Function and regulation of the histone methyltransferase Dot1
Frederiks, F.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 1

Introduction and general discussion

Floor Frederiks¹ and Fred van Leeuwen¹

¹ Division of Gene Regulation, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam

PREFACE

The subject of this thesis is the function and regulation of the histone methyltransferase Dot1. Dot1 is an unusual enzyme that modifies an unusual location on one of the histone proteins. As a consequence, many of the methods commonly used to study histone methylation cannot be applied to this enzyme. This chapter will start with an overview of current knowledge, mostly focussing on the function of Dot1 in the budding yeast Saccharomyces cerevisiae. In the second part of the chapter, important open questions will be discussed with a focus on the different genetic and (bio)chemical methods that are currently used to study Dot1 regulation. These different approaches are largely complementary and all help to solve a different part of the puzzle.
PART I: Introduction into the role of Dot1

Chromatin and histone modifications

Relative to its size, each nucleus of a eukaryotic cell contains a very large amount of DNA. This simple fact means that the DNA has to be strongly compacted to fit in the available space. To ensure proper transcription and replication of the genetic information, an intricate DNA packaging system has evolved. The first layer of packaging is the winding of 147 base pairs of the DNA double helix around an octamer of histone proteins. This complex of histones and DNA is called a nucleosome. There are four canonical histones: H2A, H2B, H3 and H4, and two copies of each are present in one nucleosome. The next level of compaction is the stacking of a chain of nucleosomes, with the help of the ‘linker histone’ H1, to form the so-called 30 nm fiber. Beyond this level, the packaging of DNA is very poorly understood.

The histone-DNA complex interacts with a very large number of associating proteins. The resulting protein-DNA complex is called chromatin. Traditionally, a distinction has been made between two types of chromatin called euchromatin and heterochromatin. Euchromatin is thought to be ‘open’ chromatin that is well accessible to proteins and therefore transcriptionally active, whereas heterochromatin is more tightly packed and therefore transcriptionally silent. Changes in the chromatin context and packing of a gene can have a major effect on its activity. This means that two genetically identical cells can have very different gene expression patterns due to differences in their chromatin composition. Posttranslational modifications of histone proteins are a very important factor in this differential regulation of gene expression. These histone modifications affect the structure and thereby the function of chromatin in many different ways.

To date, more than fifty different histone modifications have been identified. Among others, histones can be methylated, acetylated, phosphorylated and ubiquitinated on many different residues. Histone methylation in particular is a very complex modification, since one amino acid can accommodate two (for an arginine) or even three (for a lysine) methyl groups. The general view is that these different methylation states of a single arginine or lysine can have different functions, often through the binding of proteins specific for one of the methylation states. Histone modifications not only play a role in transcription, but essentially in all processes that take place on DNA, such as replication, DNA repair and signaling.
Dot1: Disruptor of Telomeric Silencing

Here we focus on methylation of histone H3 on lysine 79 (abbreviated as H3K79) by the histone methyltransferase Dot1. Dot1, which stands for Disruptor of telomeric silencing -1, was discovered in *Saccharomyces cerevisiae* in a screen for proteins that disrupt silencing of telomeric genes when overexpressed (3). The principle of the assay for telomeric silencing, which is used in several of the chapters of this thesis, is explained in Fig. 1. In contrast to that of mammals, the genome of yeast almost entirely consists of active chromatin. Three regions are known to be silenced: the telomeres, the silent mating type loci HML and HMR, and the rDNA locus. Telomeres are silenced by a complex of silencing proteins called Sir2, Sir3, and Sir4. Sir2 is a histone deacetylase and removes acetyl groups from histones. The deacetylated nucleosomes are then bound by Sir3 and Sir4, which renders the chromatin less accessible to other proteins. Even though proteins involved in transcription initiation are still able to bind to silent chromatin, access of the machinery responsible for transcription elongation is blocked (4;5). In this way transcription is inhibited. Sir3/4 in turn can recruit Sir2 which then deacetylates adjacent histones, a process that leads to the ‘spreading’ of silent chromatin (6;7).

When Dot1 was first discovered as a protein whose overexpression can disrupt telomeric silencing, its biological activity was unclear. A few years later several groups found that the protein harbored a methyltransferase activity that was specific for H3K79 (8-11). In two aspects, Dot1 is different from all other

---

**Figure 1. An assay for telomeric silencing.** A reporter gene is inserted near a telomere in the yeast genome; telomeres V-R and VII-L are often used. Since the reporter gene is placed very close to the telomere, it will be silenced by Sir-protein mediated mechanisms in wild type cells. When the *ADE2* reporter gene is silenced, yeast cells will accumulate a red pigment and therefore have a red color. In contrast, when the *ADE2* reporter is activated by loss of telomeric silencing, the cells are white. A second reporter gene, *URA3*, confers sensitivity to the drug 5-fluoro-orotic acid (5-FOA) when it is expressed. Consequently, cells with intact telomeric silencing of *URA3* can grow on media containing 5-FOA, whereas cells expressing *URA3* cannot survive on this media. In this way, the color of the cells and/or their growth on FOA-containing media is a convenient measure for the strength of telomeric silencing.
methyltransferases that can methylate lysine residues on histones. First, unlike these other enzymes, Dot1 does not contain a so-called SET catalytic domain. Instead, it possesses a completely different class-I methyltransferase domain, which is also found in arginine methyltransferases (12;13). Second, the residue that is methylated by Dot1, H3K79, is not located on the tail of histone H3 like most other methylated lysines, but instead on the surface of the core of the nucleosome (see below).

