

A Crack in Histone Lysine Methylation

Minireview

Stefan Kubicek and Thomas Jenuwein*
Research Institute of Molecular Pathology (IMP)
The Vienna Biocenter
Dr. Bohrgasse 7
A-1030 Vienna
Austria

Histone lysine methylation is regarded as a very stable modification with important functions in epigenetic gene control and for organizing chromatin domains. While more robust modifications of the chromatin template are essential to stabilize epigenetic information, there is now the first evidence for a histone lysine demethylase that reverts an activating methyl mark to the unmodified state (Shi et al., 2004 [this issue of *Cell*]).

In the nuclei of almost all eukaryotic cells, DNA is highly folded and compacted with histone and nonhistone proteins in a dynamic polymer called chromatin. This physiological template of the genetic information is subject to a great variety of alterations that include histone modifications and DNA methylation. The sum of these biochemical additions is referred to as “epigenetic information” and has been proposed to facilitate response of the chromatin template to transcriptional activity or other intrinsic and environmental signals, thereby stabilizing gene expression programs and chromatin architecture during different developmental options. Consistent with transient versus stable chromatin alterations to reflect short-term versus long-term epigenetic memory, some modifications (e.g., histone acetylation and phosphorylation) are highly dynamic, whereas others (e.g., histone methylation and DNA methylation) appear more robust. In particular, histone lysine methylation is thermodynamically very stable and the amino-methyl group is resistant to direct cleavage of the N-CH₃ bond. While several other mechanisms involving cleavage of histone N termini (Jenuwein and Allis, 2001), exchange with histone variants (Ahmad and Henikoff, 2002), or destabilization by oxidation or radical attack (Bannister et al., 2002) have been proposed, no enzymatic reaction has been identified that would remove a histone lysine methyl mark to generate the unmodified state.

The recent years have seen many surprising discoveries in the field of epigenetic control, and the first experimental evidence for a “histone lysine demethylase” is now available through the work of Y. Shi and colleagues documented in this issue of *Cell* (Shi et al., 2004). The enzyme cannot directly cleave the N-CH₃ bond but induces amine oxidation of methylated histone H3 lysine 4 (H3-K4) to generate unmodified lysine and formaldehyde. This is a plausible mechanism with solid *in vitro* evidence using recombinant enzyme and various histone peptides, where conversion of the substrates is detected by mass spectrometry. Oxidation of amino-

methyl requires the cofactor FAD and a protonated nitrogen (see Figure 1) and therefore can only attack mono- or di-methylated lysines. The reaction is not compatible with tri-methylated H3-K4 or, in general, with any tri-methylated lysine substrate. There is also a surprising specificity for the methylated H3-K4 position (an active mark), and none of the other major lysine or arginine methylation sites in the histone H3 and H4 N termini are converted. The enzyme, termed lysine-specific demethylase 1 (LSD1), shares extensive sequence homology to metabolic FAD-dependent amine oxidases but contains a nuclear localization signal (NLS) and a SWIRM domain that is often found in chromatin-associated proteins. LSD1, together with HDACs, is a component of CoREST and other corepressor complexes, some of which silence neuronal genes in non-neuron cells (Lunyak et al., 2002; Shi et al., 2003). Indeed, knocking down LSD1 in HeLa cells results in increased H3-K4 di-methylation at endogenous target promoters of several neuron-specific genes concomitant with their transcriptional upregulation. LSD1 is evolutionarily conserved from *S. pombe* to mammals, but there appears to be no direct ortholog in *S. cerevisiae*, although budding yeast has very prominent H3-K4 methylation (Santos-Rosa et al., 2002).

