Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression

Claire Rougeulle*, Pablo Navarro and Philip Avner

Unité de Génétique Moléculaire Murine, Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France

Received August 6, 2003; Revised September 30, 2003; Accepted October 13, 2003

INTRODUCTION

Modifications of histone tails are known to play an important role in the epigenetic regulation of gene expression (1,2). Among presently characterized modifications, histone methylation is unique in that it is thought to be highly stable, no histone demethylase having been identified so far. The major targets for histone methylation are conserved residues located mainly in the tails of histone H3 and H4, although these modifications also occur on H2A and H2B (2). The complexity of histone H3 methylation comes both from the presence of multiple residues that can be methylated, and from the variable number of methyl groups that can be added to each residue.

Methylation of H3 Lys 4 is, in general, associated with active, euchromatic domains, although the distribution of di- and tri-methylation within transcription units can vary: in S. cerevisiae, for example, di-methylated H3K4 is found distributed throughout the entire gene whereas tri-methylated H3K4 is mainly found associated with the 5′ region of genes (3). Methylation of H3 Lys 9 (K9), unlike that of H3K4, is a widespread marker of heterochromatin in many systems, including the inactive X chromosome in mammals (4–7). We and others have recently shown that the enrichment for methylated H3K9 on the future inactive X chromosome occurs very early during X-inactivation, in conjunction with H3K9 deacetylation and the loss of H3K4 methylation (5,6). Intriguingly, we also observed that, whereas the promoters of the X-linked genes were enriched for methylated H3K9 solely on the inactive X chromosome, in the body of such X-linked genes methylation was associated with both the inactive and the active X chromosomes (5). Our results imply not only that H3K9 methylation of promoters is associated with transcriptional silencing of the inactive X but also that the presence of methylated H3K9 within exons does not impair transcription.

Here we have used chromatin immunoprecipitation (ChIP) to quantify H3K4 methylation within several X-linked and autosomal genes. We find that X-linked genes are distinguished from most autosomal genes by a specific distribution of H3K4 di-methylation, high levels being restricted to their promoter regions. The only autosomal genes to display such a profile are imprinted genes. We suggest that promoter restricted H3K4 di-methylation is involved in the establishment of monoallelic expression.

RESULTS

X-linked genes are marked by a specific profile of H3K4 di-methylation in somatic cells

The observation that the body of X-linked genes are enriched for methylated H3K9 prompted us to analyse, in an extended number of X-linked genes, the profile of H3K4 methylation, which is usually the mirror image of H3K9 methylation (8). Using an antibody specific for di-methylated H3K4 in chromatin immunoprecipitation (ChIP) experiments, we have

*To whom correspondence should be addressed. Email: rougeull@pasteur.fr

Human Molecular Genetics, Vol. 12, No. 24 © Oxford University Press 2003; all rights reserved
analysed in female fibroblasts two positions for each gene, one in the promoter region (around the transcription initiation site), the other randomly chosen within a downstream exon. Eight X-linked genes were studied and, with the exception of Hprt (see below), all were found to display enrichment for H3K4 methylation solely over their promoter region (Fig. 1A). Given the specific enrichment for H3K9 methylation in the promoter region of genes on the inactive X (5), the peak of H3K4 methylation we observed in the promoter regions most likely correspond to the active X (Xa) chromosome, as reported previously (4,9). In support of this, we could detect similar enrichment for H3K4 methylation in the promoters of all the X-linked genes tested except Xist, in male fibroblasts, which contain a single, active X chromosome (Supplementary Fig. 1). In agreement with the previously established H3K9 methylation profile, the overall low levels of methylated H3K4 in the exons indicate that exonic regions on both the active and the inactive X are depleted for methylated H3K4.

The absence of H3K4 di-methylation that we observe within the body of X-linked genes contrasts with results reported from other systems, such as yeast (S. cerevisiae and S. pombe) and chicken (3,10,11), where high H3K4 di-methylation levels are found over the entire active domain. This led us to analyse mouse autosomal genes to clarify whether the restriction of H3K4 di-methylation to promoter regions is specific for X-linked genes or for mammalian genes in general. Strikingly, as shown in Figure 1B, the H3K4 methylation profile of the four autosomal genes tested is similar to that reported in other systems, but differs markedly from the X-linked genes profile: enrichment for methylated H3K4 is detected in exonic regions, in levels equal to or above that of promoters. The only apparent exception to this rule is the Oct3/4 gene, which was uniformly depleted for di-methylated H3K4. This most likely reflects the expression pattern of Oct3/4, which is expressed specifically in undifferentiated cells and is silent in fibroblasts. Indeed, in ES cells where it is expressed, both the promoter and exon of Oct3/4 are found to be enriched for methylated H3K4 (Fig. 2B). We conclude from our observations that in somatic cells, X-linked genes have a highly specific pattern of H3K4 di-methylation characterized by the restriction of this modification to the promoter region.