About 90% of yeast histone H3 is methylated on K79 and the modification is found in active chromatin (8;14). Binding of Sir3 to nucleosomes is negatively influenced by H3K79 methylation (see below). The current model is that methylation of H3K79 in active chromatin by Dot1 prevents Sir proteins from binding to these parts of the genome (Fig. 2 and see below). Consequently, the Sir proteins are confined to the silent chromatin regions, where they in turn prevent methylation of H3K79 by Dot1 (8;14). In this way the competition between Sir protein binding and methylation by Dot1 results in a dynamic balance between active and silent chromatin. This model explains the phenotype of Dot1 overexpression that led to its discovery: when Dot1 is overexpressed, the normal balance between Dot1 and Sir proteins is lost and H3K79 methylation also occurs in normally silenced regions of the genome, such as the telomeres. The result of this imbalance is therefore the loss of telomeric silencing. Interestingly, loss of Dot1 results in a similar loss-of-silencing phenotype. When Dot1 is absent, the Sir proteins are no longer prevented from binding to active chromatin. Since the amount of Sir proteins in the nucleus is limiting, their binding to active chromatin regions leads to a loss of binding at normally silenced regions, and therefore to a loss of silencing (8).

**Other functions of Dot1 in budding yeast**

Besides its role in telomeric silencing, several other functions for Dot1 have been described in budding yeast. The enzyme is involved in the pachytene checkpoint, which prevents the progression of meiosis when the chromosomes have not yet been properly aligned. In the absence of Dot1, two known checkpoint mutants inappropriately progress through meiosis and sporulation, which results in the production of inviable spores (15). This suggests that Dot1 has a backup function in the pachytene checkpoint, promoting cell cycle arrest in cells that have a partial checkpoint defect.
Introduction

In addition to this meiotic checkpoint function, Dot1 also has a role in the DNA damage checkpoint after damage caused by UV- or ionizing radiation or the alkylating agent methyl methanesulfonate (MMS). Two groups reported that in cells lacking Dot1, the central checkpoint kinase Rad53 is not properly activated upon UV- or IR-induced damage (16;17). After damage induction yeast cells normally arrest to allow for repair of the lesions. In dot1Δ cells this cell cycle arrest is impaired and cells re-enter the cell cycle faster than wild type cells do (16;17). However, dot1Δ cells are hardly, if at all, more sensitive to UV light or ionizing radiation than wild type cells, as measured by their survival at different doses of irradiation (17-19). Interestingly, loss of DOT1 results in increased MMS resistance and suppresses the MMS sensitivity of various known DNA repair mutants (20), indicating that the role of Dot1 in the different damage response pathways is complex. The DNA damage checkpoint role of Dot1 is the subject of Chapter 7 of this thesis.

Finally, Dot1 also affects cell wall biogenesis, a function that is not very well characterized. It has been reported that deletion of DOT1 is lethal when combined with deletion of CHS5 or GAS1, two genes involved in cell wall assembly.
This synthetic lethality suggests that Dot1 is also involved in some aspect of cell wall metabolism. We obtained more support for this idea through electron microscopy experiments showing that overexpression of Dot1 alters the outer layer of the cell wall. In addition, the N-terminus of Dot1 seems to mediate sensitivity to Calcofluor White, which is thought to be indicative of a defect in cell wall function. The cell wall function of Dot1 is discussed in more detail in Chapter 6.

Dot1 enzymes in other organisms

Quickly after Dot1 was discovered in yeast, a number of Dot1 homologs were identified in other organisms. The protein is conserved from budding yeast to mammals, with proteins known to be active towards H3K79 in the unicellular parasite *Trypanosoma brucei* (23), the fruit fly *Drosophila melanogaster* (24), mouse (25), and humans (10).

Interestingly, *T. brucei* expresses two distinct Dot1 proteins, DOT1A and DOT1B, which seem to have different catalytic activities and together play a role in cell cycle regulation (23). Chapter 5 of this thesis focuses on the catalytic activities of DOT1A and DOT1B. In addition to *T. brucei*, the nematode worm *Caenorhabditis elegans* also harbors multiple Dot1-like genes, but their biological activity and function are unknown (15). Yeast, flies, mice and humans contain only a single Dot1 protein. Attempts to generate mice deficient for the murine homologue Dot1L showed that the protein is essential: targeted deletion results in embryonic lethality (26). Mouse ES cells deficient for Dot1L show an altered heterochromatin structure and the enzyme is required for the early stages of ES cell differentiation, most likely through the regulation of a number of known cell cycle genes (26;27). The *Drosophila* Dot1 homologue, named *grappa*, is also essential (24).

The human Dot1 homolog has been named hDot1L (10). Efficient knockdown of hDot1L in cultured cells leads to cell death, suggesting that the enzyme is also essential in humans (28). Methylation of H3K79 by hDot1L has been implicated in different types of leukemia. In these leukemias, the well-known mixed lineage leukemia gene MLL is translocated and its DNA-binding domain is fused to one of a number of different translocation partners. Several of these MLL fusion partners, such as AF4, AF9, and AF10, have been shown to bind to hDot1L. Through this interaction hDot1L is recruited to MLL target genes and novel targets, which become hypermethylated on H3K79, resulting in an altered expression pattern (29). These changes in gene expression, for example at the Hox gene cluster, then drive leukemic transformation of the cells in which the
translocation has taken place. This role of hDot1L in leukemia development has led to the suggestion that the protein might be a target for therapeutic intervention, for example by the development of a specific inhibitor (30). In addition, the fact that Dot1 has a very different active site from all other lysine methyltransferases means it might be possible to target the enzyme quite specifically (31).

