69. Y. Habu, T. Kakutani, J. Paszkowski, *Curr. Opin. Genet. Dev.* **11**, 215 (2001).
70. M. Wassenegger, *Plant Mol. Biol.* **43**, 203 (2000).
71. M. A. Matzke, A. J. Matzke, J. M. Kooter, *Science* **293**, 1080 (2001).
72. J. Bender, *Trends Biochem. Sci.* **23**, 252 (1998).
73. E. U. Selker, *Cell* **97**, 157 (1999).
74. M. N. Raizada, M. I. Benito, V. Walbot, *Plant J.* **25**, 79 (2001).
75. R. F. Ketting, T. H. Haverkamp, H. G. van Luenen, R. H. Plasterk, *Cell* **99**, 133 (1999).
76. H. Tabara *et al.*, *Cell* **99**, 123 (1999).
77. B. H. Ramsahoye *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5237 (2000).
78. P. Svoboda, P. Stein, H. Hayashi, R. M. Schultz, *Development* **127**, 4147 (2000).
79. C. Cogoni *et al.*, *EMBO J.* **15**, 3153 (1996).
80. G. Faugeron, *Curr. Opin. Microbiol.* **3**, 144 (2000).
81. L. Jackson-Grusby *et al.*, *Nature Genet.* **27**, 31 (2001).
82. J. P. Vielle-Calzada, R. Baskar, U. Grossniklaus, *Nature* **404**, 91 (2000).
83. P. S. Springer, D. R. Holding, A. Groover, C. Yordan, R. A. Martienssen, *Development* **127**, 1815 (2000).
84. J. P. Vielle-Calzada *et al.*, *Genes Dev.* **13**, 2971 (1999).
85. R. Vinkenoog *et al.*, *Plant Cell* **12**, 2271 (2000).
86. S. Adams, R. Vinkenoog, M. Spielman, H. G. Dickinson, R. J. Scott, *Development* **127**, 2493 (2000).
87. M. Byrne, M. Timmermans, C. Kidner, R. Martienssen, *Curr. Opin. Plant Biol.* **4**, 38 (2001).
88. E. B. Cambareri, R. Aisner, J. Carbon, *Mol. Cell. Biol.* **18**, 5465 (1998).
89. P. SanMiguel *et al.*, *Science* **274**, 765 (1996).
90. R. Mauricio, *Nature Rev. Genet.* **2**, 370 (2001).
91. P. Cubas, C. Vincent, E. Coen, *Nature* **401**, 157 (1999).
92. R. Martienssen, *Curr. Opin. Genet. Dev.* **8**, 240 (1998).
93. Z. J. Chen, C. S. Pikaard, *Genes Dev.* **11**, 2124 (1997).
94. L. Comai *et al.*, *Plant Cell* **12**, 1551 (2000).
95. H. S. Lee, Z. J. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6753 (2001).
96. We thank E. Selker, E. Richards, V. Chandler, S. Kaeppler, S. Jacobsen, and J. Bender for communicating results prior to publication, two anonymous referees for suggestions for improvement, and our colleagues for many interesting discussions. R.M. and V.C. receive grant support from the NSF (DBI 1057338).

REVIEW

Translating the Histone Code

Thomas Jenuwein[1] and C. David Allis[2]

Chromatin, the physiological template of all eukaryotic genetic information, is subject to a diverse array of posttranslational modifications that largely impinge on histone amino termini, thereby regulating access to the underlying DNA. Distinct histone amino-terminal modifications can generate synergistic or antagonistic interaction affinities for chromatin-associated proteins, which in turn dictate dynamic transitions between transcriptionally active or transcriptionally silent chromatin states. The combinatorial nature of histone amino-terminal modifications thus reveals a "histone code" that considerably extends the information potential of the genetic code. We propose that this epigenetic marking system represents a fundamental regulatory mechanism that has an impact on most, if not all, chromatin-templated processes, with far-reaching consequences for cell fate decisions and both normal and pathological development.