**Promoter-restricted H3K4 di-methylation of X-linked genes is already present in ES cells**

In order to determine whether the X-specific methylation profile we have identified precedes X-inactivation or is rather a consequence of this process, we next wished to assess, in undifferentiated female ES cells, the distribution of H3K4 di-methylation within the genes already tested in fibroblasts. Undifferentiated female ES cells contain two active X chromosomes, one of which is randomly inactivated upon differentiation. Figure 2 shows that, taken overall, the H3K4 methylation profiles for both autosomal and X-linked genes observed in differentiated cells are already established in undifferentiated ES cells. Similar results were obtained for undifferentiated male ES cells (data not shown). The analysis of four additional exonic positions within the G6pd gene and one within Xpct gave identical results to those described above, i.e. very low levels of H3K4 di-methylation (Supplementary Fig. 2). Hprt, which was the only X-linked gene in fibroblasts to show some exonic enrichment for di-methylated H3K4, displays the profile typical of X-linked genes in ES cells. These results indicate that the bodies of genes on both X chromosomes in female are depleted for H3K4 di-methylation before X-inactivation takes place, suggesting that this specific profile of H3K4 methylation could play a precocious role in X-inactivation, either in initiation, by marking the X chromosomes, or during the propagation of inactivation, by facilitating the rapid accumulation of H3K9 methylation. Strikingly, out of the eight X-linked genes studied, the two that have an ‘autosomal-like’ H3K4 methylation profile in ES cells (Fig. 2A) are both resistant to X-inactivation, albeit differently. One of these is Xist, the key player of X-inactivation and the only gene to be expressed solely from the inactive X. The other is Smcx, which is one of the few genes in mouse known to escape
X-inactivation, at least partially (12). In humans, SMCX is embedded within a domain that contains several genes escaping X-inactivation. Smcx is however the only gene of this otherwise conserved gene cluster to escape inactivation in both species (13). Examination of the Smcx promoter failed to detect any unusual features, the identified features on the whole being shared with many other housekeeping genes (13). The H3K4 di-methylation profile we describe here is, to our knowledge, the first report of a difference between genes subject to and genes escaping X-inactivation. Our finding that this difference is observed most strikingly in ES cells compared to fibroblasts strongly suggests that the profile of H3K4 di-methylation in undifferentiated cells is triggering the status of the corresponding gene relative to X-inactivation rather than being a consequence of this process.

H3K4 di-methylation profile of imprinted genes resembles that of X-linked genes

X-linked genes differ from autosomal genes in that they undergo X-inactivation. In other words, unlike the majority of autosomal genes, X-linked genes must have the potential to be monoallelically expressed. To test whether promoter-specific enrichment for H3K4 methylation is a feature of X-linked genes only or marks monoallelically expressed genes in general, we have analysed the distribution of di-methylated H3K4 within several imprinted genes. Two maternally imprinted (paternally expressed) genes, Igf2 and Peg3 and one paternally imprinted (maternally expressed) gene, Ube3A were tested. Strikingly, the H3K4 di-methylation profile of these imprinted genes differs from that of the other autosomal genes and resembles that of X-linked genes (Fig. 3). Moreover, as observed for X-linked genes, this profile is already detected in ES cells, even though Igf2 is not expressed and Ube3A is biallelically expressed in this cell type. Several reports have already correlated histone H3 methylation and imprinting, with the expressed allele being methylated on the K4 residue but not on the K9 whilst the silent allele showed an opposite profile (14,15). These studies were however performed on adult tissues and they focused for the most part on ‘differentially methylated regions’ (DMRs), some of which control the imprinting status of large genomic regions. No allelic differences in histone H3 methylation were found in the promoter regions of several imprinted genes within the Prader–Willi Syndrome and Angelman Syndrome regions, including Ube3A (14).
DISCUSSION

We show here that X-linked genes carry a highly specific pattern of H3K4 di-methylation in both somatic and ES cells, involving the restriction of this modification to the promoter region. This modification mark is distinct from that recently reported by O’Neill and colleagues, in which X-linked genes in female ES cells can be distinguished from autosomal genes (and from X-linked genes in male ES cells) by their higher overall levels of H3K4 di-methylation, which led to the conclusion that both Xs in female cells carry an epigenetic mark prior to inactivation that could be related to the sensing of the number of X chromosomes present in the cell (16). Comparison of autosomal and X-linked genes in our system did not provide additional support for this model. Since levels of H3K4 methylation between autosomal and X-linked genes examined by O’Neill and colleagues varied only 2-fold, it is possible that variations in the ChIP technique used, the methods of analysing the immunoprecipitated DNA and the gene positions analysed account for this difference. At least in the case of G6pd, such position effects are unlikely to be concerned since we could not detect H3K4 methylation in any of the exonic positions tested. Instead we observed that levels of H3K4 di-methylation vary along the X chromosome according to the promoter region studied, with for instance genes within the Xic (Xpct, Ftx, Chic1 and Cdx4) showing elevated levels compared to genes located elsewhere on the X, at least in ES cells (Fig. 2A). It is important to underline that the promoter specific enrichment of H3K4 di-methylation associated with X-linked genes which we observe reaches levels that are up to 70-fold higher than the levels of H3K4 di-methylation within exons (Xpct, Fig. 2A).