**Putative H3K79 binding proteins**

The functions of many histone modifications are mediated via proteins that bind to the modified residue, in the case of histone methylation often specifically to one of the methylation states. Important chromatin binding domains are the bromodomain, which mediates binding to acetylated lysines, and the chromodomain, the PhD finger and the Tudor domain, which all bind to methylated lysines (32). For methylated H3K79, no binding proteins have been identified to date. It has been suggested that in yeast the Rad9 protein, a mediator of the DNA damage checkpoint, might bind to H3K79 after damage has occurred (33). Rad9 contains a Tudor domain and it has been shown that loss of Dot1 impairs the binding of Rad9 to a double-strand break (17). However, a direct interaction between Rad9 and (methylated) H3K79 has not been demonstrated. In mammalian cells, the checkpoint protein 53BP1, a Rad9 homolog, was proposed to bind methylated H3K79 after DNA damage (34), but this finding was later challenged (35). A systematic study with a large set of proteins containing known chromatin-binding domains showed that none of these showed affinity for methylated H3K79 (36). Therefore, it is not yet clear if the signal of H3K79 methylation is interpreted through binding proteins or perhaps directly through the changes in chromatin structure that are the result of the methylation.

**Part II: Understanding Dot1 regulation**

In order to get a better understanding of the regulation of Dot1, genome-wide chromatin immunoprecipitation (ChIP) experiments have been performed in yeast, flies, and mammals. In most of these studies the distribution of H3K79me1, -me2 and/or -me3 was determined, but the localization of Dot1 itself has also been addressed in yeast and in human cells. The first genome-wide study in yeast compared the levels of H3K79me3 with the transcriptional frequency of the genes and found that all genes are enriched for H3K79me3 compared to intergenic regions. However, there was no correlation between the amount of H3K79me3
and the transcriptional frequency of the gene (37). The same result was later found for H3K79me1 and -me2 on the inducible GAL genes: these genes were methylated regardless of whether they were active or repressed. In contrast, Dot1 occupancy on these genes was only detectable when they were expressed (38). A recent genome-wide study again confirmed that there is no clear correlation between the transcriptional activity of genes and the enrichment of H3K79me2 and -me3 (39). There was one exception: very highly expressed genes were found to have very low levels of H3K79me2 and -me3 in both their promoters and their open reading frames. In addition, regions enriched for H3K79me2 or H3K79me3 showed very little overlap (39).

Interestingly, in flies and humans, there is a very clear correlation between transcriptional frequency and enrichment of H3K79 methylation. In Drosophila, H3K79me2 is restricted to actively transcribed genes (40). In human cells, where hDot1L associates with active genes, H3K79me2 and -me3 are also strongly correlated with gene activity (41). This effect is much less clear for H3K79me1, which is also found in intergenic regions (41). Furthermore, in human and mouse leukemias which are caused by a rearrangement of the MLL gene, H3K79me2 is strongly enhanced on many genes including the HOX gene cluster, which results in a strong increase in expression of these genes (30). Therefore, in yeast and metazoans, H3K79 methylation seems to act as a co-transcriptional mark that is introduced efficiently in (coding) regions that are transcribed by RNA polymerase II. The quantitative differences in the relationship between transcription and H3K79 methylation between yeast and metazoans will be discussed below.

ChIP experiments provide important information about the distribution of H3K79 methylation and Dot1 itself throughout the genome, and thereby often give hints as to what the function of the methylation might be. However, they provide a static picture and do not give mechanistic or molecular insights into the regulation of Dot1. A prerequisite for the use of Dot1 as a therapeutic target to treat leukemia is a thorough understanding of how the protein functions catalytically and how its activity is regulated by other factors in the cell. These questions are the subject of the major part of this thesis and the rest of this chapter will focus on the different methods that have been used and are being used to study them.

**Regulation of Dot1 activity by ubiquitination of histone H2B**

Most of what we know about the functions and regulation of Dot1 was derived from genetic experiments. Yeast is a popular model organism for genetic chromatin research, since homologs of many mammalian chromatin proteins are
Introduction

present in yeast and many mechanisms are conserved. It is relatively easy to knock out, knock in, or tag a specific gene in the yeast genome. In addition, a complete collection of all viable yeast single-gene knockout strains is available (42). Here, the power of yeast genetics will be illustrated by the way in which the crosstalk between ubiquitination of histone H2B and H3K79 methylation was discovered.

Ubiquitination is an abundant protein modification that is often linked to quality control and proteasomal degradation. However, poly- and monoubiquitination can also be used as a signal. In yeast and mammals histone H2B is monoubiquitinated on the C-terminal tail (K123 in yeast and K120 in mammals). In higher eukaryotes ubiquitination of histone H2A is also found. Ubiquitination of H2BK123 was one of the first histone modifications to be identified (43), but its biological function was poorly understood. This changed when the first "trans-histone" pathway was identified: two groups independently found that H2B ubiquitination is required for wild type levels of methylation of H3K4 (44;45). The first study was hypothesis-driven, specifically deleting genes that might have an effect on methylation of H3K4; these genes were selected based on their silencing phenotype, which they shared with the H3K4 methyltransferase Set1 (44). The second study used a genome-wide approach, using the yeast knockout collection to screen systematically for knockout strains that showed a loss of H3K4 methylation on immunoblot (45).

