A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin

Gunnar Schotta,1,4 Monika Lachner,1,4 Kavitha Sarma,2 Anja Ebert,3 Roopsha Sengupta,1 Gunter Reuter,3 Danny Reinberg,2 and Thomas Jenuwein1,5

1Research Institute of Molecular Pathology (IMP), The Vienna Biocenter, A-1030 Vienna, Austria; 2Howard Hughes Medical Institute, Division of Nucleic Acids Enzymology, Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, USA; 3Martin-Luther University Halle-Wittenberg, Biologicum, 06120 Halle, Germany

Histone lysine methylation is a central modification to mark functionally distinct chromatin regions. In particular, H3-K9 trimethylation has emerged as a hallmark of pericentric heterochromatin in mammals. Here we show that H4-K20 trimethylation is also focally enriched at pericentric heterochromatin. Intriguingly, H3-K9 trimethylation by the Suv39h HMTases is required for the induction of H4-K20 trimethylation, although the H4 Lys 20 position is not an intrinsic substrate for these enzymes. By using a candidate approach, we identified Suv4-20h1 and Suv4-20h2 as two novel SET domain HMTases that localize to pericentric heterochromatin and specifically act as nucleosomal H4-K20 trimethylating enzymes. Interaction of the Suv4-20h enzymes with HP1 isoforms suggests a sequential mechanism to establish H3-K9 and H4-K20 trimethylation at pericentric heterochromatin. Heterochromatic H4-K20 trimethylation is evolutionarily conserved, and in Drosophila, the Suv4-20 homolog is a novel PEV modifier to regulate position-effect variegation. Together, our data indicate a function for H4-K20 trimethylation in gene silencing and further suggest H3-K9 and H4-K20 trimethylation as important components of a repressive pathway that can index pericentric heterochromatin.

Keywords: Histone code; histone H4 Lys 20; mono-, di-, trimethylation; Suv4-20h HMTases; heterochromatin; combinatorial histone methyl marks

Supplemental material is available at http://www.genesdev.org.

Received February 19, 2004; revised version accepted April 5, 2004.

Chromatin is the physiological template of the genetic information and is composed of the DNA polymer, histones, and other chromosomal proteins. The basic repeating unit of chromatin is the nucleosome octamer, which is built by the core histones H2A, H2B, H3, and H4 wrapping 147 bp of DNA [Luger et al. 1997]. Histone N termini [tails] are flexible and protrude from the nucleosome octamer. Histone tails are subject to post-translational modifications, including acetylation, methylation, phosphorylation, ubiquitination, and ADP-ribosylation [van Holde et al. 1988]. During the recent years, many histone modifying enzymes and several chromatin-associated proteins that specifically bind to these modifications have been identified. Particular histone modification patterns have been associated with distinct chromatin states and are proposed to represent an indexing mechanism that could extend the information potential of the genetic code [Strahl and Allis 2000; Turner 2000; Jenuwein and Allis 2001].

Histone modifications can act as either repressive or active marks. For histone lysine methylation, modification of H3-K4, H3-K36, and of H3-K79 have been correlated with transcriptional activation, whereas methylation of H3-K9, H3-K27, and H4-K20 are marks of repressive chromatin states [for review, see Fischle et al. 2003; Lachner et al. 2003; Vaquero et al. 2003]. Histone lysine residues can be mono-, di-, or trimethylated [Paik and Kim 1971; DeLange et al. 1973], thereby extending the coding potential of a methylatable lysine position. Indeed, the distinctions between di- and trimethylation of H3-K4 [Santos-Rosa et al. 2002] or of H3-K9 residues [Czermin et al. 2002; Kuzmichev et al. 2002; Tamaru et al. 2003] have been shown to be relevant for transcriptional regulation or Polycomb-mediated gene silencing. Distinct methylation states or combinations between several methylation marks could further discriminate
different chromatin regions or entire chromosomes. For example, H3-K27 trimethylation in conjunction with H3-K9 dimethylation are considered epigenetic imprints of the inactive X chromosome [Plath et al. 2003; Silva et al. 2003; Okamoto et al. 2004], whereas H3-K9 trimethylation and H3-K27 monomethylation are associated with pericentric heterochromatin [Peters et al. 2003; Rice et al. 2003]. The combinatorial nature of histone lysine modifications requires controlled interplay between different histone lysine methylation systems. Loss of a given HMTase may also affect methylation on lysine residues for which the enzyme has no intrinsic activity. This is exemplified by the conversion of pericentric H3-K27 methylation to H3-K27 trimethylation in the absence of Suv39h enzymes [Peters et al. 2003].

Methylation of H4-K20 has been described as another mark of repressive chromatin domains [Nishioka et al. 2002; Sims et al. 2003]. However, previous studies were focused toward detection of H4-K20 methylation, regardless of the methylation status and characterized PR/SET7 [Nishioka et al. 2002a; Rice et al. 2002] or SET8 [Fang et al. 2002] as the responsible enzyme. Here, we examine all three H4-K20 methylation states in mammalian chromatin with highly selective antibodies that discriminate H4-K20 mono-, di-, and trimethylation. We demonstrate that H4-K20 trimethylation is a novel and evolutionarily conserved mark of pericentric heterochromatin. In addition, we identify two murine SET domain HMTases that localize to pericentric heterochromatin and induce H4-K20 trimethylation in a Suv39h-dependent manner. In Drosophila, mutants of the corresponding gene locus are dominant suppressors of position-effect variegation. These data establish H4-K20 trimethylation as a repressive mark in gene silencing mechanisms and suggest that the sequential induction of H3-K9 and H4-K20 trimethylation by distinct histone lysine methylation systems can index repressive chromatin domains.

