Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability

Nan Liu,1,2,3,6 Zhuqiang Zhang,3,6 Hui Wu,4,6 Yonghua Jiang,2 Lingjun Meng,2 Jun Xiong,2 Zuodong Zhao,2 Xiaohua Zhou,2 Jia Li,2 Hong Li,2 Yong Zheng,2 She Chen,2 Tao Cai,2 Shaorong Gao,5 and Bing Zhu2,3

1College of Life Sciences, Beijing Normal University, Beijing 100875, China; 2National Institute of Biological Sciences, Beijing 102206, China; 3National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; 4National Engineering Laboratory for AIDS Vaccine, School of Life Sciences, Jilin University, Changchun, Jilin Province 130012, China; 5School of Life Sciences and Technology, Tongji University, Shanghai 200092, China

GLP and G9a are major H3K9 dimethylases and are essential for mouse early embryonic development. GLP and G9a both harbor ankyrin repeat domains that are capable of binding H3K9 methylation. However, the functional significance of their recognition of H3K9 methylation is unknown. Here, we report that the histone methyltransferase activities of GLP and G9a are stimulated by neighboring nucleosomes that are premethylated at H3K9. These stimulation events function in cis and are dependent on the H3K9 methylation binding activities of ankyrin repeat domains of GLP and G9a. Disruption of the H3K9 methylation-binding activity of GLP in mice causes growth retardation of embryos, ossification defects of calvaria, and postnatal lethality due to starvation of the pups. In mouse embryonic stem cells (ESCs) harboring a mutant GLP that lacks H3K9me1-binding activity, critical pluripotent genes, including Oct4 and Nanog, display inefficient establishment of H3K9me2 and delayed gene silencing during differentiation. Collectively, our study reveals a new activation mechanism for GLP and G9a that plays an important role in ESC differentiation and mouse viability.

Keywords: GLP; G9a; H3K9 methylation; gene repression; cell differentiation

Supplemental material is available for this article.

Received October 17, 2014; revised version accepted January 6, 2015.
third of the genome and undergoes drastic changes during cell fate transitions [Wen et al. 2009; Lienert et al. 2011; Chen et al. 2012]. However, GLP and G9a exhibit very low activities toward nucleosomal H3 in vitro [Tachibana et al. 2005]. This leads to questions of whether any activation mechanisms are involved in efficiently establishing H3K9me2 during cell fate transitions and whether the enzymatic activities of GLP and G9a are also regulated by chromatin environment.

Intriguingly, in addition to the catalytic activities for methylating H3K9, GLP and G9a also carry the ankyrin repeat domains that specifically recognize H3K9me1 and H3K9me2, respectively [Brent and Marmorstein 2008; Collins et al. 2008]. We reason that GLP and G9a may use their H3K9 methylation-binding activities to sense methylated H3K9 and achieve their full activity as well as maintain H3K9me2 levels during mitotic divisions or efficiently establish H3K9me2 at genomic regions that are to be silenced during cell fate transition events.

Using oligonucleosome substrates containing unmodified nucleosomes and nucleosomes premethylated at the H3K9 position to various states, we observed strong stimulation of the enzymatic activities of GLP and G9a by neighboring nucleosomes containing H3K9me1 and H3K9me2, respectively. The binding abilities to H3K9me1 and H3K9me2 of GLP and G9a were required for the enzymatic stimulation. ESCs harboring the H3K9me1-binding mutant form of GLP failed to efficiently establish H3K9me2 during retinoic acid (RA)-induced differentiation across a large number of genes, and a subset of these genes displayed delayed transcription repression. Notably, several genes encoding critical pluripotent factors, including Oct4 and Nanog, were among them, suggesting that GLP's H3K9 methylation-binding activity is required for efficient establishment of H3K9me2 and rapid shutdown of the pluripotency transcription program in ESCs undergoing differentiation. Homozygous knock-in mice carrying a H3K9me1-binding mutant form of GLP exhibited postnatal lethality. These mice displayed embryonic growth retardation and defects in calvaria bone formation. Together, these results provided evidences for the significance of GLP's H3K9me1-binding activity in efficient H3K9me2 establishment and gene silencing during cell differentiation and mouse viability but not in mitotic inheritance of H3K9 methylation.

Results

Histone methyltransferase activities of GLP and G9a are stimulated by neighboring nucleosomes premethylated at H3K9 in an ankyrin repeat-dependent manner

To test whether GLP and G9a can be stimulated by neighboring nucleosomes premethylated at H3K9, we developed a histone methyltransferase activity stimulation assay. We prepared premethylated histones using chemical reactions to install methyl-lysine analogs [MLAs] at the K9 position of histone H3 [Simon et al. 2007]. The MLAs installed by MLA reactions preserve most of the features of the methyl-lysine, including correct recognition by specific methyl-lysine-binding proteins [Simon et al. 2007] and compatibility in serving as substrates for histone methyltransferases and demethylases [Jia et al. 2009]. Histone octamers that contained H3 premethylated at the K9 position with various modifications [Kc9me0/1/2/3 or unreacted K9C] and histone octamers that contained unmodified H3 were assembled separately and then pooled together at a 1:1 ratio. The octamers were assembled with DNA into oligonucleosomes. The assembled oligonucleosomes were purified by gel filtration on a Sepharose CL-4B column to remove free histone octamers. The final substrates contained unmodified nucleosomes that were flanked by nucleosomes harboring premethylated H3K9 [Fig. 1A]. To distinguish the two types of histone H3, the premethylated H3 histones were tagged and therefore migrated slower on SDS-PAGE gels [Fig. 1B].