Genomic DNA is the ultimate template of our heredity. Yet despite the justifiable excitement over the human genome, many challenges remain in understanding the regulation and transduction of genetic information (*1*). It is unclear, for example, why the number of protein-coding genes in humans, now estimated at ~35,000, only doubles that of the fruit fly *Drosophila melanogaster*. Is DNA alone then responsible for generating the full range of information that ultimately results in a complex eukaryotic organism, such as ourselves?

We favor the view that epigenetics, imposed at the level of DNA-packaging proteins (histones), is a critical feature of a genome-wide mechanism of information storage and retrieval that is only beginning to be understood. We propose that a "histone code" exists that may considerably extend the information potential of the genetic (DNA) code. We review emerging evidence that histone proteins and their associated covalent modifications contribute to a mechanism that can alter chromatin structure, thereby leading to inherited differences in transcriptional "on-off" states or to the stable propagation of chromosomes by defining a specialized higher order structure at centromeres. Under the assumption that a histone code exists, at least in some form, we discuss potential mechanisms for how such a code is "read" and translated into biological functions.

Throughout this review, we have chosen epigenetic phenomena and underlying mechanisms in two general categories: chromatin-based events leading to either gene activation or gene silencing. In particular, we center our discussion on examples where differences in "on-off" transcriptional states are reflected by differences in histone modifications that are either "euchromatic" (on) or "heterochromatic" (off) (Fig. 1A). We also point out that, despite many elegant genetic and biochemical insights into chromatin function and gene regulation in the budding yeast *Saccharomyces cerevisiae*, some of the heterochromatic mechanisms (e.g., HP1-based gene silencing) discussed here do not exist in an obvious form in this organism. Thus, we will need to pursue other model systems, such as *Schizosaccharomyces pombe*, *Caenorhabditis elegans*, *Drosophila*, and mice, to "crack" the histone code.

Chromatin Template and Histone Code

In the nuclei of all eukaryotic cells, genomic DNA is highly folded, constrained, and compacted by histone and nonhistone proteins in a dynamic polymer called chromatin. For example, chromosomal regions that remain transcriptionally inert are highly condensed in the interphase nucleus and remain cytologically visible as heterochromatic foci or as the "Barr body," which is the inactive X chromosome in female mammalian cells (*2*). The distinct levels of chromatin organization are dependent on the dynamic higher order structuring of nucleosomes, which represent the basic repeating unit of chromatin. In each nucleosome, roughly two superhelical turns of DNA wrap around an octamer of core histone proteins formed by four histone partners: an H3-H4 tetramer and two H2A-H2B dimers (*3*). Histones are small basic proteins consisting of a globular domain and a more flexible and charged NH_2-terminus (histone "tail") that protrudes from the nucleosome. It remains unclear how nucleosomal arrays containing linker histone (H1) then twist and fold this chromatin fiber into increasingly more compacted filaments leading to defined higher order structures.

Central to our current thinking is that chromatin structure plays an important regulatory role and that multiple signaling pathways converge on histones (*4*). Although histone proteins themselves come in generic or specialized forms (*5*), exquisite variation is provided by covalent modifications (acetylation, phosphorylation, methylation) of the histone tail domains, which allow regulatable contacts with the underlying DNA. The enzymes transducing these histone tail modifications are highly specific for particular amino acid positions (*6, 7*), thereby extending the information content of the genome past the genetic (DNA) code. This hypothesis predicts that (i) distinct modifications of the

[1]Research Institute of Molecular Pathology (IMP) at the Vienna Biocenter, Dr. Bohrgasse 7, A-1030 Vienna, Austria. E-mail: jenuwein@nt.imp.univie.ac.at
[2]Department of Biochemistry and Molecular Genetics, University of Virginia Health Science Center, Charlottesville, VA 22908, USA. E-mail: allis@virginia.edu

histone tails would induce interaction affinities for chromatin-associated proteins, and (ii) modifications on the same or different histone tails may be interdependent and generate various combinations on any one nucleosome.