Results

H4-K20 trimethylation is a novel mark of pericentric heterochromatin

Pericentric heterochromatin in mammals mainly consists of major satellite repeats and can easily be visualized by the fluorochrome DAPI, which preferentially intercalates with A/T-rich repeat sequences. Differential association of H3-K9 and H3-K27 methylation states in euchromatin versus heterochromatin [Peters et al. 2003] prompted us to analyze other potentially repressive marks, such as histone H4-K20 methylation. Highly specific antibodies that discriminate H4-K20 methylation states were developed by Upstate Biotechnology [UBI] and quality controlled by dot blots presenting a panel of 23 histone tail peptides [Supplementary Fig. S1] and by Western blots of mouse nuclear extracts [Supplementary Fig. S2]. Antibody specificity was further confirmed by peptide competition assays in immunofluorescence analyses [Supplementary Fig. S3]. By using these and our previously characterized H3-K9 methyl antibodies, we performed comparative immunofluorescence in wild type and Suv39h double-null (dn) female mouse embryonic fibroblasts [MEFs]. In wild-type MEFs, H4-K20 monomethylation is dispersed in euchromatin but shows focal enrichment at the inactive X chromosome [Fig. 1A, arrow; Kohlmaier et al. 2004]. H4-K20 dimethylation is also broadly distributed over euchromatic regions but shows a more speckled pattern, which resembles H3-K9 dimethylation [Fig. 1A]. In contrast, H4-K20 trimethylation is strongly enriched at DAPI dense regions and reflects the characteristic accumulation of H3-K9 trimethylation at pericentric heterochromatin [Fig. 1A]. We also analyzed H4-K20 methylation states in interphase chromatin of HeLa cells. H4-K20 mono- and dimethylation are uniformly distributed throughout the nuclei [Supplementary Fig. S4]. Notably, H4-K20 monomethylation displays variable signal intensities, suggesting a potential cell-cycle regulation for this mark. In contrast, H4-K20 trimethylation is enriched at several nuclear foci, which represent pericentric heterochromatin as demonstrated by immunofluorescence analysis of metaphase spreads [Supplementary Fig. S4].

An interplay between distinct methylation systems at pericentric heterochromatin has been described previously [Peters et al. 2003]. For example, disruption of the Suv39h enzymes results in the loss of H3-K9 trimethylation by also converting H3-K27 monomethylation to H3-K27 trimethylation. Because of the strikingly similar accumulation of H4-K20 trimethylation and H3-K9 trimethylation at pericentric heterochromatin, we analyzed whether H4-K20 methylation states may also depend on the presence of the Suv39h enzymes. In Suv39h dn female MEFs, H4-K20 mono- and dimethylation are not altered, but H4-K20 trimethylation is entirely lost from pericentric heterochromatin [Fig. 1A]. These data were confirmed by analyzing metaphase chromosomes of wild-type and Suv39h dn MEFs. In wild-type mitotic spreads, H4-K20 trimethylation is strongly enriched at pericentric regions and shows a diffuse staining pattern along the chromosomal arms [Fig. 1B]. In contrast, in mitotic spreads from Suv39h dn MEFs, pericentric H4-K20 trimethylation is lost, whereas the chromosomal arms display enriched signals [Fig. 1B].

From these data, we conclude that the presence of the Suv39h enzymes can direct pericentric H4-K20 trimethylation.

Identification of novel heterochromatic SET domain proteins

The requirement of Suv39h enzymes for pericentric H4-K20 trimethylation raised the question whether Suv39h enzymes might contain an intrinsic activity toward the H4-K20 position. However, in previous and extended in vitro HMTase assays [data not shown], the recombinant Suv39h enzymes only target the H3-K9 position, with a weak activity also toward histone H1 [Rea et al. 2000;
Therefore, other enzymes must exist that can trimethylate the H4-K20 position. H4-K20 methylation is conserved among eukaryotes, such as *Schizosaccharomyces pombe, Caenorhabditis elegans, Drosophila*, and mammals, but is not present in *Saccharomyces cerevisiae* [Fang et al. 2002; data not shown]. We therefore chose a candidate approach to identify H4-K20-specific HMTases by comparing all SET domain proteins that are shared among mouse, *Drosophila*, and *S. pombe*. We reasoned that HMTases with similar enzymatic specificities could be classified by their relative homology within the SET domain [Fig. 2]. Within each of the resulting subgroups, the catalytic specificity of known SET domain HMTases is indicated [Fig. 2, gray boxes]. Candidate proteins that are conserved between *Drosophila* and mouse were selected for further analyses and are shown by an asterisk [Fig. 2]. Full-length IMAGE clones (RZPD) were used to express cDNAs as EGFP fusion proteins in MEFs under control of a cytomegalovirus (CMV) promoter. Of the 12 candidate cDNAs tested, Cgi-85 (Suv4-20h1, see below) and Mgc2705 (Suv4-20h2, see below) show a heterochromatic accumulation in wild-type MEFs, which is lost in *Suv39h* dn cells [Fig. 3B]. In contrast, all other candidate SET domain proteins, including a previously described H4-K20 HMTase [Fang et al. 2002; Nishioka et al. 2002; Rice et al. 2002], displayed broad nuclear staining patterns [data not shown].

