993–994.
Schmitt AP, and McEntee K. 1996. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in Saccharomyces cerevisiae. *PNAS* **93**: 5777–82.
Schneider J, Dover J, Johnston M, and Shilatifard A. 2004. Global proteomic analysis of S. cerevisiae (GPS) to identify proteins required for histone modifications. *Methods enzymology* **377**: 227–234.
Schneider J, Wood A, Lee J-S, Schuster R, Dueker J, Maguire C, Swanson SK, Florens L, Washburn MP, and Shilatifard A. 2005. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Molecular cell* **19**: 849–56.
Schroeder SC. 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes & development* **14**: 2435–2440.
Schulze JM, Hentrich T, Nakanishi S, Gupta A, Emberly E, Shilatifard A, and Kobor MS. 2011. Splitting the task: Ubp8 and Ubp10 deubiquitinate different cellular pools of H2BK123. *Genes & development* **25**: 2242–2247.
Sekito T, Thornton J, and Butow RA. 2000. Mechanism of metabolic control. Target of rapamycin signaling links nitrogen quality to the activity of the Rtg1 and Rtg3 transcription factors. *Molecular biology of the cell* **11**: 2103–15.
Seward DJ, Cubberley G, Kim S, Schonewald M, Zhang L, Tripet B, and Bentley DL. 2007. Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nature structural & molecular biology* **14**: 240–2.
Shahbazian MD, Zhang K, and Grunstein M. 2005. Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. *Molecular cell* **19**: 271–7.
Sheldon KE, Mauger DM, and Arndt KM. 2005. A requirement for the Saccharomyces cerevisiae Paf1 complex in snoRNA 3' end formation. *Molecular cell* **20**: 225–36.
Shi L, Sutter BM, Ye X, and Tu BP. 2010. Trehalose is a key determinant of the quiescent metabolic state that fuels cell cycle progression upon return to growth. *Molecular biology of the cell* **21**: 1982–1990.
Shi X, and Gozani O. 2005. The Fellowships of the INGs. *Journal of Cellular Biochemistry* **1136**: 1127–1136.
Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Peña P, Lan F, Kaadige MR, et al. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–9.
Shi X, Kachirskaia I, Walter KL, Kuo J-HA, Lake A, Davrazou F, Chan SM, Martin DGE, Fingerman IM, Briggs S, et al. 2007. Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *The journal of biological chemistry* **282**: 2450–5.
Shilatifard A. 2006. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annual review of biochemistry* **75**: 243–69.
Shilatifard A. 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current opinion in cell biology* **20**: 341–8.
Shukla A, Stanojevic N, Duan Z, Sen P, and Bhaumik SR. 2006. Ubp8p , a histone deubiquitinase whose association with SAGA is mediated by Sgf11p , differentially regulates lysine 4 methylation of histone H3 in vivo. *Molecular and cellular biology* **26**: 3339–3352.
Silverman SJ, Petti AA, Slavov N, Parsons L, Briehof R, Thiberge SY, Zenklusen D, Gandhi SJ, Larson DR, Singer RH, et al. 2010. Metabolic cycling in single yeast cells from unsynchronized steady-state populations limited on glucose or phosphate. *PNAS* **107**: 6946–51.
Sims RJ, Chen C-F, Santos-Rosa H, Kouzarides Tony, Patel SS, and Reinberg D. 2005. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *The journal of biological chemistry* **280**: 41789–92.
Singer MS, Kahana A, Wolf AJ, Meisinger LL, Peterson SE, Goggin C, Mahowald M, and Gottschling DE. 1998. Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae. *Genetics* **150**: 613–32.
Singh N, Ma Z, Gemmill T, Wu X, Defiglio H, Rossettini A, Rabeler C, Beane O, Morse RH, Palumbo MJ, et al. 2009. The Ess1 prolyl isomerase is required for transcription termination of small noncoding RNAs via the Nrd1 pathway. *Molecular cell* **36**: 255–66.
Slavov N, Airoldi EM, Van Oudenaarden A, and Botstein D. 2012. A conserved cell growth cycle can account for the environmental stress responses of divergent eukaryotes. *Molecular biology of the cell* **23**: 1986–97.
Slavov N, and Botstein D. 2011. Coupling among growth rate response, metabolic cycle, and cell division cycle in yeast. *Molecular biology of the cell* **22**: 1997–2009.
Slavov N, Macinskas J, Caudy A, and Botstein D. 2011. Metabolic cycling without cell division cycling in respiring yeast. *PNAS* **108**: 19090–19095.
Smith E, Eisen A, Gu W, Sattah M, Pannuti A, Zhou J, Cook RG, Lucchesi JC, and Allis CD. 1998. ESA1 is a histone acetyltransferase that is essential for growth in yeast. *PNAS* **95**: 3561–5.