Recombinant GLP and G9a proteins covering ankyrin repeats and SET domains were purified from Escherichia coli and assayed for methyltransferase activity using the above-described substrates. Strikingly, despite the fact that equal amounts of enzymes and unmodified H3 histones were present in the reactions, G9a displayed 20-fold more activity on the unmodified nucleosomes that were flanked by H3Kc9me2-containing nucleosomes than on unmodified nucleosomes that were flanked by nucleosomes containing H3K9C, H3Kc9me0, or H3Kc9me3 [Supplemental Fig. 1A]. These results are consistent with the fact that H3K9me2 is G9a's favored binding partner [Collins et al. 2008]. Similarly, GLP displayed the best activity on unmodified nucleosomes flanked by nucleosomes containing H3Kc9me1 [Fig. 1B], which was in agreement with its preferred binding to H3K9me1 [Collins et al. 2008]. Catalytic mutants of GLP [C1201A] or G9a [C1007A] displayed no detectable activity on any of the substrates [Fig. 1C; Supplemental Fig. S1B]. Notably, G9a also displayed much higher enzymatic activity at the tagged premethylated H3 histones in the case of H3Kc9me2 [Supplemental Fig. 1A, lane 4]. Similar observations were also obtained for GLP at the tagged premethylated H3 histones in the case of H3Kc9me1 [Fig. 1B, lane 3]. This is most likely due to G9a's and GLP's in vitro catalytic activities at H3K27, as previously reported [Tachibana et al. 2001; Wu et al. 2011]. Indeed, when residue Lys27 of H3 was mutated to alanine, the stimulation activity was observed at only the unmodified H3 histones but not the tagged premethylated H3 histones [Supplemental Fig. 1C, lane 2]. Furthermore, neither GLP nor G9a activity was stimulated by nucleosomes premethylated at H3K27 [Fig. 1D, Supplemental Fig. 1D], in agreement with their inability to recognize H3K27 methylation [Collins et al. 2008]. Together, these results suggested that the activities of GLP and G9a on nucleosomal histones could be strongly stimulated when their substrates were flanked by neighboring nucleosomes harboring H3K9 premethylation.

Ankyrin repeats of GLP and G9a were responsible for their recognition toward H3K9 methylation [Collins et al. 2008]. GLP W877A/W882A/E885A and G9a W791A/W796A/E799A mutants [abbreviated as GLP 3A and G9a 3A mutants, respectively] lost methylated H3K9-binding activity due to the disruption of their aromatic cages and hydrogen bonds within the ankyrin repeats that confine
the methylated lysine [Collins et al. 2008], which was confirmed by isothermal titration calorimetry (ITC) [Supplemental Fig. 2]. Neither GLP 3A (Fig. 1E) nor G9a 3A (Supplemental Fig. 1E) was stimulated by premethylated neighboring nucleosomes, indicating that recognition of premethylated neighboring nucleosomes by ankyrin repeats is an essential step for stimulation.

The stimulation of GLP and G9a activities by premethylated nucleosome functions in cis

A second set of substrates was assembled to allow simultaneous testing of stimulation events at nucleosomes existing in cis and in trans. Oligonucleosomes containing Flag-H3 were mixed with the substrates described in Figure 1A at a 1:1 ratio [Fig. 2A]. In the final substrate mixture, the unmodified H3 histones existed in cis with the premethylated H3 histones, while the Flag-H3 histones existed in trans with the premethylated H3 histones [Fig. 2A]. These three types of H3 histones could be distinguished in SDS-PAGE gels due to their different migration velocities [Fig. 2B]. GLP [Fig. 2B] and G9a [Supplemental Fig. 3A] displayed robust activity on unmodified H3 histones existing in cis with nucleosomes containing H3K9me1 and H3K9me2, respectively, but not on Flag-H3 nucleosomes existing in trans with the premethylated nucleosomes.
To exclude that the Flag tag might affect the enzymatic activities, control experiments were performed using substrates in which Flag-H3 histones were placed in cis with the pre-modified His-H3 histones, and untagged H3 histones were placed in trans with the pre-modified His-H3 histones. GLP and G9a displayed robust activities on Flag-H3 nucleosomes existing in cis with the premethylated H3 and minimal activities on H3 nucleosomes existing in trans with the premethylated H3 (Supplemental Fig. 3B,C). These results indicated that the stimulation events were dependent on spatial position of the nucleosomes and functioned in cis.

H3K27 methyltransferase (PRC2) can be allosterically activated by peptides containing H3K27me3 (Margueron et al. 2009). However, neither GLP nor G9a was activated by the addition of methylated H3K9 peptides (Fig. 2C; Supplemental Fig. 3D), which suggests that GLP and G9a adapted distinct strategies for spreading chromatin modifications different from the allosteric activation approach that PRC2 uses. Taken together, GLP and G9a were able to sense H3K9me1/2 through their ankyrin repeats and facilitate their H3K9 methylation activities in vitro.

GLP 3A mutant mice were postnatal-lethal with growth retardation and craniofacial defects

To understand the in vivo role of the H3K9 methylation recognition activities of GLP and G9a, we generated knock-in mice bearing the GLP 3A or G9a 3A mutations. Mutations were introduced into R1 ESCs by homologous recombination (Supplemental Fig. 4). We confirmed that the targeting events did not affect the expression level of the targeted GLP and G9a alleles in the corresponding heterozygote ESC lines (Supplemental Fig. 4). F1 heterozygous mice of the GLP 3A and G9a 3A mutants (termed GLP+/M and G9a+/M) were healthy and fertile. Surprisingly, homozygous mice (GLPM/M and G9aM/M) displayed strikingly different phenotypes. G9aM/M mice were born at the Mendelian ratio (Supplemental Fig. 5A) and were healthy and fertile. In contrast, GLPM/M mice were mostly postnatal-lethal (Fig. 3A). All GLPM/M mice were born smaller in size (Fig. 3B), and 120 out of 128 of them died within 2 d of birth (Fig. 3A). Only eight GLPM/M mice survived to adulthood, and they suffered from severe growth retardation within the first 4 wk after birth (Supplemental Fig. 5B). GLP and G9a double-mutant mice displayed a phenotype similar to that of GLPM/M mice (Fig. 3C), suggesting that the H3K9 methylation recognition activity of GLP played a dominant role in vivo. Therefore, we focused our study on GLPM/M in all subsequent experiments.