Cgi-85 [876 amino acids] and Mgc2705 [468 amino acids] are two closely related proteins present in mammals, whereas only one corresponding gene product is found in *Drosophila* and *S. pombe* [Supplementary Fig. S2]. As we will show below, the *Drosophila* homolog, CG13363, is a dominant suppressor of position-effect variegation. We therefore classify *Cgi-85* and *Mgc2705* as novel *Su(var)* genes and suggest to rename these loci as *Suv4-20h1* and *Suv4-20h2*.

With the exception of the SET domain, which is located close to the N terminus, no other conspicuous domains could be identified in *Suv4-20h* proteins. However, the SET domain is surrounded by regions that are highly conserved within these orthologs. In addition, there is a small region (~10 amino acids) in the C termini that is shared between the various *Suv4-20h* proteins [Fig. 3A; Supplementary Fig. S2].

**Figure 1.** H3-K9 and H4-K20 methylation states in wild-type and *Suv39h* double-null (dn) MEFs. (A) Female wild-type and *Suv39h* dn MEFs were stained with antibodies directed against H3-K9 mono-, di-, and trimethylation [upper panel] or H4-K20 mono-, di-, and trimethylation [lower panel]. DAPI dense foci represent pericentric regions. The inactive X chromosome is enriched for H4-K20 monomethylation and indicated by an arrow. (B) Metaphase spreads of wild-type and *Suv39h* dn MEFs were stained with the H4-K20 trimethylation antibody. Pericentric enrichment of this methylation mark is lost in *Suv39h* dn cells.

**Suv4-20h** proteins mediate H4-K20 trimethylation at pericentric heterochromatin

Next, we investigated the function of the putative *Suv4-20h* enzymes by RNA interference [RNAi] in MEF cells.
MEFs were cotransfected with a pSUPER vector expressing hairpin RNAs and pEGFP-N1 conferring G418 resistance. Transfected cells were cultured for 5 d under selection medium and then analyzed for H4-K20 and H3-K9 methylation patterns by indirect immunofluorescence. As controls, we included RNAi for both Suv39h enzymes.

RNA hairpins directed against both Suv39h1 and Suv39h2 result in the loss of H3-K9 and H4-K20 trimethylation at pericentric heterochromatin (Fig. 3C). Thus, the Suv39h-specific RNAi knock-down accurately reflects the histone methylation defects observed in Suv39h dn cells [Fig. 1]. In contrast, RNAi knock-down of both Suv4-20h1 and Suv4-20h2 selectively impair pericentric accumulation of H4-K20 trimethylation, but do not perturb H3-K9 trimethylation [Fig. 3C].

Suv4-20h proteins are nucleosomal HMTases with specificity for H4-K20 trimethylation

Next, we characterized the activity and substrate specificity of Suv4-20h proteins. N-terminal fragments of Suv4-20h1 [amino acids 1–387] and Suv4-20h2 [amino acids 1–280] containing the SET domain were bacterially expressed as GST fusion proteins. Increasing amounts of fusion protein were used in a reaction mix containing ^3H-labeled SAM as methyl donor, and recombinant histone H4, histone octamers, or nucleosomes. Reaction products were separated by SDS-PAGE, and methyl transfer to histones was detected by fluorography. Both Suv4-20h1 and Suv4-20h2 show some HMTase activity toward recombinant octamers but display a significantly higher transfer reaction toward nucleosomes. Peptides and free histones were not accepted as substrates (Fig. 4A; data not shown). The HMTase activity appears to be specific for H4-K20, because no methyl transfer is detected when nucleosomes containing a mutated histone H4 (K20A) were used.

Because histone lysine positions can be mono-, di-, or trimethylated, we next investigated the methylation state of the reaction products. Bacterially expressed N termini of Suv4-20h1 and Suv4-20h2 show some HMTase activity toward recombinant octamers but display a significantly higher transfer reaction toward nucleosomes. Peptides and free histones were not accepted as substrates [Fig. 4A, data not shown]. The HMTase activity appears to be specific for H4-K20, because no methyl transfer is detected when nucleosomes containing a mutated histone H4 (K20A) were used.