Intriguingly, the only X-linked genes tested that did not show promoter-specific H3K4 di-methylation are genes that are known to resist/escape inactivation. Since in mouse very few genes escape X-inactivation compared with human—where some 20% of genes escape inactivation (17)—it would be of clear interest to extend our analysis to the human.

Imprinted genes are the only autosomal genes analysed to display the promoter specific H3K4 di-methylation profile otherwise typical of X-linked genes. Our data collectively indicate that monoallelically expressed genes carry a specific profile of H3K4 di-methylation, with methylated H3K4 restricted to the promoter region. This is in contrast to the global enrichment for methylated H3K4 observed within biallelically expressed genes, whether autosomal or X-linked but escaping X-inactivation, which affects promoter and/or exonic parts. Our findings that such a profile is already present in ES cells where X-linked genes have yet to be inactivated and where some imprinted genes are biallelically expressed and some are not expressed suggest that the absence of H3K4 methylation within the body of a gene may be causally related to the long-term programming of allelic expression. Promoter-restricted H3K4 di-methylation could, in this view, be an epigenetic mark for monoallelic expression. In contrast to the epigenetic marks previously reported for X-linked and imprinted genes that differentiate between parental alleles, we believe this to be the first report of a mark that distinguishes monoallelically expressed genes from other genes. We propose that the establishment of this differential H3K4 di-methylation profile occurs early during development and results in a spatial and/or functional segregation between biallelically expressed genes and genes dedicated to be monoallelically expressed (Fig. 4). The establishment of such distinct profiles could be triggered through the recruitment of distinct H3K4 methyltransferases, which could in turn facilitate the segregation by specific association with factors or nuclear compartments. Grouping monoallelically expressed genes together could be important in establishing efficient allelic extinction of gene expression, which could require cross-talk between alleles. Indeed, significant physical proximity between chromosome homologues has been observed for several imprinted regions in mouse (18). Cross-talk between the two X chromosomes is also suspected to play a role in the counting phase of the X-inactivation process, whereby a cell senses how many Xs are present in order to keep only a single X active per diploid chromosomal complement.

MATERIALS AND METHODS

Cell culture

Mouse female fibroblasts were prepared from 13.5 dpc sex-typed embryos and cultured in DMEM (GIBCO), 10% fetal calf serum (FCS, ES cell grade, GIBCO). Female ES cells (LF2) were grown in DMEM, 15% FCS and 1000 u/ml LIF (Chemicon). Female LF2 cells were cultured on gelatin-coated flasks in the absence of feeder cells.
**Chromatin immunoprecipitation**

Immunoprecipitations of DNA associated with methylated histone H3 were performed as described previously (5). Briefly, cells were crosslinked with 1% formaldehyde for 15 min at RT. Isolated nuclei were sonicated to an average length of 300–1000 bp. We used 5 μl of di-methylated H3 lysine 4 antibody (Upstate Biotechnology, Lake Placid, NY, USA).

**Quantitative analysis by real-time PCR**

Immunoprecipitated DNA and input DNA were analysed by real-time PCR using SYBR Green Universal Mix and an ABI Prism 7700 (Perkin Elmer Applied Biosystem). Each PCR was run in triplicate to control for PCR variation. To standardize each experiment, the results are represented as a percentage of the IP by the average value of the corresponding input, both values being first normalized by the dilution factor. Each experiment was repeated on independent chromatin extracts two to three times.

**Primers**

All primer sequences are given in the 5′–3′ orientation. Xpet promoter: forward primer agagagataagggagag; reverse primer cctggtccgttgcct. Xpet exon 3: forward primer ctaagacccagatcaaat; reverse primer tetcgctgttctagac. Ftx promoter: forward primer ggccgcatgaccc; reverse primer ggacctctgttcatggt. 

Chic1 promoter: forward primer tgtgctgtcag; reverse primer ggcccggtctttc. G6pd promoter: forward primer ggacctctgccaga; reverse primer cggccgttttcacggt. 

Smcx promoter: forward primer gcgcaagttgaatctgcaaa; reverse primer ccacaggactagaa. Cdx4 promoter: forward primer ggcgccggccacat. 

**REFERENCES**