The implications of identifying LSD1 as the first histone lysine demethylase are several-fold. Probably the biggest surprise is its exquisite selectivity for H3-K4 methylation, despite the rather broad mechanism of amine oxidation by many metabolic enzymes or polyamine oxidases (Landry and Sternglanz, 2003). Although one would predict that comparable (nuclear) amine oxidases, particularly if recruited to chromatin by auxiliary factors, could target distinct methyl-lysine positions, such as the other active marks H3-K36 and H3-K79 methylation, there are only around 10 LSD1-related mammalian amine oxidases. Six of those contain a putative NLS (see Figure 2) and one is very similar to LSD1. By contrast, there are ≥ 50 SET-domain HMTases in mammals that confer methylation of the five major lysine positions in the H3 and H4 N termini. Thus, the 3D resolution of substrate bound LSD1 will be required to understand its specificity and to define which residues are important for histone lysine demethylation rather than generic amine oxidation. These studies will also be informative for cofactor binding and formaldehyde neutralization and could lead to high-throughput screens to identify small molecule inhibitors for LSD1 and related demethylases. The second surprise is that LSD1 can only work on mono- and di-methylated H3-K4 residues. Since LSD1 is a corepressor, it is conceivable that comparable demethylating enzymes may be found in association with HDAC complexes. In the absence of these enzymes, LSD1 appears as an “attenuated” demethylase since it will not be able to repress a fully activated promoter that is marked by H3-K4 tri-methylation (Santos-Rosa et al., 2002). As a third surprise, it is remarkable that LSD1 exclusively targets an active histone lysine methylation signal. This last result addresses the critical issues of whether there may be *activating* histone lysine demethylases that associate with HAT or nucleosome-