Because histone lysine positions can be mono-, di-, or trimethylated, we next investigated the methylation state of the reaction products. Bacterially expressed N termini of Suv4-20h1 and Suv4-20h2 were used in an HMTase assay with nonradioactive SAM and recombinant substrates. Reaction products were separated by SDS-PAGE, and the methylation state of histone H4 was...
analyzed by probing Western blots with H4-K20 mono-, di-, and trimethyl-specific antibodies. No H4-K20 monomethylated product could be observed [Fig. 4B]. The weak signal visualized by the H4-K20 monomethyl antibody is a cross-reactivity with unmodified histone H4 (shown by an asterisk). In contrast, reaction products of Suv4-20h enzymes display some H4-K20 dimethylation but are strongly enriched for H4-K20 trimethylation [Fig. 4B]. The absence of H4-K20 monomethylation in reaction products of Suv4-20h2 was further confirmed by a time-course experiment. Reactions of Suv4-20h2 with recombinant nucleaseos were stopped after sequential incubation times and analyzed by Western blotting. In the reaction products, H4-K20 di- and trimethylation increased over time, but even at early time points [2 min], no H4-K20 monomethylation is detected [Supplemental Fig. S3]. The collective interpretation of the results presented above allow us to conclude that Suv4-20h enzymes are nucleosomal HMTases with a preferential activity to direct H4-K20 trimethylation.

Targeting of Suv4-20h2 by HP1 interactions

To identify the targeting mechanism[s] of Suv4-20h enzymes to heterochromatic regions, we expressed various protein truncations as EGFP fusions in MEFs. For this analysis we focused on Suv4-20h2 [468 amino acids]. The N terminus of Suv4-20h2 [S#1: amino acids 1–280] comprising the SET domain is broadly dispersed in the nucleus, whereas the C terminus [S#2: amino acids 280–468] shows prominent association with pericentric heterochromatin [Fig. 5A]. To further define the domain required for heterochromatic targeting, we also analyzed subfragments of the C terminus. One subfragment [S#4: amino acids 280–349] shows aberrant nuclear distribution. Although there are EGFP speckles in the nucleus, these do not correspond to DAPI dense heterochromatin. All other protein fragments display heterochromatic associations. Thus, the minimal heterochromatic targeting module of Suv4-20h2 is localized within amino acids 348–441.

A possible mechanism for Suv39h-dependent targeting of Suv4-20h2 could be an interaction with other heterochromatic proteins, such as HP1, the localization of which also depends on Suv39h function [Bannister et al. 2001; Lachner et al. 2001]. Therefore, we investigated whether Suv4-20h2 could directly interact with HP1 isoforms. Protein fragments of the above-mentioned localization assay were generated as recombinant GST fusion proteins and bound to glutathione-Sepharose beads. Full-length myc-tagged HP1α, HP1β, and HP1γ were expressed by in vitro translation and incubated with GST-Suv4-20h2 protein variants. After extensive washing steps, bound proteins were separated on SDS-PAGE, and Western blots were probed with α-myc antibody for detection of HP1 isoforms. The N terminus of Suv4-20h2 cannot bind HP1 isoforms, but the C terminus interacts with HP1α, HP1β, and HP1γ [Fig. 5B]. Recombinant GST does not interact with HP1, excluding unspecific binding to the GST tag [Fig. 5B]. Subfragments of the C terminus show a differential interaction pattern with HP1 isoforms. Suv4-20h2 S#5, which was found to be the minimal heterochromatic targeting module [Fig. 5A], is also the minimal subfragment that interacts with HP1α, HP1β, and HP1γ. However, the aberrantly localized subfragment S#4 does not interact with HP1α or HP1β but can still associate with HP1γ. Thus, although there are several distinct HP1 binding sites in amino acid region 280–441 (which are different from the conserved C-terminal domain), it is the interactions with HP1α and HP1β that appear to localize Suv4-20h2 to pericentric heterochromatin [Fig. 5C].

The in vitro binding data together with the heterochromatic localization of Suv4-20h2 suggest a directed targeting of this HMTase by interaction with HP1. If this model is correct, Suv4-20h enzymes and H4-K20 trimethylation should be lost in HP1-deficient cells. This question cannot easily be addressed in mammals, because of the lack of available HP1 mutants. Moreover, our attempts using RNAi-mediated knock-down of all three HP1 isoforms in MEFs failed to deplete HP1 func-
tion. We therefore extended our studies to *Drosophila*, in which H4-K20 trimethylation is present at pericentric heterochromatin and in a high number of bands (as judged by comparison with propidium iodide labeling of DNA) on salivary gland chromosomes (Fig. 6). In *Drosophila*, pericentric H4-K20 trimethylation is severely reduced in *Su(var)*3-9 mutants (Fig. 6A), indicating that cross-talk between H3-K9 and H4-K20 trimethylation systems is evolutionarily conserved. In *Su(var)*2-5 (HP1) mutants, H3-K9 methylation and pericentric accumula-

Figure 5. Definition of a heterochromatic targeting module in Suv4-20h2. (A) Various truncations of Suv4-20h2 [amino acids 1–387] and Suv4-20h2 [amino acids 1–280] were expressed in HMTase assays with recombinant histone H4 [rH4], recombinant histone octamers [rOct], recombinant nucleosomes [rNuc], and recombinant nucleosomes containing a point mutation in H4 [K20A]. The top panel shows HMTase activity, the bottom panel shows coomassie staining of the membrane used for the HMTase assay. (B) Reaction products of HMTase assays, as in A, were Western blotted and probed with α-monomethyl, α-dimethyl, and α-trimethyl H4-K20 antibodies (top panels, antibody; bottom panels, Ponceau staining). As a control, recombinant and native nucleosomes that were not processed by the enzyme were loaded (asterisks indicate slight cross-reactivities, observed with long exposure times).
tion of SU(VAR)3-9 remains unaltered (Schotta et al. 2002). Intriguingly, H4-K20 trimethylation is nearly lost from polytene chromatin in these mutants (Fig. 6A). This result indicates that the induction of heterochromatic histone lysine methylation marks occurs in a sequential pathway. At pericentric heterochromatin Su-(var)3-9-mediated H3-K9 methylation would precede H4-K20 trimethylation by the Suv4-20 enzymes. Other H3-K9-specific HMTases could direct HP1 and subsequently H4-K20 trimethylation to euchromatic bands.