Very soon afterwards, it was shown that H2B ubiquitination is also required for H3K79 methylation (46;47). Here again, the approach taken was hypothesis-driven, based on the fact that the two modifications are in close proximity in the 3D structure of the nucleosome (Fig. 3). Two possibilities were addressed: H2B ubiquitination might affect the levels of H3K79 methylation, or vice versa. Using a yeast strain expressing a tagged version of H2B, it was shown that the levels of ubiquitinated H2B were not affected by deletion of Dot1. However, a strain in which Rad6, the ubiquitin conjugating enzyme for H2BK123, was deleted did show a dramatic reduction in H3K79 methylation (47). Crucial in all of these early studies were antibodies specific for methylated H3K4 and H3K79, which allowed direct visualization of the effect of various deletions on methylation levels. Soon, similar genetic experiments identified the ubiquitin ligase Bre1 as the enzyme responsible for ubiquitination of H2B and thereby also required for H3K79 methylation (48). Similar pathways of histone crosstalk involving homologs of Bre1, Rad6, and Dot1 have been identified in humans and in flies (49;50).
Regulation of Dot1 activity by other factors

Apart from H2B ubiquitination, several other mechanisms of Dot1 regulation have been described in yeast. The first is again a trans-histone regulatory pathway: a basic patch in the tail of histone H4, consisting of the amino acids R17H18R19, is required for H3K79 di- and trimethylation both in vitro and in vivo (51;52). In contrast, H3K79 monomethylation is still present when this basic patch is mutated (52). The regulatory effect is due to a direct interaction: an acidic patch in the C-terminus of Dot1 mediates binding of the protein to the basic region in the H4 tail (52). Interestingly, the silencing protein Sir3 binds to the same basic patch on histone H4 through its C-terminus (53), suggesting that Dot1 and Sir3 might compete for the same binding site. Indeed, methylation of H3K79 by Dot1 in in vitro methyltransferase assays was strongly inhibited by the addition of recombinant Sir3 C-terminus or Sir3 protein purified from yeast (51;52). Finally, Sir3 also binds to the nucleosome surface around H3K79 through its C-terminus (51).

These data suggest a model where binding of Dot1 and Sir3 to a nucleosome are mutually exclusive: methylation of H3K79 and/or binding of Dot1 to histone H4 prevent Sir3 binding to nucleosomes, whereas binding of Sir3 through its C-terminus to the H4 tail and the region surrounding H3K79 prevents methylation by Dot1. This competition is one of the factors that determines the establishment and maintenance of the balance between euchromatin and heterochromatin. In Chapter 2 we present data to extend this model further: our genetic experiments strongly suggest that Sir3 binds to the H3K79 nucleosome surface not only via its C-terminus, but also via its N-terminal BAH domain. Others have demonstrated that the BAH domain of Sir3 can interact with nucleosomes in vitro (54). The proper recognition of unmethylated H3K79 by Sir3 seems to be dependent on acetylation of the N-terminal alanine on Sir3: without this acetylation the protein loses its specificity for unmethylated H3K79 (Chapter 2). Finally, a recent biochemical study has shown that the N-terminus of Sir3 can bind to nucleosomes in vitro and that this interaction is inhibited by methylation of H3K79 (55). Thus, it seems that Sir3 uses three different mechanisms of interaction with the nucleosome core: binding via its C-terminus to the tail of H4, and binding via its C- or N-terminus to the region around H3K79. All of these interactions together serve to prevent interaction of Dot1 with its target H3K79 and thereby regulate the levels of H3K79 methylation. We found that the DOT1A and DOT1B enzymes from T. brucei do not require the tail of histone H4 for their activity (Chapter 5), indicating that this mechanism of regulation does not exist in...
trypanosomes. Whether mammalian Dot1 requires interactions with histone tails has not been established.

H4 is not the only histone that can affect Dot1 activity. Nucleosomes containing H2A.Z, a variant of histone H2A, have low levels of H3K79me2. In addition, the presence of H2A.Z in a nucleosome inhibits the activity of Dot1 \textit{in vitro} (56). The mechanism of this inhibition is currently unclear. H2A.Z is present in ~5-10% of the genome and is predominantly localized to the two nucleosomes flanking the nucleosome free region of the transcription start site (56-58).

Recently, another possible mode of regulation of Dot1 activity was described. In a screen for proteins that affect the levels of H3K79 dimethylation but not trimethylation on immunoblots (39), the proteins Swi4 and Swi6, together forming the SBF complex, were found to alter Dot1 activity. The levels of H2B ubiquitination were not affected, ruling out the possibility that the regulation occurred via this pathway. It is unclear by which mechanism Swi4/6 regulate Dot1 activity, but the authors suggest that the SBF complex might recruit factors required for Dot1 activity to chromatin (39). However, we were unable to detect any effect of deletion of Swi4 or Swi6 on H3K79 methylation levels (data not shown), making the impact of this finding somewhat uncertain.

Proteins that bind directly to Dot1 are also good candidates to be involved in regulation of its activity. However, no binding partners for Dot1 have been described in yeast. Large-scale tandem affinity purification screens have yielded

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{nucleosome.png}
\caption{H3K79 and H2BK123 are located in close proximity on the nucleosome core. Space-filling model of the yeast nucleosome (1) with arrows indicating the positions of H3K79 and H2BK123.}
\end{figure}
some candidate binders (59), but we were unable to confirm any of these using direct co-immunoprecipitation experiments (data not shown). However, in mammalian cells several Dot1 binding proteins have been identified. Murine Dot1L can bind to the MLL fusion partner AF17 and this interaction promotes a cytoplasmic localization of Dot1L, resulting in decreased overall levels of H3K79 methylation (60). In addition, direct interactions with several other MLL fusion partners have been described (50). Interestingly, hDot1L binds to a number of proteins from the Wnt signaling pathway, such as β-catenin. These interactions serve to recruit Dot1 to Wnt target genes and thereby alter the expression of these genes (50), suggesting that the regulation takes place at the level of localization rather than activity.

Finally, Dot1 activity can be indirectly regulated by differential recruitment of H2B deubiquitinating enzymes. Through an interaction with Sir4, the deubiquitinating enzyme UBP10 is recruited to heterochromatin, which results in low levels of H2B ubiquitination and thereby also inefficient H3K79 methylation (61;62).