*Correspondence : jenuwein@imp.univie.ac.at

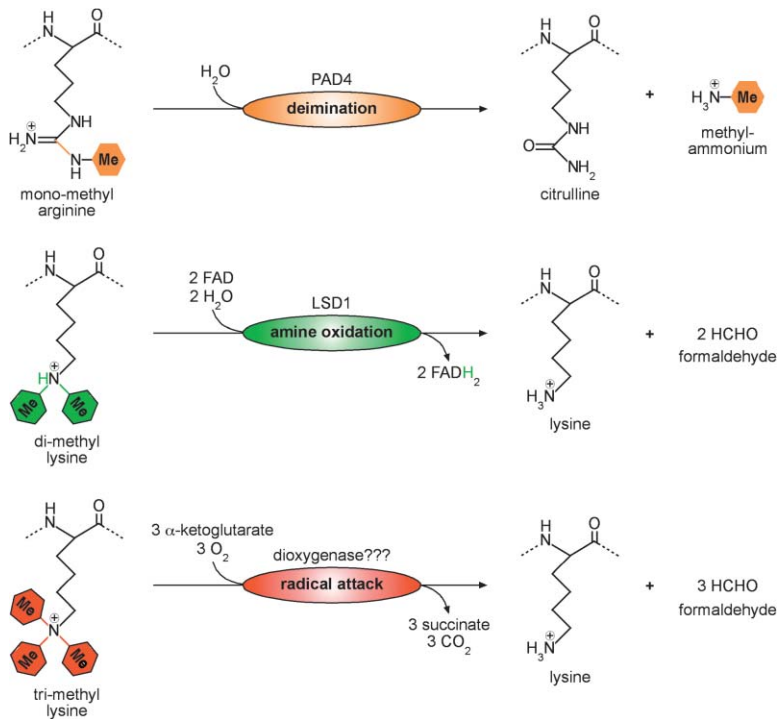


Figure 1. Enzymatic Mechanisms to Remove Histone Methylation

Arginine methylation can be by cleaved at the guanidino C-N bond (shown in orange) by protein arginine deiminases (PADs) to generate citrulline and methyl-ammonium (Cuthbert et al., 2004; Wang et al., 2004). For lysine methylation, amine oxidation represents a genuine demethylating mechanism that generates unmodified lysine and formaldehyde (Shi et al., 2004). The formation of an imine intermediate by transfer of two hydrogen atoms to the cofactor FAD requires a protonated nitrogen (shown in green), therefore restricting this reaction to mono- and di-methylated lysine substrates. Hydroxylation has been proposed as an alternative demethylating mechanism (Trewick et al., 2002; Falnes et al., 2002) that could also convert tri-methylated lysine. Direct radical attack on the methyl-carbon by Fe(II)- and α -ketoglutarate-dependent dioxxygenases would lead to the formation of an unstable carbinolamine, followed by the generation of unmodified lysine and formaldehyde. There is no experimentally proven evidence that this hydroxylation mechanism works on methylated histone lysine residues. Activating histone lysine marks (e.g., H3-K4) are shown in green, whereas repressive histone lysine methylation marks (e.g., H3-K9) are indicated in red.

remodeling activities to help revert repressive marks, such as H3-K9, H3-K27, and H4-K20 methylation, and whether there will be an active enzymatic mechanism to convert histone lysine tri-methylation to an unmodified state.

In the histone N termini, both arginine and lysine residues are subject to various degrees of methylation, and these complexities invoke further questions. Arginine methylation has mainly been associated with active transcription and can be mono- or di-methylated in symmetric or asymmetric configurations (Bannister et al., 2002; see also Figure 1). Since the adjacent imino-group can destabilize the guanidino C-N bond, arginine methylation—in contrast to the relatively inert amino-methyl in lysine methylation—is more prone to enzymatic conversion and was the first histone methylation to fall in a reaction involving de-imination (Cuthbert et al., 2004; Wang et al., 2004). The responsible enzymes, termed protein arginine demethylases or PADs, are not very prominent in mammals (as are the corresponding protein arginine methyltransferases or PRMTs) and only comprise five members (see Figure 2). Although four contain an NLS, only one is concentrated in the nucleus. The described reaction of arginine de-imination is not a true reversion of the methyl mark since it generates an altered amino acid (citrulline) and methyl-ammonium. It also works unspecifically with several mono-methylated arginine residues in the histone N termini and even on unmodified arginine (Cuthbert et al., 2004; Wang et al., 2004). To date, there is no histone arginine demethylase that would be selective for a distinct arginine position or for arginine di-methylation.

For histone lysine methylation, there are three methylation states including mono-, di-, and tri-methylation. Although amine oxidation and subsequent demethyl-

ation by LSD1 reflect a specific and genuine reversion to an unmodified histone lysine residue, it is only compatible with mono- and di-methylation. Since di-methylation is converted to the ground state, LSD1 appears to be a processive enzyme, but it cannot remove trimethyl groups from modified lysine positions. For this reaction, probably more potent enzymatic mechanisms are required. Oxidative demethylation or hydroxylation by radical attack involving Fe(II)- and α -ketoglutarate-dependent dioxxygenases (Trewick et al., 2002; Falnes et al., 2002) could represent these missing enzymes. They would have the potential to destabilize histone lysine tri-methylation by direct oxidation of the methyl-carbon, resulting in an unmodified lysine position and formaldehyde (see Figure 1). Such a mechanism has been shown to be able to remove chemically stable alkylation damage of DNA via AlkB dioxxygenases, and, although not yet experimentally proven, it has been proposed to also function for methylated protein or histone substrates (Trewick et al., 2002). Intriguingly, there are ≥ 30 dioxxygenase-related proteins in mammals that also contain an NLS.

Histone demethylases should be particularly active during dynamic transitions in transcriptional regulation, such as at early stages of lineage commitment, thereby facilitating developmental plasticity of stem cells. They would, however, be tightly controlled in differentiated cells to allow for long-term transcriptional memory and during cell divisions to stabilize chromatin organization and chromosome segregation. Indeed, histone acetylation and arginine—but not lysine—methylation is transiently restructured in early mouse embryogenesis and coincides with elevated expression of PAD enzymes (Wang et al., 2004). By contrast, there are mitotically stable repressive histone lysine modifications, such as