**Drosophila Suv4-20 is a dominant suppressor of position-effect variegation**

Our data suggest H4-K20 trimethylation to be a mark of silenced chromatin domains. We therefore investigated whether this modification would indeed be important for gene silencing in well-described PEV models in *Drosophila* [Reuter and Spierer 1992]. A single, homozygous viable P-element insertion (P[GT1]BG00814) into the third exon of *Suv4-20* has been identified in the course of the *Drosophila* gene disruption project (Flybase). H4-K20 trimethylation at polytene chromatin is nearly lost in homozygous mutant larvae (Fig. 6A), demonstrating that the P-element insertion (Suv4-20 BG00814) represents a strong hypomorphic allele of *Suv4-20*. Because the *Suv4-20* locus maps on the X chromosome, the classical PEV rearrangement In(1)wm4 [Reuter and Spierer 1992] cannot be used to analyze a potential modifier effect of *Suv4-20*. Therefore, we analyzed another PEV rearrangement that translocates a different marker, *Stubble* (Sb), close to pericentric heterochromatin [T(2;3)SbV; Sinclair et al. 1983]. The dominant mutation *Stubble* induces short bristles, but heterochromatin-induced silencing of *Sb* results in wild-type (long) bristles [Fig. 6B]. Homozygous *Suv4-20* BG00814 as well as control wild-type females were crossed to T(2;3)SbV males. In the progeny, the extent of Sb reactivation was determined as the ratio of short bristles [active Sb] to long bristles [inactive Sb]. In males and females of the wild-type crosses, only 1%–2% of bristles show a Sb phenotype, indicating that SbV is largely inactivated [Fig. 6B]. In contrast, SbV becomes derepressed in the progeny of *Suv4-20* BG00814 flies, because now ~25% of the bristles are short [Fig. 6B]. This result classifies *Suv4-20* as a dominant PEV modifier and further indicates a functional role for *Suv4-20*-dependent H4-K20 trimethylation in gene silencing.

**Discussion**

A combinatorial histone methyl code for constitutive heterochromatin

One of the predictions of the “histone code” hypothesis suggests that distinct combinations of histone modifications participate in the formation of different chromatin domains [Strahl and Allis 2000; Turner 2000; Jenuwein and Allis 2001]. Of the five lysine positions in the histone N termini that display prominent methylation (Peters et al. 2003), H3-K9, H3-K27, and H4-K20 are supposed to have repressive functions [Fischle et al. 2003; Lachner et al. 2003; Vaquero et al. 2003]. A combination of H3-K27 trimethylation together with H3-K9 dimethylation and H4-K20 monomethylation has been described to index the inactive X chromosome [Silva et al.
Prominent marks of constitutive heterochromatin are H3-K9 trimethylation and H3-K27 monomethylation [Peters et al. 2003, Rice et al. 2003]. However, the contribution of the three distinct H4-K20 methylation states in the formation of chromosomal subdomains has not been analyzed. Previous studies indicated H4-K20 methylation is broadly enriched at mitotic chromosomes [Fang et al. 2002, Nishioka et al. 2002a, Rice et al. 2002]. By using highly selective antibodies, developed by UBI, which discriminate H4-K20 mono-, di-, and trimethylation, we now describe H4-K20 trimethylation as a third component of repressive histone lysine methyl marks in pericentric heterochromatin.

Based on our novel data, we propose the following model for sequential induction of H3-K9 and H4-K20 trimethylation at constitutive heterochromatin in mammals [Fig. 7]. In a first step, Suv39h enzymes would be targeted to repeat-rich sequences at pericentric heterochromatin, presumably by interacting with protein complexes containing components of the RNAi machinery [Hall et al. 2002; Jenuwein 2002; Volpe et al. 2002; Schramke and Allshire 2003; Pal-Bhadra et al. 2004; Verdel et al. 2004]. Because pericentric heterochromatin is also enriched for H3-K27 monomethylation, and the Suv39h enzymes prefer an H3-K9 monomethylated substrate [Peters et al. 2003], this first step most likely also requires the activity of currently unknown H3-K27 and H3-K9 monomethylases. Following stabilized binding of HP1α and HP1β to H3-K9 trimethylated nucleosomes [Bannister et al. 2001, Lachner et al. 2001], HP1 molecules, probably in conjunction with nucleosomal surfaces, would then recruit the Suv4-20h enzymes, which in turn trimethylate H4-K20. This tentative model would involve collaboration of at least four distinct HMTases to induce the observed combination of H3-K9 trimethyl, H4-K20 trimethyl, and H3-K27 monomethyl marks at pericentric heterochromatin. Future studies are aimed to identify whether HP1 binding is stabilized by a combination of H3-K9 and H4-K20 trimethylation or if other heterochromatin-specific components may exist that could selectively be recruited by this combinatorial histone lysine methylation pattern.