South PF, Fingerman IM, Mersman DP, Du H-N, and Briggs S. 2010. A conserved interaction between the SDI domain of Bre2 and the Dpy-30 domain of Sdc1 is required for histone methylation and gene expression. *The journal of biological chemistry* **285**: 595–607.

Spellman PT, Sherlock G, Zhang MQ, Vishwanath R, Anders K, Eisen MB, Brown PO, and Futcher B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridisation. *Molecular biology of the cell* **9**: 3273–3297.

Stassen MJ, Bailey D, Nelson S, Chinwalla V, and Harte PJ. 1995. The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mechanisms of Development* **52**: 209–223.

Steinmetz EJ, Conrad NK, Brow DA, and Corden JL. 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* **413**: 327–31.

Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, and Brow DA. 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Molecular cell* **24**: 735–46.

Sterner DE, Belotserkovskaya R, and Berger SL. 2002. SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *PNAS* **99**: 11622–7.

Sterner DE, Grant PA, Roberts SM, Duggan LJ, Belotserkovskaya R, Pacella LA, Winston F, Workman JL, and Berger SL. 1999. Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-Binding Protein interaction. *Molecular and cellular biology* **19**: 86–98.

Steward MM, Lee J-S, O'Donovan A, Wyatt M, Bernstein BE, and Shilatifard A. 2006. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature structural & molecular biology* **13**: 852–4.

Strahl BD, and Allis CD. 2000. The language of covalent histone modifications. *Nature* **403**: 41–45.

Strahl BD, Grant PA, Briggs S, Sun Z-W, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, et al. 2002. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Molecular and cellular biology* **22**: 1298–1306.

Sun Z-W, and Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**: 104–8.

Takahashi Y-H, Lee J-S, Swanson SK, Saraf A, Florens L, Washburn MP, Trievel RC, and Shilatifard A. 2009. Regulation of H3K4 trimethylation via Cps40 (Spp1) of COMPASS is monoubiquitination independent: implication for a Phe/Tyr switch by the catalytic domain of Set1. *Molecular and cellular biology* **29**: 3478–86.

Takahashi Y-H, Westfield GH, Oleskie AN, Trievel RC, Shilatifard A, and Skiniotis G. 2011. Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *PNAS* **108**: 20526–20531.

Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li Haitao, Baker L, Boyle J, Blair LP, Chait BT, et al. 2006. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Molecular cell* **24**: 785–96.

Taverna SD, Li Haitao, Ruthenburg AJ, Allis CD, and Patel DJ. 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology* **14**: 1025–40.

Taylor CM, Hardre R, and Edwards PJB. 2005. The impact of pyrrolidine hydroxylation on the conformation of proline-containing peptides. *Journal of organic chemistry* 1306–1315.

Terzi N, Churchman LS, Vasiljeva L, Weissman JS, and Buratowski S. 2011. H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Molecular and cellular biology* **31**: 3569–83.

Thiebaut M, Colin J, Neil H, Jacquier A, Séraphin B, Lacroute F, and Libri D. 2008. Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in S . cerevisiae. *Molecular cell* **31**: 671–682.

Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, and Libri D. 2006. Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the Nrd1-Nab3 pathway in genome surveillance. *Molecular cell* **23**: 853–64.

Tompa R, and Madhani HD. 2007. Histone H3 lysine 36 methylation antagonizes silencing in Saccharomyces cerevisiae independently of the Rpd3S histone deacetylase complex. *Genetics* **175**: 585–93.