Here, we wish to extend this concept for overall chromosome structure by proposing that (iii) distinct qualities of higher order chromatin, such as euchromatic or heterochromatic domains (*7*), are largely dependent on the local concentration and combination of differentially modified nucleosomes (Fig. 1A). We envision that this "nucleosome code" then permits the assembly of different epigenetic states (*7*), leading to distinct "readouts" of the genetic information, such as gene activation versus gene silencing or, more globally, cell proliferation versus cell differentiation. Any such model must account for how these epigenetic states are established, maintained, and stably inherited through mitosis and meiosis. Although there is clear evidence for a "cellular memory" (*8*), sudden switches in cell fate do occur, leading to variegating phenotypes. If the histone code hypothesis is correct, at least in part, deciphering how that code is translated into biological response remains an important and nontrivial challenge. On the basis of current knowledge, other possibilities are likely to exist, including less stringent "charge patches" in histone tails (*9*).

Clear evidence is beginning to link alterations in chromatin structure to cell cycle progression, DNA replication, DNA damage and its repair, recombination, and overall chromosome stability (*10*). It also seems likely that the fundamental nature of chromatin-based epigenetics will have an impact on X inactivation, imprinting, developmental reprogramming of cell lineages, and the plasticity of stem cells. The implications for human biology and disease, including cancer and aging, are far-reaching.

Su(var)s, Histone Methylation, and Heterochromatin

It is now widely recognized that heritable, but reversible, changes in gene expression can occur without alterations in DNA sequence. Pioneering studies on radiation-induced chromosomal translocations (*11*) provided some of the earliest findings that epigenetic "on-off" transcriptional states are largely dependent on the position of a gene within an accessible (euchromatic) or an inaccessible (heterochromatic) chromatin environment. This phenomenon, known as position-effect variegation (PEV), allowed the development of genetic screens in *Drosophila* (*12*) and *S. pombe* (*13, 14*) that have identified ~30 to 40 loci involved in modifying PEV. Similar to PEV, mating-type switching in budding (*15*) and fission (*16*) yeast represents another paradigm for a variegating mechanism where the location of a gene within a distinct chromatin environment, the *mat* region, dictates the establishment of an active or a silent transcriptional state. In particular for *S. pombe*, which appears to contain a higher order chromatin structure more closely resembling that of multicellular eukaryotes, inheritance of silenced chromatin domains has been shown to be remarkably stable during both mitosis and meiosis (*16*).