Consistent with the smaller size of GLPM/M newborns, the GLPM/M embryos also showed growth retardation as early as embryonic day 10.5 (E10.5) (Supplemental Fig. 5C). Although no drastic abnormalities of the main organs were observed and no catastrophic events occurred in breathing
and heart beating, we consistently observed that no milk was present in the stomachs of GLPM/M pups, so it was highly possible that GLPM/M mice died of starvation. To understand at which stage the GLPM/M pups failed to be fed, we adapted an assay to dissect multiple steps of suckling processes (Hongo et al. 2000). The GLPM/M mice were capable of actively searching and licking the nipples of mother mice (Fig. 3D), suggesting that their ollaction and sensory–motor systems were probably fine. However, these pups were unable to remain stably attached to the nipples; while most wild-type and heterozygote pups can stably attached to the nipples for up to 20 min, the GLPM/M pups failed to remain attached for even just 1 min in ~70% of the attempts [Fig. 3E], which suggested that they might suffer from craniofacial development abnormalities (Turgeon and Meloche 2009). Consistent with this hypothesis, skeletal preparations at various developmental stages showed obvious delayed calvaria and nasal bone ossification in GLPM/M mice [Fig. 3F, Supplemental Fig. 5D]. The most severe delayed ossification occurred in the sagittal and metopic sutures as well as in the nasal bones [Fig. 3F]. Even for the few surviving adult GLPM/M mice, bone loss was observed in their metopic sutures [Fig. 3G; Supplemental Fig. 5E], which was likely caused by ossification defects during embryonic development. In addition, several of the surviving adult GLPM/M mice [two out of eight] displayed an obviously bent nose and malocclusion [Supplemental Fig. 5E,F].

The H3K9me1-binding activity of GLP is largely dispensable for maintenance of global H3K9me2 levels in ESCs and tissues

To study the in vivo role of H3K9 methylation-binding activity of GLP at molecular level, we derived wild-type and GLP 3A mutant ESCs [termed wild-type and GLP 3A ESCs, respectively] from littermate embryos at the blastocyst stage. These ESCs displayed the correct karyotypes [Supplemental Fig. 6A] and similar expression levels of GLP [Supplemental Fig. 6B]. Western analysis revealed comparable bulk H3K9me2 levels between wild-type and GLP 3A ESCs [Supplemental Fig. 6C]. Next, we analyzed brain and liver samples harvested from multiple littermate pairs of wild-type and GLPM/M newborn pups [Supplemental Fig. 6D]. We did not observe consistent changes of H3K9me2 between wild-type and GLPM/M livers or brains [Supplemental Fig. 6D]. Considering that Western analysis is not an ideal quantitative assay, especially for subtle changes, we performed H3K9me2 ChIP-seq [chromatin immunoprecipitation [ChIP] combined with deep sequencing] experiments using ESCs and livers from newborn pups with wild-type and GLPM/M mutant backgrounds. We first divided the mouse genome into 10-kb windows. This window size was chosen because H3K9me2 often exist in large blocks of chromatin regions termed LOCKs, and the minimal size of LOCKs is ~20 kb [Wen et al. 2009]. We plotted the H3K9me2 fold enrichment changes between mutant and wild-type samples for all of the 10-kb windows. In ESCs and livers from newborn pups, only 1% and 2.2% of the windows displayed a reduction of H3K9me2 in mutant newborn GLPM/M mice, bone loss was observed in their metopic sutures [Fig. 3F; Supplemental Fig. 5E], which was likely caused by ossification defects during embryonic development. In addition, several of the surviving adult GLPM/M mice [two out of eight] displayed an obviously bent nose and malocclusion [Supplemental Fig. 5E,F].

The H3K9me1-binding activity of GLP is largely dispensable for maintenance of global H3K9me2 levels in ESCs and tissues

To study the in vivo role of H3K9 methylation-binding activity of GLP at molecular level, we derived wild-type and GLP 3A mutant ESCs [termed wild-type and GLP 3A ESCs, respectively] from littermate embryos at the blastocyst stage. These ESCs displayed the correct karyotypes [Supplemental Fig. 6A] and similar expression levels of GLP [Supplemental Fig. 6B]. Western analysis revealed comparable bulk H3K9me2 levels between wild-type and GLP 3A ESCs [Supplemental Fig. 6C]. Next, we analyzed brain and liver samples harvested from multiple littermate pairs of wild-type and GLPM/M newborn pups [Supplemental Fig. 6D]. We did not observe consistent changes of H3K9me2 between wild-type and GLPM/M livers or brains [Supplemental Fig. 6D]. Considering that Western analysis is not an ideal quantitative assay, especially for subtle changes, we performed H3K9me2 ChIP-seq [chromatin immunoprecipitation [ChIP] combined with deep sequencing] experiments using ESCs and livers from newborn pups with wild-type and GLPM/M mutant backgrounds. We first divided the mouse genome into 10-kb windows. This window size was chosen because H3K9me2 often exist in large blocks of chromatin regions termed LOCKs, and the minimal size of LOCKs is ~20 kb [Wen et al. 2009]. We plotted the H3K9me2 fold enrichment changes between mutant and wild-type samples for all of the 10-kb windows. In ESCs and livers from newborn pups, only 1% and 2.2% of the windows displayed a reduction of H3K9me2 in mutant newborn GLPM/M mice, bone loss was observed in their metopic sutures [Fig. 3F; Supplemental Fig. 5E], which was likely caused by ossification defects during embryonic development. In addition, several of the surviving adult GLPM/M mice [two out of eight] displayed an obviously bent nose and malocclusion [Supplemental Fig. 5E,F].

The H3K9me1-binding activity of GLP is critical for efficient establishment of H3K9me2 and gene silencing at the Oct4 gene during ESC differentiation

Considering that the H3K9me1-binding activity of GLP stimulated its catalytic activity in vitro [Fig. 1], it displayed...
such a difference was not observed at the lower than that in the wild-type cells (Fig. 4B). In contrast, expression was delayed for Oct4 gene, and complete silencing of transcription factor association at the treatment, which most likely reflects changes of transcription levels were significantly increased at the Oct4 locus in wild-type cells (Fig. 4B), in agreement with previous reports (Feldman et al. 2006). However, in RA-treated GLP 3A cells, H3K9me2 levels at the Oct4 locus were much lower than that in the wild-type cells (Fig. 4B). In contrast, such a difference was not observed at the Magea2 gene (Fig. 4A,B), which is constitutively repressed by GLP and G9a (Tachibana et al. 2002).