**Suv4-20h enzymes are nucleosomal-specific HMTases**

Recombinant Suv4-20h HMTases possess selective activity toward the H4-K20 position in a nucleosomal context. Recombinant histone H4 is not used as a substrate, and only poor activity is detected with histone octamers [Fig. 4]. In contrast, Suv39h enzymes preferably methylate free histone H3 and recombinant octamers, whereas nucleosomes are only very poor in vitro substrates [data not shown]. Thus, selective generation of H3-K9 and H4-K20 methylation marks could occur in a stepwise mechanism [Vermaak et al. 2003], in which histone H3–H4 dimers [Tagami et al. 2004] or tetramers might be methylated by Suv39h and, after nucleosome assembly, Suv4-20h enzymes would trimethylate the H4-K20 position.

In vitro, Suv4-20h enzymes can directly trimethylate unmodified H4-K20 positions [Fig. 4], without requirement for a monomethyl substrate. This is in contrast to Suv39h function, which prefers an H3-K9 monomethylated position [Peters et al. 2003]. In Suv39h dn cells, H3-K9 monomethylation becomes enriched at pericentric heterochromatin [Peters et al. 2003]; however, we did not detect an increase for H4-K20 monomethylation in Suv4-20h RNAi knockdown experiments [data not shown]. In addition, H3-K9 trimethylation is not required for Suv4-20h activity in vitro [Fig. 4], although cross-tail interactions between nucleosomal H3-K9 and H4-K20 positions may exist in vivo that could affect enzymatic activity of Suv4-20h HMTases.

PR/SET7 is another HMTase with selective activity toward nucleosomal H4-K20 [Nishioka et al. 2002a, Fang et al. 2002], but in contrast to Suv4-20h enzymes, it appears to be an exclusively monomethylating enzyme [K. Sarma, D. Reinberg, and S. Gamblin, unpubl.]. Furthermore, PR/SET7 fails to localize to pericentric heterochromatin when expressed as an EGFP fusion protein [data not shown]. N- and C-terminal extensions of the SET domain, such as pre- and post-SET domains of Suv39h enzymes, are involved in substrate binding and in conferring enzyme selectivity toward mono-, di-, and trimethylation [Xiao et al. 2003]. Suv4-20h enzymes contain extensions to the SET domain that are distinct from pre- and post-SET domains, and that are only conserved.
H4-K20 trimethylation is a novel hallmark of pericentric pathway between H3-K9 and H4-K20 methylation. Indeed, the two H4 tails of one nucleosome are presented in different molecular projections, such that one H4 tail is directed toward the DNA backbone, whereas amino acids 16–25 of the second histone H4 molecule contact the H2A/H2B dimer of the adjacent nucleosome (Luger et al. 1997). It is according to these structural definitions that the H4 tail may be particularly suited for directing transitions between 11- and 30-nm chromatin fibers, and H4-K20 trimethylation appears as a good candidate mark to contribute in facilitating these higher order nucleosomal organizations.

H4-K20 trimethylation is evolutionarily conserved

In mammals, ~50 SET domain genes have been identified (Lachner and Jenuwein 2002; Kouzarides 2002), of which a subset is evolutionarily conserved in *Drosophila* and *S. pombe*. Intriguingly, nearly all putative SET domain HMTases that target repressive lysine positions (H3-K9, H3-K27, H4-K20) are absent in *S. cerevisiae*. In agreement, H3-K9, H3-K27, and H4-K20 methylation cannot be detected in *S. cerevisiae* (R. Sengupta and T. Jenuwein, unpubl.), indicating that this organism largely lacks extensive heterochromatic domains and uses other repressive systems (e.g., SIR proteins) to silence gene activity (Kurdistan and Grunstein 2003).

H3-K9 methylation has emerged as an important and evolutionarily conserved mark to silence gene activity at repressive chromatin domains. The concerted function of Suv39h and Suv4-20h enzymes described here raises the question as to the functional role of H4-K20 methylation in well-described epigenetic paradigms. Homologs of the *Suv39h/Suv4-20h* methylation systems are also conserved in *Drosophila* (Su(var)3-9 and Su(var)20) and *S. pombe* (Clr4 and SET9). Surprisingly, H3-K9 methylation appears severely underrepresented (and can hardly be detected) in bulk histone preparations from *S. pombe* extracts, which instead display robust signals for all three H4-K20 methylation states (R. Sengupta and T. Jenuwein, unpubl.). Because HP1 is also conserved in *S. pombe* [as the Swi6 protein], a functional synergy between H3-K9 and H4-K20 methylation systems might exist in *S. pombe* that could be operative to safeguard gene silencing and/or to index distinct chromosomal regions. Together, our definition of novel H4-K20 trimethylating HMTases and the presence of H4-K20 trimethylation from *S. pombe* to mammals provides evidence for an evolutionary conservation of the proposed silencing pathway between H3-K9 and H4-K20 methylation.