Trésaugues L, Dehé P-M, Guérois R, Rodriguez-Gil A, Varlet I, Salah P, Pamblanco M, Luciano P, Quevillon-Cheruel S, Sollier J, et al. 2006. Structural characterization of Set1 RNA recognition motifs and their role in histone H3 lysine 4 methylation. *Journal of molecular biology* **359**: 1170–81.

Tse C, Sera T, Wolffe AP, and Hansen JC. 1998. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Molecular and cellular biology* **18**: 4629–4638.

Tu BP, Kudlicki A, Rowicka M, and McKnight SL. 2005. Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. *Science* **310**: 1152–8.

Tu BP, Mohler RE, Liu JC, Dombek KM, Young ET, Synovec RE, and McKnight SL. 2007. Cyclic changes in metabolic state during the life of a yeast cell. *PNAS* **104**: 16886–91.

Tu S, Bulloch EMM, Yang L, Ren C, Huang W-C, Hsu P-H, Chen C-H, Liao C-L, Yu H-M, Lo W-S, et al. 2007. Identification of histone demethylases in Saccharomyces cerevisiae. *The journal of biological chemistry* **282**: 14262–71.

Turner BM. 2000. Histone acetylation and an epigenetic code. *BioEssays* **22**: 836–45.

Utley RT, Ikeda K, Grant PA, Côté J, Steger D, Eberharter A, John S, and Workman JL. 1998. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**: 1–5.

Vanácová S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, and Keller W. 2005. A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS biology* **3**: e189.

Vasiljeva L, and Buratowski S. 2006. Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Molecular cell* **21**: 239–48.

Vasiljeva L, Kim M, Mutschler H, Buratowski S, and Meinhart A. 2008. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature structural & molecular biology* **15**: 795–804.