What are the molecular mechanisms of Dot1 regulation?

Genetics has taught us about modifications and proteins that are involved in the regulation of Dot1 activity. Regulatory pathways have been delineated, mainly by specific deletion of putative regulators and assessment of the effect on H3K79 methylation using immunoblots. To gain more knowledge about the exact mechanism by which the regulation takes place and about the molecular details, different experiments are needed.

Traditionally, in vitro approaches have been used to get down to the molecular level. To identify specific interactions between proteins and histone modifications, pulldown assays are often used. In these assays modified peptides are fused to an epitope tag of choice and added to a cell extract. The epitope tag can be used for easy purification of the peptides, which will also bring down proteins from the cell extract that have bound to the modification. Through comparison between modified and unmodified peptides specific binding proteins can be identified (63). However, the fact that H3K79 is located on the core of the nucleosome in close contact with many other (modified) residues means that a simple linear peptide does not recapitulate the in vivo context of this residue very well. In addition, since Dot1 is not active on peptide substrates (see below), it is questionable whether pulldown assays give a reliable indication of its affinity for the different methylated states of H3K79.
Another widespread technique is the use of peptides for *in vitro* methyltransferase assays. Unmodified and methylated peptides are used as a substrate for a recombinant enzyme, to determine its affinity for and activity on the different methylation states of a residue. Often these assays are done with radiolabeled methyl donor, to allow reliable quantification of the resulting amount of methylation on the different peptides. However, it is difficult to study the activity of Dot1 in this type of assay. In contrast to the methyltransferases that modify the histone tails, Dot1 is not active on peptide substrates or recombinant free histones. Yeast, human, and *T. brucei* recombinant Dot1 proteins can only methylate H3K79 in the context of chromatin or the nucleosome, which implies that interactions with the nucleosome core are required for its activity (8-11;23;64;65). One solution to this problem is to use not peptides or recombinant histones as a substrate in the *in vitro* methyltransferase reactions, but purified chromatin from a *dot1Δ* yeast strain, which does not contain any methylated H3K79.

Recently, a novel approach has been developed to generate recombinant histones containing methyl-lysine analogs at any desired position (66). In this method a cysteine residue in the recombinant protein is chemically converted into N-methylated aminoethylcysteine, an analog of methylated lysine. By using different alkylating agents, specific analogs of mono-, di- or trimethylated lysine can be generated. The recombinant histones can be used to assemble recombinant nucleosomes containing specifically methylated lysines (66). This technique has been used to study the effect of methylation of H3K79 on the crystal structure of the nucleosome (67). Nucleosomes containing these methylated lysine analogs could also be used as a substrate in *in vitro* assays to determine the affinity and activity of Dot1 for (methylated) H3K79. However, even though several methylated lysine analogs closely mimic the function of their natural counterparts in functional assays (66), the chemical environment is still slightly different and this could affect the kinetics of the methylation reactions.

Several unconventional approaches have been taken to circumvent these problems. To determine the kinetic mechanism of methylation by Dot1, we have combined *in vitro* methyltransferase assays with *in vivo* methods using quantitative mass spectrometry (Chapter 3). We next developed a mathematical model that describes distributive histone methylation. This model allowed us to address questions of Dot1 regulation and the dynamics of H3K79 methylation during the cell cycle (Chapter 4). Other groups have used new techniques based on mass spectrometry and chemical biology to study the dynamics of methylation
and the regulation of Dot1 by ubiquitinated H2B, respectively. These approaches will be discussed in the next sections.

**Kinetic mechanism of Dot1 and functional implications**

There are two different catalytic mechanisms of multiple methylation: methyltransferases can be processive or distributive (Fig. 4). A processive enzyme performs consecutive rounds of methylation without dissociation from the substrate residue. In contrast, a distributive methyltransferase dissociates from the substrate after each round of methylation. Most SET domain histone methyltransferases are processive enzymes (13;68). However, the *Drosophila* SU(VAR)3-9 enzyme, which contains a SET domain, has been shown to act in a non-processive manner *in vitro* (69). When targeted to a specific location, processive enzymes can introduce a specific methylation state. The processivity of SET enzymes has been suggested to be regulated by enzyme-binding partners or crosstalk by other histone modifications. Distributive enzymes cannot generate specific methyl states and are therefore expected to have different modes of regulation. Therefore, to understand the function and regulation of H3K79 methylation, it is important to understand the catalytic mechanism of multiple methylation by Dot1. Both a distributive and a processive mechanism of methylation have been suggested, based on the crystal structures of the yeast and human enzymes (64;70).

Traditionally, *in vitro* methyltransferase assays have been used to determine the kinetic mechanism of a specific enzyme. Using *in vitro* assays with purified chromatin as the substrate, we found that Dot1 was most likely a distributive enzyme (Chapter 3). However, the most important question is what
mechanism the enzyme uses in vivo, where the kinetic mechanism can be affected by associating factors or other interactions. We decided to design an in vivo methyltransferase assay: using an inducible expression system we manipulated the levels of Dot1 in vivo and quantified the resulting levels of H3K79 methylation using mass spectrometry. This novel approach confirmed that Dot1 is indeed a distributive enzyme in vivo (Chapter 3).

The distributive mechanism of methylation has functional implications, as we show in Chapter 3. It suggests that H3K79me1 and –me2 cannot be generated specifically, since these methylation states are by definition substrates for further methylation by Dot1. We reasoned that this could imply that these methylation states of H3K79 do not have specific functions. Indeed, we were able to show that the level of telomeric silencing is determined by the overall methylation of H3K79 and not by a specific methylation state (Chapter 3). This model is also supported by data showing that Sir3 binding to histone H3 is inhibited by all three methylation states of H3K79 (51).