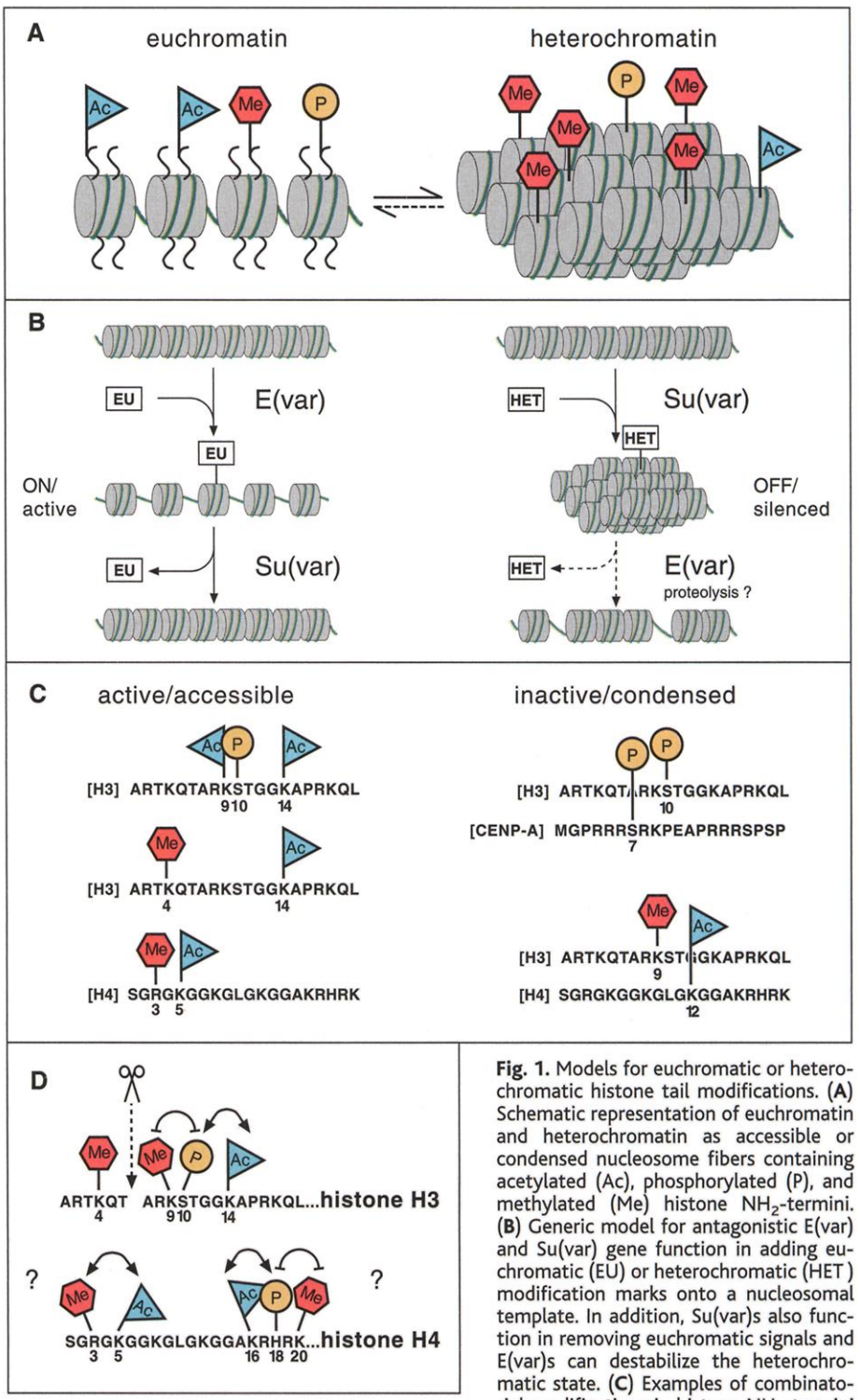

Fig. 1. Models for euchromatic or heterochromatic histone tail modifications. (**A**) Schematic representation of euchromatin and heterochromatin as accessible or condensed nucleosome fibers containing acetylated (Ac), phosphorylated (P), and methylated (Me) histone NH_2-termini. (**B**) Generic model for antagonistic E(var) and Su(var) gene function in adding euchromatic (EU) or heterochromatic (HET) modification marks onto a nucleosomal template. In addition, Su(var)s also function in removing euchromatic signals and E(var)s can destabilize the heterochromatic state. (**C**) Examples of combinatorial modifications in histone NH_2-termini that are likely to represent "imprints" for active or inactive chromatin. Single-letter abbreviations for amino acid residues: A, Ala; E, Glu; G, Gly; H, His; K, Lys; L, Leu; M, Met; P, Pro; Q, Gln; R, Arg; S, Ser; and T, Thr. (**D**) Proposed synergistic (connected arrowheads) or antagonistic (blocked oval line) modifications in histone H3 and H4 NH_2-termini. The arrow with the scissors indicates possible proteolytic cleavage of the H3 NH_2-terminus.

Among the modifier genes identified in the above model systems, one subclass suppresses variegation [the Su(var) group] and comprises gene products such as histone deacetylases (HDACs), protein phosphatases (PPTases), and S-adenosylmethionine (SAM) synthetase (*17*), as well as chromatin-associated components that are best characterized by the heterochromatin protein HP1 [Su(var)2-5] (*18*). In addition to the Su(var) group of genes, an antagonizing class of PEV modifiers enhances variegation [E(var) group] (*12*) and counteracts the Su(var)-induced silent chromatin state. Several E(var) gene products are components of adenosine triphosphate (ATP)–dependent nucleosome-remodeling machines, such as the SWI/SNF and brahma complexes (*19, 20*), which increase overall nucleosome mobility.