Wild-type and GLP 3A ESCs displayed comparable levels of Oct4 expression [Fig. 4C], and, in both cells, Oct4 expression was quickly reduced within the first 3 d of RA treatment, which most likely reflects changes of transcription factor association at the Oct4 gene. However, GLP 3A cells displayed an obvious delay in shutting down Oct4 expression, and complete silencing of Oct4 transcription was delayed for ~6 d in GLP 3A cells [Fig. 4C], although establishment of DNA methylation at the Oct4 promoter was largely normal in these cells during RA-induced differentiation [Fig. 4D]. To rule out a potential clonal effect of ESCs, the above experiments were repeated in ESCs derived from a separate litter of mice, and similar results were observed [Supplemental Fig. 7A,B].

We also monitored H3K9me2 establishment and Oct4 repression in a related differentiation system involving the formation of embryoid bodies (EBs) from ESCs, which to some extent mimics the in vivo post-implantation embryo development (Desbaillets et al. 2000). In this system, we observed similar defects in H3K9me2 establishment and transcription repression at the Oct4 gene [Supplemental Fig. 7C–E]. These results collectively suggested that the H3K9me1-binding activity of GLP was required for efficient establishment of H3K9me2 and rapid repression of the Oct4 gene during differentiation.

Figure 4. H3K9me2 establishment and transcription silencing of Oct4 during ESC differentiation were delayed in GLP 3A cells. [A,B] ChIP-qPCR results showing H3K9me2 occupancy at the indicated loci in ESCs [A] or RA-treated cells at the indicated time points [B]. [C] RT–PCR results showing Oct4 expression levels during differentiation. [D] DNA methylation levels at the Oct4 promoter during differentiation. Error bars represent the standard error.
only 11% (362 of 3307) of these regions displayed comparable or higher H3K9me2 levels in RA-treated GLP 3A cells with or without RA treatment. The fold enrichment values of three biological replicates were averaged, centered to the mean of rows, and normalized to standard deviation by rows [Z-score in A–C]. (A) Clustering of the 10-kb genomic windows according to their H3K9me2 levels for wild-type (WT) and GLP 3A cells with or without RA treatment. The values are averages of three biological replicates. (B) Cluster dendrogram of the 10-kb genomic windows according to their H3K9me2 levels for all replicates in wild-type and GLP 3A cells with or without RA treatment. (C) Clustering of the 3307 10-kb genomic windows (shown in A) according to their H3K9me1 levels for wild-type and GLP 3A cells with or without RA treatment. The values are averages of three biological replicates. (D) Accumulative distribution of H3K9me2 fold changes between RA-treated cells and ESCs at genic regions. (E) Number of genes with significantly changed H3K9me2 in wild-type or GLP 3A cells after RA treatment. Red dots represent genes with significantly increased H3K9me2 levels after RA treatment in wild-type cells. (F) Box plot of H3K9me2 fold changes between ESCs and RA-treated cells at genes acquired by H3K9me2 (shown in G) in wild-type and GLP 3A cells. (G) H3K9me2 profile at genic regions of genes acquired by H3K9me2 (shown in G) in ESCs or RA-treated wild-type [left] and GLP 3A [right] cells. “Rep 1–3” represents the three biological repeats. Lengths of genes were normalized to 100 percentiles and extended 20% upstream of and downstream from the gene body, respectively.

Figure 5. The H3K9 methylation-binding activity of GLP was required for efficient H3K9me2 establishment during ESC differentiation at a genome-wide level. (A) Clustering of the 10-kb genomic windows according to their H3K9me2 levels for wild-type (WT) and GLP 3A cells with or without RA treatment. The fold enrichment values of three biological replicates were averaged, centered to the mean of rows, and normalized to standard deviation by rows [Z-score in A–C]. (B) Clustering of the 10-kb genomic windows according to their H3K9me2 levels for all replicates in wild-type and GLP 3A cells with or without RA treatment. (C) Clustering of the 3307 10-kb genomic windows (shown in A) according to their H3K9me1 levels for wild-type and GLP 3A cells with or without RA treatment. The values are averages of three biological replicates. (D) Accumulative distribution of H3K9me2 fold changes between RA-treated cells and ESCs at genic regions. (E) Number of genes with significantly changed H3K9me2 in wild-type or GLP 3A cells after RA treatment. Red dots represent genes with significantly increased H3K9me2 levels after RA treatment in wild-type cells. (F) Box plot of H3K9me2 fold changes between ESCs and RA-treated cells at genes acquired by H3K9me2 (shown in G) in wild-type and GLP 3A cells. (G) H3K9me2 profile at genic regions of genes acquired by H3K9me2 (shown in G) in ESCs or RA-treated wild-type [left] and GLP 3A [right] cells. “Rep 1–3” represents the three biological repeats. Lengths of genes were normalized to 100 percentiles and extended 20% upstream of and downstream from the gene body, respectively.

To test the reproducibility of the above analysis, we performed clustering analysis for the same 3307 10-kb windows for all 12 experiments (two cell types × two treatments × three replicates). Essentially, almost all replicates clustered together, indicating good reproducibility among them [Fig. 5B]. The only exception was that one experiment of the GLP 3A cells at the ESC stage was grouped together with the wild-type ESCs, which was in agreement with our observation that wild-type and GLP 3A ESCs were highly similar for their H3K9me2 patterns [Fig. 5B]. These observations were also supported by quantitative results calculating the correlation coefficient between each experiment pair [Supplemental Fig. 8B]. Interestingly, RA-treated GLP 3A cells, which failed to establish H3K9me2 at most of these regions, were clustered together with ESCs but not RA-treated wild-type cells [Fig. 5B], implying that the differentiation process was impaired.
Given that GLP can also generate H3K9me1 products and used them as substrates for further methylation (Tachibana et al. 2005), we performed ChiP-seq experiments for H3K9me1 using the same samples. The same above-mentioned 3307 10-kb genomic regions were first chosen for analysis, and ~66% of them displayed obvious gain of H3K9me1 after RA treatment in wild-type cells (Fig. 5C). Most of these regions displayed correctly established H3K9me1 in GLP 3A cells after RA treatment (Fig. 5C). However, some of the regions that are defective in establishing H3K9me2 during differentiation displayed a reduction of H3K9me1 at the ESC stage of the GLP 3A cells (Fig. 5C). To analyze whether the reduced H3K9me1 is correlated with reduced H3K9me2 in GLP 3A cells before or after RA-induced differentiation, we first selected all 10-kb regions with more than twofold increase of H3K9me2 during RA-induced differentiation in wild-type cells, categorized them into five subgroups according to their H3K9me1 ratio [GLP 3A/wild type] at the ESC stage, and calculated their H3K9me2 ratio [GLP 3A/wild type] before or after RA treatment. Regions with lower H3K9me1 levels tended to display relatively lower H3K9me2 levels in GLP 3A ESCs [Supplemental Fig. S8C]. On the other hand, all of these regions displayed similar degrees of defect in their H3K9me2 levels in GLP 3A cells after RA treatment regardless of their initial levels of H3K9me1 defect at the ESC stage.