Notably, there are also studies that have suggested there might be specific functions for specific H3K79 methylation states (39;71;72). In particular, a recent report found very different genome-wide localizations for H3K79me2 and –me3 in yeast. Furthermore, genes specifically regulated in M/G1 were enriched for H3K79me2 but not –me3, whereas G2-regulated genes were enriched for H3K79me3 but not –me2 (39). Unfortunately, the distribution of H3K79me1 was not addressed in this study. Finally, the absence of known H3K79 demethylases suggests that the H3K79me3 mark might be very stable (73). Once a gene has acquired H3K79 trimethylation, it will stay trimethylated unless histone turnover occurs. This means that it is still very well possible that especially H3K79me3 has specific functions. However, no such functions have been described to date. In addition, a systematic search has not identified any proteins that can bind specifically to any methylation state of H3K79 (36).

SILAC labeling and mathematical modeling reveal dynamics of histone methylation

The histone modification pattern of a cell is an important factor in determining the gene expression pattern and thereby the identity of that cell. The general view is that this pattern of histone modifications is more or less fixed, with the exception of acetylation and methylation on gene promoters which are quite dynamic. The pattern is disrupted during replication, because the histones that are deposited on newly duplicated DNA contain few modifications (74). However, it is generally
thought that the histone modifications are rapidly re-established on the new nucleosomes, resulting in two daughter cells that are indistinguishable from the mother cell. But in fact, depending on the catalytic rate of the histone modifying enzymes, it might take considerable time to recapitulate the histone modification pattern that is present on the parental histones. In addition, turnover of histones and the activity of histone demodifying enzymes also contribute to a dynamic situation, suggesting that some histone modifications might be more static than others.

Several studies have addressed the establishment of specific histone modifications after replication using mass spectrometry. Stable Isotope Labeling by Amino acids in Cell culture, or SILAC, is a recently developed method that is very useful to study the kinetics of histone modifications. In SILAC experiments, cells are arrested and then transferred to media containing an isotopically labeled form of an essential amino acid as the only source of this amino acid. Consequently, all newly synthesized proteins are stably labeled with this isotope and can be distinguished from the proteins that were already present using mass spectrometry (75). Using SILAC, it was shown that after release from a cell cycle block, newly synthesized molecules of histone H4 in HeLa cells become progressively and irreversibly methylated on K20. It takes two to three cell cycles before virtually all new H4 is dimethylated on K20, demonstrating that the establishment of this mark is quite slow (76). Extending on this work, another group showed using SILAC that the patterns of lysine acetylation are rapidly re-established within two hours after histone deposition. In contrast, six hours after labeling, methylation patterns on newly incorporated histones were still very different from those on the parental histones. In fact, for several methyl marks the re-establishment took most of the cell cycle (77). The use of SILAC to study histone modification dynamics was taken one step further by the development of a quantitative mathematical model to characterize the methylation rates of specific lysine residues (73). In this study, the authors used the quantitative proteomics data from their SILAC experiments to determine the rates of formation of the mono-, di- and trimethylated states of a lysine residue. Interestingly, they found that in general the rate of formation becomes progressively slower going from mono- to di- to trimethylation. Furthermore, the set-up allowed simultaneous determination of the rate constants for demethylation of the lysine residues. No detectable demethylation was found for H3K79, in line with the fact that there is no known demethylase (73).
We developed a mathematical model to describe *in vivo* H3K79 methylation in replicating yeast cells (Chapter 4) and found that the H3K79 methylation pattern does not reach a steady state. This was also demonstrated experimentally showing that cell cycle arrest and extension of the cell cycle length both result in higher levels of H3K79 methylation. Our data suggest that newly synthesized histones are progressively methylated during the cell cycle and that H3K79 methylation accumulates on histones as they age. Simulation of the methylation pattern of a single cell during the cell cycle supported this model (Fig. 5). In addition, we have also been able to demonstrate this concept directly by purifying old and newly synthesized histones and analyzing them by mass spectrometry. The ratio of trimethylated over monomethylated H3K79 was significantly higher on old histones, indicating that H3K79 methylation is indeed a measure of histone age (Chapter 4). Therefore, our results and those of others suggest that the histone modification pattern of a cell is not necessarily static. For some modifications, such as lysine acetylation, the pattern can be rapidly adjusted to a new steady state. However, for methylation of H3K79, it is not possible to define ‘the’ modification pattern of a cell, since this pattern changes throughout the cell cycle.

**Modeling and chemical biology shed light on the regulation by ubiquitinated H2B**

Genetic experiments have firmly established the connection between H2B ubiquitination and H3K79 methylation, as outlined above. However, how the regulation takes place on the molecular level is still unclear. One important question is whether H2B ubiquitination regulates all methylation states of H3K79.
or specifically di- and trimethylation. Strains lacking H2B ubiquitination show a loss of H3K79me2 and –me3, but an increase in H3K79me1 (38). This result led to the suggestion that H2B ubiquitination specifically regulates the higher methylation states of H3K79 (38).

We examined the effect of loss of H2B ubiquitination in a wild type strain and confirmed that it results in loss of H3K79me2 and –me3. However, when a catalytically compromised Dot1 enzyme was used, which only generates H3K79me1 and –me2, loss of H2B ubiquitination led to a complete loss of all H3K79 methylation states (Chapters 3 and 4). This result suggests that ubiquitination of H2B affects the general activity of Dot1, leading to slower accumulation of methyl groups on H3K79 in the absence of H2B ubiquitination. Due to the distributive mechanism of methylation, a partial reduction in Dot1 activity will result in loss of mainly H3K79me3 and –me2. This idea was confirmed by the mathematical model of methylation by Dot1, which showed that the rate constants of all three methylation reactions were reduced (Chapter 4), indicating that it is indeed the general activity of Dot1 that is stimulated by ubiquitination of H2B.