Extending these parallels even further, Su(var) and E(var) gene products contain several conserved protein domains—the bromo-, chromo-, and SET domains—that are also shared with two other classes of antagonizing chromatin regulators: the Polycomb (Pc-G) and trithorax (trx-G) groups. The Pc-G and trx-G genes are important for maintaining the expression boundaries of the homeotic selector genes and several other key developmental genes (*21, 22*), presumably by modulating the chromatin structure of their target loci. The bromodomain (*23*) is found in SNF2, $TAF_{II}250$, and mammalian trithorax (HRX/Mll); the chromodomain (*24, 25*) is shared between Polycomb and HP1; and the SET domain (*26*) is found in Su(var)3-9, in the Pc-G member E(z), and in trithorax. These modules have been widely used during evolution to generate a considerable functional diversity among proteins specialized in modulating chromatin structure.

Histone acetylation (*27, 28*) and histone phosphorylation (*29*) modification systems have been characterized in detail. A further class of enzymatic activities that regulate the site-specific addition of methyl groups to histones has recently been described. Originally identified as the PEV modifier *Su(var)3-9* in *Drosophila*, homologs from fission yeast (*Clr4*) to human (*SUV39H1*) have been shown to encode histone methyltransferases (HMTases) that selectively methylate histone H3 at Lys^9 (*30*). The HMTase function in the Su(var)3-9 family maps to the highly conserved SET domain but also requires adjacent Cys-rich regions. Notably, generation of the H3-Lys^9 methyl epitope induces a heterochromatic affinity for HP1 proteins that recognize this epigenetic signal through their chromodomains (*31, 32*). These results provide a strong link among Su(var) function, gene-silencing activity, and the assembly of heterochromatin (*31–35*).

By contrast, an enzymatic HMTase function has not yet been demonstrated for Pc-G and trx-G proteins. Instead, E(z) has been associated with a Pc-G complex containing HDAC activity (*36*), and trx or HRX have been shown to interact with components of chromatin-remodeling machines (*37*). In general terms, Su(var) and Pc-G gene function would be characterized by transducing the addition of heterochromatic marks and the removal of euchromatic marks on the chromatin template. Conversely, the antagonizing activity of E(var) and trx-G gene function would involve the establishment of euchromatic signals (e.g., increased nucleosome mobility) and destabilize or degrade (see below) heterochromatic "imprints" (Fig. 1B).

Translating the Histone Code

The histone code hypothesis predicts that the modification marks on the histone tails should provide binding sites for effector proteins. In agreement with this notion, the bromodomain has been the first protein module to be shown to selectively interact with a covalent mark (acetylated lysine) in the histone NH_2-terminal tail (*23, 38, 39*). In addition to the proteins discussed above, the bromodomain is also present in many transcriptional regulators having intrinsic histone acetyltransferase (HAT) activity (e.g., GCN5, PCAF, $TAF_{II}250$). Consistent with the second prediction of the histone code (that there be combinatorial readout), $TAF_{II}250$, which itself harbors several histone-modifying activities, contains two tandem copies of the bromodomain. In this configuration it preferentially binds diacetylated histone peptides presenting acetyl-lysine moieties that are appropriately spaced (*40*). Use of the Simple Modular Architectural Research Tool (SMART; http://smart.embl-heidelberg.de) indicates that there are ~75 bromodomain-containing proteins in humans. Several of these proteins, such as human poly-bromodomain protein 1, exhibit many copies (six) of regularly spaced bromodomains, which could conceivably bind to a specific combination of acetyl groups presented on one or several histone tails.

Chromodomains, on the other hand, appear to be targeting modules for methylation marks. The chromodomain of HP1 is highly selective for methylated H3 at Lys^9, and little if any binding is observed with H3 peptides containing a methylated Lys^4 position (*32*). Thus, although chromodomains are highly conserved, it seems likely that not all chromodomains—nor their methyl targets—behave similarly. In support, chromodomain swapping experiments have not uniformly indicated functional conservation in silencing assays (*41, 42*). Interestingly, Su(var)3-9 HMTase family members also contain a chromodomain, whose integrity is critical for silencing in vivo (*33, 43*). Several repressive chromatin-remodeling complexes comprise components such as the Mi-2/CHD ATPase subunit of the NuRD complex (*44*), which harbors two chromodomains and might conceivably recognize dimethylated histone tails in a manner analogous to double bromodomains. In this regard, we note that Lys^9 and Lys^{27} in the H3 tail are embedded in similar sequence motifs, and both positions are "hot spots" for methylation by the SET domain–containing HMTase G9a (*45*).