We also directly analyzed the H3K9me1 levels in the four samples above. All 10-kb windows with at least one case of more than twofold change between any pair of samples [5642 in total] were selected for clustering analysis. Clearly, the main differences among the four samples resided between ESCs and RA-treated cells but not between wild-type and GLP 3A cells [Supplemental Fig. S8D]. Taken together, the above results collectively support that the primary defect of GLP 3A cells during RA treatment was the efficient establishment of H3K9me2 but not of H3K9me1.

Notably, 89% [2955 of 3307] of the regions with the highest H3K9me2 variation among the four samples were genic regions of 3079 genes [P < 1E-300, binomial test], which was consistent with previous studies reporting that GLP and G9a primarily functioned at genic regions during differentiation of ESCs [Lienert et al. 2011]. Therefore, we focused on the establishment of H3K9 methylation at genic regions for further analysis. Globally, H3K9me2 levels at genic regions were comparable between wild-type and GLP 3A ESCs [Supplemental Fig. 8E]. However, after RA treatment, GLP 3A cells displayed significantly lower H3K9me2 levels [P < 2.2E-16] at genic regions than wild-type cells [Supplemental Fig 8F]. We also calculated the H3K9me2 ratio between RA-treated and untreated cells and plotted the accumulative percentage curve against the fold change of the H3K9me2 ratio. GLP 3A cells displayed a clear shift in comparison with wild-type cells [Fig. 5D], indicating a wide range of defects in establishing H3K9me2 at genic regions in GLP 3A cells. Such a defect was not observed for H3K9me1; instead, the GLP 3A cells displayed modestly higher H3K9me1 enrichment at genic regions after RA treatment than their wild-type counterparts [Fig. 5E], and this probably reflected that less H3K9me1 was converted to H3K9me2 in the GLP 3A cells.

To further analyze the defect of H3K9me2 establishment at genes in detail, we calculated the H3K9me2 ratio between RA-treated cells and ESCs for each gene and defined genes with significantly changed H3K9me2 levels during RA-induced differentiation in wild-type cells and GLP 3A cells [P < 0.05 in three biological replicates with at least 1.5-fold change]. Clearly, far more genes gained H3K9me2 than those lost H3K9me2 (3091 vs. 101). In contrast, in GLP 3A cells, only 295 genes gained H3K9me2, while 1463 genes displayed decreased H3K9me2 levels [Fig. 5F].

We next plotted the H3K9me2 changes against the average H3K9me2 fold enrichment of ESCs and RA-treated cells. First, the overall distribution of the plot for the wild-type cells was much higher than the one for GLP 3A cells [Fig. 5G], indicating that the GLP 3A cells suffered from a wide range of defects in establishing H3K9me2 at genes during differentiation. Strikingly, 93.9% of the genes that acquired H3K9me2 in wild-type cells displayed obviously reduced H3K9me2 levels in GLP 3A cells [Fig. 5G, right panel], and the average difference in H3K9me2 changes between wild-type and GLP 3A cells at these genes was 0.8 [log2] [Fig. 5H]. We next plotted the average fold enrichment of H3K9me2 at these genes for all 12 samples, and the results clearly indicated defective H3K9me2 establishment in GLP 3A cells [Fig. 5I].

We also analyzed genes with significantly changed H3K9me2 levels during differentiation in GLP 3A cells. The results showed that a large number of genes in GLP 3A cells failed to maintain the H3K9me2 levels during differentiation [Supplemental Fig. 9A]. These results again implicated that the GLP 3A cells failed to establish H3K9me2 at a wide range of genes during differentiation. Notably, at the 295 genes that acquired H3K9me2 in GLP 3A cells, the average difference in H3K9me2 changes between wild-type and GLP 3A cells was only 0.2 [log2] [Supplemental Fig. 9B], much more modest than the difference observed in Figure 5H.

Together, these results indicated that GLP 3A cells had normal H3K9me2 levels at the ESC stage but failed to correctly establish H3K9me2 levels at genic regions during RA-induced cell differentiation.

The H3K9me1-binding activity of GLP is required for efficient silencing of pluripotent genes during ESC differentiation

To examine whether the defective establishment of H3K9me2 in GLP 3A cells was associated with inefficient gene repression, we performed RNA sequencing (RNA-seq) experiments for the above-mentioned samples. We chose genes acquired by H3K9me2 during RA-induced differentiation in wild-type cells [Fig. 5G, red dots] and plotted their expression fold changes against their expression levels. The majority of these genes did not show more than twofold changes during RA treatment in wild-type cells, probably because H3K9me2 changes do not
lead to strict linear changes in gene expression (Lienert et al. 2011), and sometimes H3K9me2 can also occur during transcription (Vakoc et al. 2005, 2006; Lee et al. 2006; Yuan et al. 2007; Chaturvedi et al. 2009). Nevertheless, 551 genes were repressed (Fig. 6A, red bars), while only 215 genes were activated (Fig. 6A, green bars) during RA treatment in wild-type cells, in line with a role of H3K9me2 establishment in silencing a subset of genes during differentiation. Remarkably, these 551 genes displayed similar expression levels in wild-type and GLP 3A ESCs (Fig. 6B), but 110 of them were expressed at much higher levels [more than twofold] in GLP 3A cells than in wild-type cells after RA treatment (Fig. 6C). Moreover, even for the rest of the 441 genes that did not cross the twofold threshold, we observed an overall trend of upshift, which suggests that inefficient establishment of H3K9me2 contributed to inefficient silencing of these genes.