Venkatasubrahmanyam S, Hwang WW, Meneghini MD, Tong A, and Madhani HD. 2007. Genome-wide, as opposed to local, antisilencing is mediated redundantly by the euchromatic factors Set1 and H2A.Z. *PNAS* **104**: 16609–14.
Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamon AJ, Jain P, Rolleri NS, Jiang C, Hemeryck-Walsh C, et al. 2011. A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. *Molecular cell* **41**: 480–92.
Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen J V, Hyman AA, Stunnenberg HG, et al. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**: 967–80.
Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, Van Schaik FMA, Varier RA, Baltissen MPA, Stunnenberg HG, Mann M, and Timmers HTM. 2007. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**: 58–69.
Vettese-Dadey M, Grant PA, Hebbes TR, Crane- Robinson C, Allis CD, and Workman JL. 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *The EMBO journal* **15**: 2508–18.
Vignali M, Steger D, Neely KE, and Workman JL. 2000. Distribution of acetylated histones resulting from Gal4-VP16 recruitment of SAGA and NuA4 complexes. *The EMBO journal* **19**: 2629–40.
Vitaliano-Prunier A, Menant A, Hobeika M, Géli V, Gwizdek C, and Dargemont C. 2008. Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. *Nature cell biology* **10**: 1365–71.
De Vos D, Frederiks F, Terweij M, Van Welsem T, Verzijlbergen KF, Iachina E, De Graaf EL, Altelaar AFM, Oudgenoeg G, Heck AJR, et al. 2011. Progressive methylation of ageing histones by Dot1 functions as a timer. *EMBO reports* **12**: 956–62.
Walter W, Clynes D, Tang Y, Marmorstein R, Mellor J, and Berger SL. 2008. 14-3-3 interaction with histone H3 involves a dual modification pattern of phosphoacetylation. *Molecular and cellular biology* **28**: 2840–9.
Wang A, Kurdistani SK, and Grunstein M. 2002. Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science* **298**: 1412–4.
Wang GG, Cai L, Pasillas MP, and Kamps MP. 2007. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nature cell biology* **9**: 804–12.
Wang S-S, Zhou BO, and Zhou J-Q. 2011. Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. *Molecular and cellular biology* **31**: 3171–81.
Wang X, Moore SC, Laszckzak M, and Ausió J. 2000. Acetylation increases the alpha-helical content of the histone tails of the nucleosome. *The journal of biological chemistry* **275**: 35013–20.
Wang Xiaodong, and Hayes JJ. 2008. Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure. *Molecular and cellular biology* **28**: 227–36.
Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ, et al. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics* **40**: 897–903.
Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences* **24**: 437–40.
Webb KJ, Laganowsky A, Whitelegge JP, and Clarke S. 2008. Identification of two SET domain proteins required for methylation of lysine residues in yeast ribosomal protein Rpl42ab. *Journal of Biological Chemistry* **283**: 35561–35568.
Webby CJ, Wolf A, Gromak N, Dreger M, Kramer H, Kessler B, Nielsen ML, Schmitz C, Butler DS, Yates JR, et al. 2009. Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science (New York, N.Y.)* **325**: 90–3.
Weiner A, Chen H V, Liu CL, Rahat A, Klien A, Soares L, Gudipati M, Pfiffner J, Regev A, Buratowski S, et al. 2012. Systematic dissection of roles for chromatin regulators in a yeast stress response ed. Jerry Workman. *PLoS Biology* **10**: e1001369.
Weiss MS, Jabs A, and Hilgenfeld R. 1998. Peptide bonds revisited. *Nature structural biology* **5**: 676.
Werner-Allen JW, Lee C-J, Liu P, Nicely NI, Wang S, Greenleaf AL, and Zhou P. 2011. cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *The journal of biological chemistry* **286**: 5717–26.
White CL, Suto RK, and Luger K. 2001. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *The EMBO journal* **20**: 5207–18.
Widlund HR, Vitolo JM, Thiriet C, and Hayes JJ. 2000. DNA sequence-dependent contributions of core histone tails to nucleosome stability: differential effects of acetylation and proteolytic tail removal. *Biochemistry* **39**: 3835–41.
Winkler GS, Kristjuhan A, Erdjument-Bromage H, Tempst P, and Svejstrup JQ. 2002. Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels in vivo. *PNAS* **99**: 3517–22.
Winter S, Simboeck E, Fischle W, Zupkovitz G, Dohnal I, Mechtler K, Ammerer G, and Seiser C. 2008. 14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation. *The EMBO journal* **27**: 88–99.
Wittschieben BØ, Otero G, Bizemont T De, Fellows J, Erdjument-Bromage H, Ohba Reiko, Li Yang, Allis CD, Tempst P, Svejstrup JQ, et al. 1999. A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. *Cell* **4**: 123–128.
Wlodarski T, Kutner J, Towpik J, Knizewski L, Rychlewski L, Kudlicki A, Rowicka M, Dziembowski A, and Ginalski K. 2011. Comprehensive structural and substrate specificity classification of the Saccharomyces cerevisiae methyltransferome. *PloS one* **6**: e23168.
Wood A, Krogan NJ, Dover J, Schneider J, Heidt J, Boateng MA, Dean K, Golshani A, Zhang Y, Greenblatt JF, et al. 2003a. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Molecular cell* **11**: 267–274.
Wood A, Schneider J, Dover J, Johnston M, and Shilatifard A. 2005. The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. *Molecular cell* **20**: 589–99.

Wood A, Schneider J, Dover J, Johnston M, and Shilatifard A. 2003b. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *The journal of biological chemistry* **278**: 34739–42.

Wood A, Shukla A, Schneider J, Lee J-S, Stanton JD, Dzuiba T, Swanson SK, Florens L, Washburn MP, Wyrick JJ, et al. 2007. Ctk complex-mediated regulation of histone methylation by COMPASS. *Molecular and cellular biology* **27**: 709–20.

Wu M, Wang PF, Lee J-S, Martin-Brown S, Florens L, Washburn MP, and Shilatifard A. 2008. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Molecular and cellular biology* **28**: 7337–44.

Wu P-YJ, and Winston F. 2002. Analysis of Spt7 function in the Saccharomyces cerevisiae SAGA coactivator complex. *Molecular and cellular biology* **22**: 5367–5379.

Wu X, Wilcox CB, Devasahayam G, Hackett RL, Arévalo-Rodríguez M, Cardenas ME, Heitman J, and Hanes SD. 2000. The Ess1 prolyl isomerase is linked to chromatin remodeling complexes and the general transcription machinery. *The EMBO journal* **19**: 3727–38.

Wyart M, Botstein D, and Wingreen NS. 2010. Evaluating gene expression dynamics using pairwise RNA FISH data. *PLoS computational biology* **6**: e1000979.