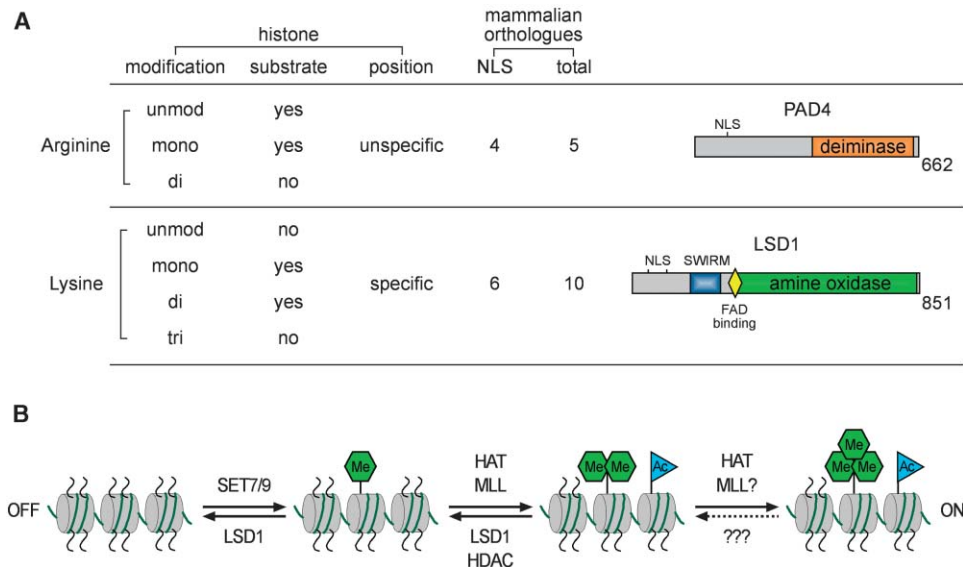


Figure 2. Substrate Specificities of Histone Demethylases and Transcriptional Stimulation by H3-K4 Methylation

(A) PAD enzymes are rather unspecific for several mono-methylated arginine positions in the histone N termini and even convert unmodified arginine (Cuthbert et al., 2004; Wang et al., 2004). By contrast, LSD1 is highly selective for H3-K4 mono- and di-methylation but cannot demethylate H3-K4 tri-methylation (Shi et al., 2004). In addition, the number of murine proteins (<http://www.ebi.ac.uk/interpro/>) that contain a deiminase or amine oxidase domain, similar to those found in PADs or LSD1, are shown. The presence of a nuclear localization signal (NLS) was predicted using PSORT (<http://psort.nibb.ac.jp/>).

(B) Diagram depicting the transition of a naïve chromatin template to the fully activated H3-K4 tri-methyl state (Santos-Rosa et al., 2002). While H3-K4 mono- and di-methylation can be removed by LSD1, no enzymatic activity has been described that would be able to revert the tri-methylated state. LSD1 is a component of Co-REST and other repressor complexes that also contain HDACs (Lunyak et al., 2002; Shi et al., 2003). In addition to H3-K4 methylation, mediated by the SET-domain HMTases SET7/9 and MLL (Santos-Rosa et al., 2002; Milne et al., 2002), transcriptional activation also involves histone acetylation by HAT enzymes.

H3-K9 and H4-K20 tri-methylation at pericentric heterochromatin and persistent H3-K27 tri-methylation at the inactive X chromosome (Xi) (Lachner et al., 2004). These patterns argue for increased stability of repressive histone lysine tri-methyl states. Very early during X inactivation, H3-K4 tri-methyl is converted to di-methyl and then mono-methyl in a process that would be incompatible with histone cleavage or exchange (O'Neill et al., 2003). Here, a specific histone lysine dioxygenase or hydroxylase may selectively remove H3-K4 tri-methylation. Similarly, there is very rapid DNA demethylation of 5-me CpG sites of the paternal genome immediately after fertilization (Mayer et al., 2000). An interesting proposal has been put forward that involves activation-induced deaminases (Aid) in the conversion of 5-me CpG to thymine (Morgan et al., 2004). Although this mechanism is highly mutagenic and primarily used for somatic hypermutation in B cells, it may potentially be allowed for the haploid genome since Guanine:Thymine mismatches could be corrected after base pairing with maternal DNA. Because DNA methylation is often dependent on preceding H3-K9 methylation, candidate histone lysine demethylases (nuclear amine oxidases and dioxygenases) may also be particularly prevalent in pluripotent tissues. By contrast, transcriptional memory mediated by Polycomb-group complexes, and enriched histone lysine tri-methyl marks at pericentric heterochromatin and at the Xi, could be protected by a more compact chromatin structure (Francis et al., 2004) that would impair access to activating transcription factors and demethylating enzymes.