As a control, we also studied genes displaying significantly elevated H3K9me2 levels during RA-induced differentiation in GLP 3A cells (Supplemental Fig. 9A, red dots). Among these genes (295 in total, 151 have detectable expression), 70 genes were repressed in GLP 3A cells during RA-induced differentiation (Supplemental Fig. 9C, red bars), and their expression differences between GLP 3A and wild-type cells did not show a particularly biased shift before or after RA treatment (Supplemental Fig. 9D).

Notably, among the 110 genes that failed to be efficiently repressed in GLP 3A cells, the gene ontology (GO) term “stem cell maintenance” was enriched for >40-fold (Fig. 7A), and several critical pluripotent genes, including Oct4, Nanog, and Fgf4, were found within the list (Supplemental Table 1). The silencing defect of these genes in GLP 3A cells during RA-induced differentiation was validated by RT-qPCR results (Fig. 7B). ChIP-seq results at these gene loci are shown in Figure 7C and were verified by ChIP-qPCR (Fig. 7D).

In conclusion, >100 genes, including several critical pluripotent genes that should have been repressed during RA-induced differentiation, were not efficiently repressed due to impaired H3K9me2 establishment when the H3K9me1-binding activity of GLP was lost.

We also found that 329 genes failed to be activated in GLP 3A cells during RA-induced differentiation, unlike in the wild-type cells. Genes with GO terms related to differentiation such as “embryonic morphogenesis” and “pattern specification process” were enriched among these genes, including Sox1, Kdr, foxa1, Bmp, and Wnt family genes, indicating a defect in differentiation (Supplemental Fig. 9E). We reason that these results are in line with the impairment in efficiently shutting down the pluripotent genes during RA-induced differentiation in GLP 3A cells.

Discussion

GLP and G9a have relatively poor activities on nucleosomes in vitro, whether they catalyze H3K9 methylation on nucleosome substrates or histone octamers prior to nucleosome assembly during replication remains an open question (Shinkai and Tachibana 2011). Our results demonstrated that GLP and G9a are capable of exerting strong catalytic activities on nucleosomal substrates when they are flanked by neighboring nucleosomes carrying H3K9 methylation (Fig. 1). This suggests that GLP and G9a can use nucleosomes as their optimal substrates.

The enzymatic activation of GLP and G9a by neighboring nucleosomes methylated at H3K9 leads to several interesting questions: Is this a directional, processive methylation-spreading activity or a dose effect? Does this contribute to the mitotic inheritance of H3K9 methylation? What are the functional roles of such H3K9 methylation-mediated activation of H3K9 methyltransferases?

The activation of G9a and GLP by neighboring methylated nucleosomes is unlikely to be a directional, processive methylation-spreading activity

Our observation that GLP and G9a can sense neighboring nucleosomes premethylated at H3K9 suggests that GLP and G9a can efficiently spread H3K9 methylation. This partially resembles the H3K9 methylation-spreading activity of Clr4 in fission yeast (Al-Sady et al. 2013). However, unlike Clr4, which guides the spreading toward

Figure 6. Inefficient gene repression in GLP 3A cells during RA-induced differentiation. (A) Histogram showing the expression change of genes acquired by H3K9me2 (Fig. 5C) during differentiation in wild-type cells. [Black bars] Unchanged, [red bars] repressed, [green bars] activated. [B, C] Expression comparison of the genes acquired by H3K9me2 in wild-type cells (Fig. 5C) between wild-type and GLP 3A ESCs (B) or RA-treated cells (C). Note: Red dots in B and C represent the same set of genes, and all of the dots shown represent the red dots in Figure 5C.
one specific direction (Al-Sady et al. 2013), we believe that the spreading activity of GLP is most probably not a directional, processive event. Otherwise, it would be predicted that we could observe unaffected nucleating centers flanked by impaired H3K9me2 establishment in their surrounding regions in GLP 3A cells during RA-induced ESC differentiation, which we did not observe. Instead, we observed a general reduction of H3K9me2 within the affected regions (Fig. 7C). This is not surprising, because a mechanism that can reinforce a directional and processive spreading activity would require much more sophisticated machinery.

Instead, we favor the idea that the stimulation activities observed were due to the stabilization of GLP and G9a association with nucleosomes bearing H3K9 methylation, which leads to increased local concentration of GLP and G9a and efficient methylation of surrounding nucleosomes. Interestingly, a recent study reported that heterozygote GLP knockout mice displayed developmental delay and cranial abnormalities (Balemans et al. 2014). These phenotypes, including the bent nose phenotype observed in ~30% of the heterozygote GLP knockout mice (Balemans et al. 2014), are highly similar to the phenotypes of the GLPM/M mice observed in our study, which suggests that both dosage and H3K9me1-mediated stimulation activity of GLP are critical for bone formation and craniofacial development. It remains unclear why bone formation appears to be one of the development processes most sensitive to the impairment of GLP-mediated H3K9 methylation. It is highly conceivable that the loss of H3K9me1 binding and the subsequent stimulation activity of GLP may affect biological processes similar to those affected by the dose reduction of GLP. This can be well explained by our understanding that the H3K9me1-binding activity of GLP helps to stabilize its association with nucleosomes bearing H3K9me1 and increase its local concentration for efficient methylation. Moreover, we note that the heterozygote GLP knockout mice display much milder defects.
H3K9me2 during RA-induced differentiation of ESCs. Our results demonstrated that the H3K9me1-mediated stimulation of GLP, many genes, including several critical pluripotent genes, displayed defective H3K9me2 establishment and delayed repression during differentiation of ESCs. In vivo, the GLPM/M embryos went through the entire embryonic development stage, which may argue against a role of an H3K9me2 establishment defect observed in the in vitro ESC differentiation system. There are several potential explanations. First, the GLPM/M embryos may indeed suffer from a delayed differentiation process even at the early development stage because the GLPM/M embryos were lighter than their wild-type littermates starting from E10.5, which may be a consequence of delayed differentiation processes at earlier stages. Second, although bone ossification appears to be the most vulnerable process, milder but similar impacts could well be perceived in other cell types. It would be interesting to interrogate other cell differentiation systems, especially the osteoblasts differentiation process, which is highly relevant to the observed bone ossification defects in GLPM/M mice.