Wyce A, Xiao T, Whelan KA, Kosman C, Walter W, Eick D, Hughes TR, Krogan NJ, Strahl BD, and Berger SL. 2007. H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Molecular cell* **27**: 275–88.

Wyers F, Rougemaille M, Badis G, Rousselle J-C, Dufour M-E, Boulay J, Régnault B, Devaux F, Namane A, Séraphin B, et al. 2005. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**: 725–737.

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, et al. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86–90.

Xiang K, Nagaike T, Xiang S, Kilic T, Beh MM, Manley JL, and Tong L. 2010. Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. *Nature* **467**: 729–33.

Xiao B, Jing C, Wilson JR, Walker PA, Vasisht N, Kelly G, Howell S, Taylor IA, Blackburn GM, and Gamblin SJ. 2003. Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* **421**: 652–6.

Xiao T, Hall H, Kizer KO, Shibata Y, Hall MC, Borchers CH, and Strahl BD. 2003. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes & development* **17**: 654–63.

Xiao T, Kao C-F, Krogan NJ, Sun Z-W, Greenblatt JF, Osley MA, and Strahl BD. 2005. Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Molecular and cellular biology* **25**: 637–651.

Xu Y-X, Hirose Y, Zhou XZ, Lu KP, and Manley JL. 2003. Pin1 modulates the structure and function of human RNA polymerase II. *Genes & development* **17**: 2765–76.

Xu Z, Wei W, Gagneur J, Clauder-Münster S, Smolik M, Huber W, and Steinmetz LM. 2011. Antisense expression increases gene expression variability and locus interdependency. *Molecular systems biology* **7**: 1–10.

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Guffanti E, Stutz F, Huber W, and Steinmetz LM. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**: 1033–1037.

Yassour M, Pfiffner J, Levin JZ, Adiconis X, Gnirke A, Nusbaum C, Thompson D-A, Friedman N, and Regev A. 2010. Strand-specific RNA sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species. *Genome biology* **11**: R87.

Yokoyama A, Wang Zhong, Wysocka J, Sanyal M, Aufiero DJ, Kitabayashi I, Herr W, and Cleary ML. 2004. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Molecular and cellular biology* **24**: 5639–5649.

Youdell ML, Kizer KO, Kisseleva-Romanova E, Fuchs SM, Duro E, Strahl BD, and Mellor J. 2008. Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Molecular and cellular biology* **28**: 4915–26.

Yu BD, Hess JL, Horning SE, Brown GA, and Korsmeyer SJ. 1995. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**: 505–508.

Zee BM, Levin RS, Xu B, LeRoy G, Wingreen NS, and Garcia BA. 2010. In vivo residue-specific histone methylation dynamics. *The journal of biological chemistry* **285**: 3341–50.

Zhang DW, Mosley AL, Ramisetty SR, Rodríguez-Molina JB, Washburn MP, and Ansari AZ. 2012. Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA polymerase II C-terminal domain is essential for viability and transcription termination. *The journal of biological chemistry* **287**: 8541–51.

Zhang K, Lin W, Latham JA, Riefler GM, Schumacher JM, Chan C, Tatchell K, Hawke DH, Kobayashi Ryuji, and Dent SYR. 2005. The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. *Cell* **122**: 723–34.

Zhang W, Bone JR, Edmondson DG, Turner BM, and Roth SY. 1998. Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *The EMBO journal* **17**: 3155–67.

Zhang Wenzheng, Xia Xuefeng, Reisenauer MR, Hemenway CS, and Kone BC. 2006. Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCalpha in an aldosterone-sensitive manner. *The journal of biological chemistry* **281**: 18059–68.

Zhang X, Yang Z, Khan SI, Horton JR, Tamaru H, and Selker EU. 2003. Structural basis for the product specificity of histone lysine methyltransferases. *Cell* **12**: 177–185.

Zheng S, Wyrick JJ, and Reese JC. 2010. Novel trans-tail regulation of H2B ubiquitylation and H3K4 methylation by the N terminus of histone H2A. *Molecular and cellular biology* **30**: 3635–45.

Zhou BO, and Zhou J-Q. 2011. Recent transcription-induced histone H3K4 methylation inhibits gene reactivation. *The journal of biological chemistry* **286**: 34770–34776.