Su4-20h enzymes are novel Su(var) genes

H4-K20 trimethylation is a novel hallmark of pericentric heterochromatin in mammals, but the current lack of mutants for Su4-20h enzymes prevents the analysis of its regulatory function during mouse development. However, the identification of a P-element insertion in *Drosophila Su4-20*, which results in nearly complete loss of H4-K20 trimethylation [Fig. 6A], allowed a first functional analysis. In *Drosophila*, there are >50 *Su(var)* loci (Reuter and Spierer 1992, Schotta et al. 2003), which encode products involved in heterochromatin-induced gene silencing. Only ~15 of these *Su(var)* genes have been identified to date, the most revealing of which include *Su(var)2-5* (HP1; Eisenberg et al. 1990), *Su(var)3-7* (Cleard et al. 1997), and *Su(var)3-9* (Tschiersch et al. 1994). Here we show that *Drosophila Su4-20* is a novel, dominant suppressor of position-effect variegation [Fig. 6B]. The *Su(var) effect of Su4-20 is about as strong as that of *Su(var)3-9* [data not shown], suggesting that H3-K9 and H4-K20 trimethylation may have a similar potential to establish heterochromatin-induced gene silencing. In addition, both the hypomorphic *Su4-20* and the *Su(var)3-9* null mutants are homozygous viable. Based on the observed synergy, it is conceivable that the combined disruption of *Su4-20* and *Su(var)3-9* genes might have a much more pronounced phenotype that could compromise establishment and stability of heterochromatin.

Major functions of heterochromatin include to safeguard a specialized chromatin structure around centromeres [Karpen and Allshire 1997; Bernard and Allshire 2002] and to protect genome stability by silencing of transposable elements [Birchler et al. 2000; Pal-Bhadra et al. 2004]. The pericentric accumulation and the silencing defect observed in *Su4-20* mutants predict H4-K20 trimethylation to participate in these mechanisms, but H4-K20 trimethylation may also be involved in gene repression at euchromatic targets or other functions of chromatin control, in which it could even be independent of Su39h activity. As histone lysine methylation is implicated in almost all epigenetic paradigms, and H4-K20 trimethylation has been shown to be enriched in aged organs (Sarg et al. 2002), future analyses into the physiological roles of H4-K20 methylation promise not only to provide a better understanding of the interplay between distinct histone lysine methylation systems but may even offer novel insights into fundamental biological processes, such as normal and perturbed development, aging, and regeneration.

Materials and methods

**Plasmid construction**

Full-length open reading frames of all candidate SET domain proteins as well as protein truncations of Su4-20h1 and Su4-20h2 were PCR amplified from IMAGE cDNA clones (purchased from Deutsches Ressourcenzentrum für Genomforschung GmbH [RZPD] and cloned into the GATEWAY entry vector pDONR207 [Invitrogen]). A list of clones used as well as primer sequences is available on request. The reference sequences of Su4-20h1 and Su4-20h2 have been deposited at GenBank [AY555192, AY555193].
C-terminal associated regions (Fig. 3A) were expressed in Suv4-20h2 (1–280) containing the SET domain and N- as well as GST fusion proteins of Suv4-20h1 (amino acids 1–387) and underrepresented in the transfection mix, cells showing EGFP as described (Peters et al. 2003). Because pEGFP-N1 vector was cells. Five days after transfection, cells were transferred on selection, medium containing 600 µg/mL G418 was applied to the using Lipofectamine (Invitrogen). One day after transfection set containing the 19mer targeting sequences and 0.2 µg pEGFP-N1 were transfected with a mixture of 2 µg pSUPER plasmids con- MEFs were plated on a 10-cm dish 1 d before transfection. Cells were trans- trichlamysynthesis, trimethylation (Peters et al. 2003). Transient transfection and immunofluorescence on interphase cells

Candidate SET domain proteins were expressed in female wild- and Suv39h dn cells. In each chamber of a multwell slide, 5000 cells were plated 1 d before transfection. Cells were trans- fected with 1 µg plasmid DNA by using the Lipofectamine proto- col (Invitrogen). One day posttransfection, cells were fixed in 2% PFA/PBS, extracted with 0.1% sodium citrate/0.1% Triton X-100, washed with PBS, and mounted with Vectashield [Vector Laboratories]. EGFP fluorescence was analyzed by fluorescence microscopy. Indirect immunofluorescence was performed according to the method of Peters et al. (2003), using antibodies against H4-K20 mono-, di-, and trimethylation [Upstate Biotech; nos. 07-440, 07-367, 07-463] and H3-K9 mono-, di-, and trimethylation [Peters et al. 2003].

RNAi knock-down

DNA oligos containing the 19mer targeting sequence, a loop region, and HindIII/BglII restriction sites were cloned into pSUPER [Brummelkamp et al. 2002]. For each target gene, two different 19mer sequences were used. Fifty thousand wild-type MEFs were plated on a 10-cm dish 1 d before transfection. Cells were transfected with a mixture of 2 µg pSUPER plasmids containing the 19mer targeting sequences and 0.2 µg pEGFP-N1 using Lipofectamine (Invitrogen). One day after transfection selection, medium containing 600 µg/mL G418 was applied to the cells. Five days after transfection, cells were transferred on chamber slides and analyzed by indirect immunofluorescence as described [Peters et al. 2003]. Because pEGFP-N1 vector was underrepresented in the transfection mix, cells showing EGFP fluorescence were assumed to also contain pSUPER vector.