Distinct roles of GLP and G9a

GLP and G9a share high sequence similarity and similar biochemical activities and are both required for H3K9me2 levels in vivo (Tachibana et al. 2005; Collins et al. 2008). However, GLP 3A mutant mice and G9a 3A mutant mice displayed dramatic phenotype differences. The H3K9 methylation-binding activity of GLP appeared to be much more important than that of G9a. It remains unclear whether this is due to their differential expression profile or their binding affinity difference toward H3K9me1 and H3K9me2. The ankyrin repeat domain of G9a preferentially associates with H3K9me2, whereas the ankyrin repeat domain of GLP preferentially associates with H3K9me1 (Collins et al. 2008). Engineering a GLP knock-in mouse line bearing an H3K9me2-binding ankyrin repeat domain may help to resolve this question in future studies. Another potential explanation is the distinct expression profile between GLP and G9a, although these two proteins can form heterodimers (Tachibana et al. 2005), they often display different tissue-specific expression profiles (Ohno et al. 2013), suggesting that they may have functions independent of each other.

Role of GLP’s H3K9me1-mediated stimulation activity in bone ossification

Bone ossification is mediated by osteocytes, which are originated from sequential differentiation of mesenchymal stem cells and osteoblasts. Chromatin modifications, including histone deacetylation and H3K9 methylation, regulate bone formation (Tan et al. 2009; Bradley et al. 2011). During mouse embryonic development, histone H3K9 methyltransferases G9a and GLP emerge on E16.5 in prehypertrophic and hypertrophic chondrocytes in the growth plate as well as mesenchyme cells in the tooth germ, resulting in increased H3K9 methylation, which may be important in regulating chondrocyte differentiation and tooth development (Ideno et al. 2013; Kamiunten...
et al. 2014). It is highly probable that during the differentiation process of mesenchymal stem cells and/or osteoblasts in GLP/M/M mice, defective establishment of H3K9me2 at genes that should be repressed led to inefficient gene silencing, incomplete differentiation of osteocytes, and defective ossification.

Finally, we note the potential relevance between the nonhistone substrates of GLP or G9α and the phenotypes that we observed in this study. In addition to histone H3, G9α has been reported to methylate other substrates, including histone H1.4 and a number of nonhistone proteins (Sampath et al. 2007; Rathert et al. 2008; Collins and Cheng 2010). It is easily conceivable that GLP may also have additional substrates. To the best of our knowledge, no other methylated proteins have been reported to be specifically recognized by the ankyrin repeats of GLP or G9α. Considering H3K9 methylation remains the only known methyl mark that GLP recognizes via its ankyrin repeats, it is most likely that the cellular and developmental defects of GLP 3A mutants observed in this study are the consequences of impaired H3K9 methylation recognition and subsequent stimulation. However, we certainly cannot completely rule out the possibility that a fraction of the defects observed in this study may be attributed to the impaired association between GLP and an unknown methylated protein.

Materials and methods

Preparation of recombinant enzymes and substrates

For histone methyltransferase activity assays, His-tagged recombinant G9α (amino acids 601–1210, NP_001012507) and recombinant GLP (amino acids 635–1296, NP_001012536) covering both the SET domain and ankyrin repeats were expressed and purified from E. coli. Briefly, His-tagged G9α or His-tagged GLP were expressed in BL21 (DE3) cells upon IPTG induction. The cells were pelleted and resuspended in lysis buffer, snap-frozen, and stored at −80°C.

All histones were purified under denatured conditions. Histone H3K9/C110A with a C-terminal Flag-His dual tag was purified with Ni-NTA beads (Qiagen). The beads were washed with 50 column volumes of lysis buffer containing 15 mM imidazole, and recombinant proteins were then eluted with lysis buffer containing 250 mM imidazole. The purified proteins were dialyzed against lysis buffer, snap-frozen, and stored at −80°C.

All histones were purified under denatured conditions. Histone H3K9/C110A with a C-terminal Flag-His dual tag was purified with Ni-NTA beads and then used for MLA reactions.

Recombinant oligonucleosomes were assembled via sequential salt dialysis using pG5E4 plasmid DNA and histone oligomers (Sampath et al. 2007; Rathert et al. 2008; Collins and Cheng 2010). It is easily conceivable that GLP may also have additional substrates. To the best of our knowledge, no other methylated proteins have been reported to be specifically recognized by the ankyrin repeats of GLP or G9α. Considering H3K9 methylation remains the only known methyl mark that GLP recognizes via its ankyrin repeats, it is most likely that the cellular and developmental defects of GLP 3A mutants observed in this study are the consequences of impaired H3K9 methylation recognition and subsequent stimulation. However, we certainly cannot completely rule out the possibility that a fraction of the defects observed in this study may be attributed to the impaired association between GLP and an unknown methylated protein.

Histone methyltransferase assays

The reactions were performed as previously described (Yuan et al. 2012). A 40-μL reaction mixture containing S-[methyl-3H]adenosyl-methionine [PerkinElmer Life Sciences], recombinant oligonucleosomes, and enzymes in HKMT assay buffer [50 mM Tris at pH 8.5, 20 mM KCl, 10 mM MgCl₂, 10 mM DTT, 250 mM succrose] was incubated for 1 h at 25°C. The reaction products were separated by 13% SDS-PAGE, transferred to PVDF membranes, and then subjected to autoradiography. For quantification, the membranes were stained with Coomassie blue G250 followed by liquid scintillation counting for each stained histone bands.

Generation of knock-in mouse model

The targeting vectors were constructed according to previously published methods [Liu et al. 2003]. Vectors were verified by sequencing and electrotransfected into R1 ESCs. After selection with G418 and ganciclovir, surviving clones were picked and then screened by genomic PCR. To confirm that the targeting events resulted in a functional and correctly mutated allele, a large fragment that covered the targeting region was amplified from cDNA and sequenced. The correct ESC clones were used to produce chimeras. F1 heterozygous mice were obtained by crossing chimera with C57BL/6 female mice.