A second important question is whether the effect of ubiquitinated H2B on Dot1 activity is direct or mediated by other proteins. For detailed mechanistic studies reasonable amounts of nucleosomes ubiquitinated on H2B are required. However, it is difficult to obtain these by purification from cells, since only a small fraction of all H2B is ubiquitinated (43). One solution is to use chromatin from cells lacking the H2B deubiquitinating enzymes Ubp8 and Ubp10, which have higher levels of ubiquitinated H2B (61;78). However, in any case the purified nucleosomes will be very heterogeneous with respect to the other histone modifications that are present (79). A solution to this problem is to site-specifically ubiquitinate H2B molecules and use these to reconstitute chemically defined nucleosomes. A semi-synthetic strategy to produce H2B ubiquitinated on K120 (the human homolog of K123) has been developed based on the expressed protein ligation (EPL) technology. In short, a synthetic peptide of the C-terminus of H2B was used with an N-terminal cysteine substitution for an alanine. This peptide was ligated to a recombinant ubiquitin molecule which contained a C-terminal α-thioester. In a second step, the intermediate product was ligated to recombinant N-terminal H2B-α-thioester. Finally, the cysteine was converted back to alanine, resulting in a traceless H2B-ubiquitin protein (80). In a subsequent study, the yield and speed of synthesis were greatly increased by using ubiquitin bearing a G76A mutation (81).
For biochemical analysis of the activity of human Dot1L, the semi-synthetic H2B molecules were incorporated into mono- and dinucleosomes. In vitro methyltransferase assays showed that recombinant hDot1L was able to methylate mononucleosomes containing ubiquitinated H2B, whereas it did not show any activity on unmodified nucleosomes. This result indicates that ubiquitinated H2B directly stimulates hDot1L activity and does not work via other proteins (80). Interestingly, this effect does not occur in trans: ubiquitination of one of the nucleosomes in a dinucleosome did not stimulate methylation of the other, non-ubiquitinated nucleosome (80). Finally, hDot1L was able to bind equally well to ubiquitinated and unmodified nucleosomes, suggesting that ubiquitination of H2B does not alter recruitment of hDot1L but specifically affects its activity, possibly through allosteric regulation (80). This is consistent with ChIP results from yeast showing that Dot1 is only detectable on active genes (38), suggesting that its recruitment is regulated by as yet unknown factors.

In line with our experimental data and modeling results, H2B ubiquitination stimulated the general activity of hDot1L in vitro, resulting in mono- and dimethylation of H3K79. No trimethylation was observed (80). However, this last finding seems to be an artefact caused by the protease that was used to prepare the samples for mass spectrometry: it has been shown very recently that hDot1L is capable of trimethylating ubiquitinated nucleosomes in similar in vitro experiments (50).

Even though we know that ubiquitinated H2B stimulates the activity of hDot1L directly, the exact molecular mechanism has not been solved yet. However, more and more details are becoming clear. Usually, ubiquitin interacts with other proteins via a hydrophobic patch, but mutation of this patch did not interfere with hDot1L stimulation, suggesting that hDot1L is activated in a non-canonical way (81). Furthermore, when three different ubiquitin-like proteins were attached to K120 instead of canonical ubiquitin, only Nedd8 was able to stimulate hDot1L activity, showing that the activation of Dot1 is not restricted to ubiquitin but also that a ubiquitin fold is not sufficient for activation (81;82). Finally, when an ubiquitin group was introduced not on K120 but on K125 of H2B or on K22 of histone H2A, hDot1L activity was stimulated to levels comparable to stimulation by H2B ubiquitinated on K120. This result indicates that the sequence around the ubiquitinated lysine is not critical for hDot1L activation (82). For a more detailed picture we will probably have to await the elucidation of the crystal structure of Dot1 or hDot1L in complex with a nucleosome containing modified or unmodified H2B, or the application of other techniques to report on structural rearrangements of the Dot1 protein upon binding to ubiquitin.
Chapter 1

Model for the regulation of H3K79 methylation

This chapter will conclude with a general model describing how the levels of H3K79 methylation are regulated in yeast (Fig. 6). Newly synthesized histones are progressively methylated on H3K79 in the course of the cell cycle. Because Dot1 is a distributive enzyme, new methyl groups are added one-by-one, meaning that the methylation pattern changes over time from predominantly lower to predominantly higher methylation states. This overall build-up of methylation on H3K79 can be modulated by factors such as ubiquitination of H2B, interactions between Dot1 and the tail of H4, proteins that bind to Dot1 and affect its targeting and a putative demethylase for H3K79. On top of these specific modes of regulation, H3K79 methylation levels are also affected by the rate of cell division (and thereby the rate of histone deposition) and the rate of histone turnover in active genes. The net level of H3K79 methylation in a specific chromatin region depends on the relative contributions of all of these factors.