HMTase assays

GST fusion proteins of Suv4-20h1 [amino acids 1–387] and Suv4-20h2 [1–280] containing the SET domain and N- as well as C-terminal associated regions (Fig. 3A) were expressed in Esch- erichia coli. HMTase assays were performed as described [Nish- ioka et al. 2002b]. In particular, samples of the Suv4-20h enzymes were incubate for 1 h at 30°C in reaction buffer (50 mM Tris-HCl at pH 8.5, 5 mM MgCl$_2$, 4 mM DTT) with 1 µM 3H-labeled SAM [Amerham Pharmacia Biotech]. Fifty nanograms of recombinant histone H4 and 2 µg of recombinant octamers/nucleosomes were used as substrates.

In vitro binding assays

GST fusion proteins of Suv4-20h2 were expressed in E. coli and purified on glutathione-S-Sepharose [Pharmacia]. Equal amounts of GST fusion proteins were incubated with in vitro translated myc$_3$-tagged HP1 isoforms [HP1a, amino acids 5–191; HP1B, amino acids 5–185; HP1γ, amino acids 5–173; Lachner et al. 2001] in IP buffer [50 mM Tris-HCl at pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.5% NP-40, 10% glycerol, proteinase inhibitor cocktail]. Incubation was carried out overnight at 4°C with gentle agitation. The beads were washed six times with IP buffer and finally resuspended in 50 µL loading buffer. Bound proteins were separated on SDS-polyacrylamidgels and detected by immunoblotting using α-myc antibody [Upstate Biotech, 05-419].

Immunostaining of polytene chromosomes

Preparation of polytene chromosomes was performed as described [Schotta et. al. 2002] with the following modifications: salivary glands of third instar larvae were dissected in 0.7% NaCl, fixed in 4% formaldehyde and squashed in 45% acetic acid/2% formaldehyde. Chromosomes were incubated with α-trimethyl H3-K9 [Peters et al. 2003] as well as α-trimethyl H4-K20 antibodies overnight at 4°C, followed by incubation with secondary alexA 488-conjugated antibodies [Molecular Probes] for 2 h at 37°C. DNA was stained with propidium iodide [Molecular Probes]. Preparations were mounted in Vectashield medium and examined with a confocal laser scanning micro- scope [LSM510, Zeiss].

Phenotypic characterization of Stubble variegation

The P-element P[GT1]BG0814 is inserted into the third exon of Suv4-20 [CG13363] and results in a strong hypomorphic al- allele (Suv4-20PG0814). The PEV modifier effect of Suv4-20PG0814 was evaluated by using T(2;3)Sh$^v$ [Sinclair et al. 1983]. This translocation brings the dominant mutation Sh$^v$ [short-bristles phenotype] near to chromosome 2R heterochromatin. Because Sh$^v$ is dominant, Su(var)s increase the stubble phenotype. Suv4-20PG0814/Suv4-20PG0814 as well as control wild-type females were crossed to T(2;3)Sh$^v$, In(3R)Mo, Sb$^v$, trx$^T$/TM3, males, and 14 defined bristles were scored as being wild type or Sh. The ratio between the total numbers of Stubble to wild-type bristles is a mark for the extent of Sb inactivation and has been evaluated for >25 flies of each genotype.

Bioinformatic methods

SET domain proteins in mouse were identified by using a hid- den Markov model [HMM] profile for the SET domain [Pfam, Pf00856]. IPI database [http://www.ebi.ac.uk/IPI] was searched by using HMMER software [http://hmmer.wustl.edu]. An align- ment of all SET domains and the neighbor-joining tree was cre- ated by using the t-coffee alignment program [Notredame et al. 2000]. Orthologs of mouse SET domain proteins in S. pombe and Drosophila were identified by using protein blast searches [http://www.ncbi.nlm.nih.gov/blast].

Acknowledgments

We are indebted to UBI for exchange of H4-K20 antibodies and thank Jim Bone, Mary-Ann Jelinek, and Judy Nisson for helpful advice. We also acknowledge Susanne Opravil and Monika Kauer for detailed characterization of the H4-K20 antibodies. We are further grateful to Alexander Schleiffer and Anton Beyer for their expert help in sequence comparison and homology searches of mouse SET domain genes. This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) to G.R. Studies conducted in the laboratory of D.R. were supported by a grant from the National Institutes of Health (GM37120) and the Howard Hughes Medical Institute. Research in the labo- ratory of T.J. is supported by the IMP through Boehringer Ingel-heim, and by grants from the Vienna Economy Promotion Fund (WWF), the European Union [EU-network HPRN-CT 2000-
References


Rea, S., Eisenhaber, F., O’Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtcher, K., Ponting, C.P., Allis, ...
Schotta et al.


A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin

Gunnar Schotta, Monika Lachner, Kavitha Sarma, et al.

Genes Dev. 2004, 18:
Access the most recent version at doi:10.1101/gad.300704