ESC derivation and cell culture

ESCs were derived from littermate blastocysts obtained by intercrossing GLP+/M. The obtained clones were genotyped and karyotyped. Two pairs of ESC lines were obtained. ESCs were maintained in the presence of feeders in standard ES medium. Prior to ChIP-seq and RNA-seq experiments, ESCs were cultured in feeder-free conditions in ES medium supplemented with 2I for two passages to eliminate feeder cells. For RA induction, 1 μM all-trans RA was added into the medium, and LIF was withdrawn simultaneously. The RA induction experiments were performed in triplicate. For EB differentiation, 20-μL liquid drops containing 1000 ESCs were cultured for 2 d in ES medium without LIF. Next, the cell aggregates were
Bisulfite conversion was performed with the EpiTect bisulfite kit with minor modifications. Briefly, 2 million cells were harvested in PBS and cross-linked with 1% formaldehyde for 2 min. After sonication, chromatin from 1 million cells was subjected to ChIP using the iDeal ChIP-seq kit (Diagenode). For ChIP-qPCR assays, the H3K9me2 enrichment was normalized to Olfr18 (a constitutively inactive gene) promoter (Magklara et al. 2011). For ChIP-seq, libraries were constructed with KAPA library preparation kits and then single-end-sequenced with HiSeq 2000. Three biological replicate experiments were performed using independently cultured cells. The antibodies used are listed in Supplemental Table 2.

RT-qPCR and RNA-seq

RNA samples were extracted using Trizol reagent, and genomic DNA contaminations were eliminated with Turbo DNase (Life Technologies). cDNA samples were synthesized using the Takara PrimeScript first strand cDNA synthesis kit. The qPCR results were normalized against a housekeeping gene, Gapdh. For high-throughput sequencing, total RNA were extracted, and the subsequent library preparation steps were carried out according to standard Illumina procedures at our in-house sequencing center and subjected to single-end sequencing with HiSeq 2000. Two biological replicates were performed.

Bisulfite sequencing

Bisulfite conversion was performed with the EpiTect bisulfite kit (Qiagen). The Oct4 promoter region was amplified with nested primers, and PCR fragments were purified and ligated to pMD18-T vectors. At least eight clones per sample were sequenced.

Bioinformatics analysis

Sequencing depths were between 24 million and 61 million reads for the four RNA-seq experiments and between 43 million and 180 million reads for ChIP-seq experiments. The sequencing depth and mapping percentage are listed in Supplemental Table 3.

H3K9me2 ChIP-seq reads were mapped to a mouse genome (mm9), and only uniquely mapped and nonredundant reads were kept for further analysis. The mapped reads were extended to the average fragment size and binned into 25-base-pair (bp) windows for visualization in the Integrative Genomics Viewer (JGV) browser. For the analysis of genomic regions, the mouse genome was divided into 10-kb windows, and the fold enrichment of H3K9me2 for each window was calculated as the ratio of normalized read density between ChIP and input samples, and the final H3K9me2 enrichment value was an average of values from three replicate experiments. Ten-kilobase windows with at least one case of more than a twofold change between any pair of samples were selected for average linkage clustering performed with Cluster 3.0 software.

RNA-seq data were mapped to the mouse genome (mm9) with TopHat (version 1.4.1), and the transcription levels were quantified with Cufflinks (version 2.0.2) software with two replicates. The ChIP-seq and RNA-seq data were deposited in the Gene Expression Omnibus database of NCBI under accession number GSE54412.

Mouse phenotype analysis

All of our study procedures complied with the guideline of the National Institute of Biological Sciences, Beijing, for the care and use of laboratory animals.

MicroCT scanning was performed with an Aloka Latheta LCT200. Skeleton preparation and Alizarin red staining were carried out according to a previous study (Dobreva et al. 2006). Briefly, sacrificed animals were skinned and fixed in EtOH and then stained in 0.1% Alizarin red solution for at least 3 d. Next, the samples were transferred into 1% KOH to remove tissues. The calvaria of animals were carefully dissected and photographed.

Mouse suckling assays were adapted from previous studies (Hongo et al. 2000). Briefly, shortly after giving birth to newborns, female mice were given 0.05 mg/g pentobarbital sodium intraperitoneally for anesthesia and laid on a heating block (37°C). GLP+/− and GLPMM/M pups were put on their mother’s abdomen at 0.5 cm distance from nipples. We divided the suckling process into two steps: (1) searching and licking nipples [known as the rooting reflex] and (2) stably attaching to the nipples with rhythmic suckling. Successfully accomplishing each step within 1 min was scored as 1, otherwise, it was scored as 0. Each newborn mouse was given two to three trials. At least seven newborns of each genotype were tested. The final “search and lick score” and “attachment score” were calculated as summed scores divided by trials. Statistical significance was tested using Fisher’s exact test.

Acknowledgments

We thank Y. Shinkai at RIKEN for cDNA and discussion. This work was supported by China National Science Foundation grant 31425013, Chinese Ministry of Science and Technology grants 2015CB856200 and 2011CB965300, Strategic Priority Research Program of the Chinese Academy of Sciences grant XDB08010103, and the Howard Hughes Medical Institute International Early Career Scientist Program to B.Z. This work was also supported by China National Science Foundation grant 81301803 to H.W., and grant 31401114 to Z.Z.

References


Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. 2005. Histone H3 Lysine 9 methylation and HP1α are associated


Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability

Nan Liu, Zhuqiang Zhang, Hui Wu, et al.

*Genes Dev.* 2015 29: 379-393 originally published online January 30, 2015

Access the most recent version at doi:10.1101/gad.254425.114

Supplemental Material

http://genesdev.cshlp.org/content/suppl/2015/01/29/gad.254425.114.DC1.html

References

This article cites 49 articles, 17 of which can be accessed free at:
http://genesdev.cshlp.org/content/29/4/379.full.html#ref-list-1

Creative Commons License

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genesdev.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.