In general, stimulation or inhibition of Dot1 activity in specific genomic locations will generate specific H3K79 methylation states. If Dot1 activity is slowed down, this will result in loss of H3K79me3 and the presence of only H3K79me1 and –me2, whereas stimulation will lead to an increase in H3K79me3. Therefore, even if H3K79me1 and –me2 are always a substrate for Dot1, by tuning the activity of the enzyme predominance of a specific methylation state can be achieved. This concept is illustrated by the best-known regulator of Dot1 activity, ubiquitinated H2B. Recent genome-wide ChIP data show that in yeast H2B ubiquitination specifically colocalizes with regions enriched for H3K79me3: about one-third of genes enriched for H3K79me3 are also marked by H2B ubiquitination. In contrast, ubiquitination is low in regions enriched for H3K79me2 (39). The consistent colocalization of H3K79me3 and H2B ubiquitination suggests that this latter modification is the main determinant of the H3K79me3 patterns. This result fits well with the model: in regions containing ubiquitinated H2B, Dot1 activity is stimulated leading to an increase in H3K79me3. The same mechanism probably applies to the interaction of Dot1 with the tail of H4: in regions where the H4 tail is bound by Sir3 (i.e. silent chromatin regions), Dot1 activity cannot be stimulated through this interaction and therefore H3K79 methylation levels are lower. In yeast, there is no clear correlation between the expression level of a gene and its H3K79 methylation pattern (37-39). In contrast, in flies and mammals, a higher transcriptional frequency is reflected in higher levels of H3K79me2 and –me3 (30;40;41). We propose that this difference results from a difference in the balance between Dot1 activity
Introduction

Figure 6. Model to summarize the main factors that regulate the levels of H3K79 methylation in yeast and to explain the observed H3K79 methylation patterns in the yeast genome. Five different regions of the yeast genome are shown: a gene with low expression, a gene with high expression, a gene located in heterochromatin, an intergenic region, and two H2A.Z containing nucleosomes flanking the nucleosome free region (NFR) of the transcription start site. Transcription leads to accumulation of H2B ubiquitination, which therefore marks genes A and B. Dot1 is recruited to active gene B by unknown mechanisms and is stimulated by the ubiquitin group present on H2B, leading to an increase in H3K79 methylation. However, transcription also leads to histone turnover, which counteracts the accumulation of methylation on gene B. Recruitment of Dot1 to the less active gene A is lower, but the turnover of this gene is also lower. Therefore, the net levels of H3K79 methylation will be roughly similar on genes A and B. An intergenic region has low levels of H2B ubiquitination and therefore the Dot1 activity is low. In addition, the histone turnover in intergenic regions is relatively high (2), resulting in the presence of only H3K79me1 and –me2. In a similar way, low Dot1 activity and very high histone turnover of the H2A.Z containing nucleosomes around the nucleosome free region prevent the accumulation of higher methylation states of H3K79. Finally, heterochromatic gene C is bound by Sir proteins, preventing methylation by Dot1. In addition, Sir4 can recruit the deubiquitinating enzyme Ubp10, which prevents the generation of ubiquitinated H2B. These two factors, combined with some histone turnover, result in very low levels of H3K79 methylation in heterochromatin. In a situation of growth arrest or slow growth, changes in the level of H3K79 methylation on genes A and B will be minor, because the overall methylation on these genes is already quite high in normally growing cells. However, intergenic regions can acquire more H3K79 methylation because there is more time for Dot1 to methylate before histones are diluted out by cell duplication. Therefore, these regions are expected to gain a methylation pattern that is more similar to that of a euchromatic coding sequence.

and histone turnover between the species. In general, the mechanism is as follows: transcription results in ubiquitination of H2B (39;78;83), which stimulates H3K79 methylation by Dot1. Since Dot1 is recruited to active genes
(38), this means that high transcription leads to high levels of H3K79 methylation. In contrast, lower expressed genes are also marked by ubiquitinated H2B (39), but have lower levels of methylated H3K79 because Dot1 is not actively recruited to these genes. However, the effect of transcription on Dot1 activity is counteracted by the increase in histone turnover that also results from the increased transcription (2;84;85). Higher histone turnover leads to more deposition of new, unmodified nucleosomes and therefore to lower levels of H3K79 methylation. In yeast the balance between these two effects is apparently such that the increase in H3K79 methylation due to transcription is cancelled by the increase in histone turnover, resulting in more or less similar levels of H3K79 methylation across all genes. However, for very highly expressed genes, which have high rates of turnover in the coding sequence (2;84), this balance breaks down: the most active yeast genes have very low levels of H3K79me2 and –me3 in their open reading frames (39). This indicates that at very high rates of histone turnover, Dot1 activity is not sufficient to achieve ‘normal’ levels of H3K79 methylation in yeast. The balance between Dot1 activity and histone turnover seems to be different in flies and mammals. The clear correlation between transcription and H3K79 methylation in these species suggests that either the Dot1 activity is higher than in yeast, or the global histone turnover is lower. In either case, this means that the enzyme can ‘keep up’ with the higher levels of histone turnover that are the result of ongoing transcription.

In mouse embryos H3K79me1/2 and H3K79me3 differentially mark the maternally and paternally inherited alleles of imprinted genes. Where expression is biased to one of the alleles, H3K79me3 is found on the CpG-methylated and therefore repressed allele, whereas the active and CpG-unmethylated allele is marked by HK79me1 and –me2 (72). This situation can be explained by higher histone turnover of the active allele, preventing accumulation of H3K79me3 on this allele. The idea that the differential marking of the two alleles of an imprinted gene is the consequence rather than the cause of allele-specific transcription is supported by experiments in which Dot1L-expression was reduced. Knockdown of Dot1L did not affect the overall expression of imprinted genes; in addition, the imprinted genes were still expressed from the correct paternal allele after knockdown (72).

In conclusion, we have presented a model for the regulation of H3K79 methylation based on the distributive mechanism of methylation by Dot1. In general, the methylation level progressively increases in the course of the cell cycle, but several different regulators can alter the balance between the methylation states
by affecting Dot1 activity. The net level of methylation therefore depends on factors that can be different in different parts of the genome, such as transcription rate and Dot1 recruitment. An important question to address in future experiments is whether H3K79 methylation is simply a measure of histone age, or whether this ‘clock’-like behaviour is critical for proper cell function.

REFERENCES


31
Chapter 1