The findings described by Y. Shi and colleagues (Shi et al., 2004) represent another milestone discovery in the fast-paced field of epigenetic control and give meaning to the idea that (almost) no modification lasts forever. Chromatin and histone modifications are strongly suggested by many studies to be at the core of imparting epigenetic memory from unicellular to multicellular organisms. This is highlighted, for example, by the elegant work in *Drosophila* (Cavalli and Paro, 1999) and *S. pombe* (Grewal and Klar, 1996), where transient pulses of an activator or a repressor can propagate transcriptional states over many cell divisions and even during meiosis. Thus, although other lysine demethylases are just around the corner and enzymes to destabilize repressive histone lysine tri-methyl marks may probably be identified, epigenetic control of the chromatin template refuses to be a "carte blanche."

Selected Reading

- Ahmad, K., and Henikoff, S. (2002). *Cell* 111, 281–284.
- Bannister, A.J., Schneider, R., and Kouzarides, T. (2002). *Cell* 109, 801–806.
- Cavalli, G., and Paro, R. (1999). *Science* 286, 955–958.
- Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J., and Kouzarides, T. (2004). *Cell* 118, 545–553.
- Falnes, P.O., Johansen, R.F., and Seeberg, E. (2002). *Nature* 419, 178–182.
- Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). *Science* 307, 5174–5177.
- Grewal, S.I., and Klar, A.J. (1996). *Cell* 86, 95–101.

- Jenuwein, T., and Allis, C.D. (2001). *Science* 293, 1074–1080.
- Lachner, M., Sengupta, R., Schotta, G., and Jenuwein, T. (2004). *Cold Spring Harbor Symp. Quant. Biol.* 69, in press.
- Landry, J., and Sternglanz, R. (2003). *Biochem. Biophys. Res. Commun.* 303, 771–776.
- Lunyak, V.V., Burgess, R., Prefontaine, G.G., Nelson, C., Sze, S.H., Chenoweth, J., Schwartz, P., Pevzner, P.A., Glass, C., Mandel, G., and Rosenfeld, M.G. (2002). *Science* 298, 1747–1752.
- Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). *Nature* 403, 501–502.
- Milne, T.A., Biggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). *Mol. Cell* 10, 1107–1117.
- Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). *J. Biol. Chem.* 279, 52353–52360. Published online September 24, 2004. 10.1074/jbc.M407695200.
- O'Neill, L.P., Randall, T.E., Lavender, J., Spotswood, H.T., Lee, J.T., and Turner, B.M. (2003). *Hum. Mol. Genet.* 12, 1783–1790.
- Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). *Nature* 419, 407–411.
- Shi, Y., Sawada, J., Sui, G., Affar, el B., Whetstine, J.R., Lan, F., Ogawa, H., Luke, M.P., Nakatani, Y., and Shi, Y. (2003). *Nature* 422, 735–738.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). *Cell* 119, this issue, 941–953.
- Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). *Nature* 419, 174–178.
- Wang, Y., Wysocka, J., Sayegh, J., Lee, Y.H., Perlin, J.R., Leonelli, L., Sonbuchner, L.S., McDonald, C.H., Cook, R.G., Dou, Y., et al. (2004). *Science* 306, 279–283.