Finally, a hallmark property of all HP1 proteins is the combination of a chromodomain with a chromoshadow domain that are separated by a short but variable hinge region. Because the chromoshadow domain of HP1 appears to self-dimerize in solution (*46, 47*), it is tempting to infer that full-length HP1 may assemble intermolecular chromodomains, thereby generating a bifunctional cross-linker that is likely to stabilize the more rigid higher order structure of heterochromatin (*35, 48*).

Combinations and Switches

The above examples provide support for modification-induced recruitment of chromatin-associated proteins to acetylated and methylated histone NH_2-termini (Fig. 2A), and it is likely that other modules exist that specifically recognize phosphorylation marks. Consistent with the second prediction of the histone code hypothesis, all four NH_2-termini of the core histones contain short "basic patches" that often comprise acetylation, phosphorylation, and methylation marks in close proximity on one individual tail (*4*). All three of these modifications can be found both in active or silenced chromatin regions, which raises the question of how combinatorial specificity is used in defining an imprint for euchromatin or heterochromatin (Fig. 1, A and C).

Some evidence is emerging about a possible combinatorial code. For example, the histone H3 NH_2-terminus appears to exist in two distinct modification states that are likely to be regulated by a "switch" between Lys^9 methylation and Ser^{10} phosphorylation (Fig. 1D). Ser^{10} phosphorylation inhibits Lys^9 methylation (*30*) but is synergistically coupled with Lys^9 and/or Lys^{14} acetylation during mitogenic and hormonal stimulation in mammalian cells (*49–51*). In this phosphorylated-acetylated state, the modified H3 tail marks transcriptional activation (Fig. 1C). H3 phosphorylation is also important for mitotic chromosome condensation (*52*), where it may be linked to other secondary signal(s) such as the nucleosomal incorporation of the pericentric H3 analog Cenp-A (*53*). Conversely, aberrant Lys^9 methylation antagonizes Ser^{10} phosphorylation, leading to mitotic chromosome dysfunction (*30, 54*). Further, deacetylation of Lys^{14} in H3 (*33*) is required to facilitate subsequent Lys^9 methylation by the Clr4 HMTase, again highlighting an ordered interplay to establish distinct histone

tail modifications. Although the single H3-Lys^9 methyl epitope appears sufficient to recruit HP1 to heterochromatic regions, acetylation of Lys^{12} in H4 is another repressive mark (*55*) that may help to reinforce a silent chromatin state (Fig. 1C).

The SUV39H1 HMTase also displays weak activity toward histone H1 (*30*), and this is likely to involve methylation of Lys^{26} (*56*). RNA interference (RNAi) for an H1 variant was recently shown to phenocopy silencing and proliferation defects in the *C. elegans* germ line (*57*). These phenotypes are similar to those seen in *mes-2* mutants. Mes-2 is a homolog of the SET domain–containing E(z) member of the Pc-G group (*58*). *Su(var)3-9* (*59*) and a few other Su(var) genes, such as *E(Pc)* (*60*), have also been shown to enhance Pc-G–dependent homeotic transformations (*60, 61*). Is there a possible mechanistic link between Su(var) and Pc-G function? Because the Polycomb protein contains a chromodomain, the dual methylation of Lys^{26} in H1 and of Lys^9 in H3 could conceivably provide a combinatorial signal to recruit a Pc-G protein complex to developmentally regulated target loci (Fig. 2C).

Collectively, these observations indicate that one histone modification can influence another in either a synergistic or an antagonistic way (Fig. 1D), providing a mechanism to generate and stabilize specific imprints. During development, stem cell divisions are often characterized by one daughter cell that continues to proliferate while the other daughter cell starts to differentiate. Could the proposed "Lys^9/Ser^{10}" switch or the discussed synergisms provide an early clue about a more general mechanism for how these cell fates are chosen and maintained? Do other histone tails or entire nucleosomes contain similar switches, and to what extent has this theme been used in other nonhistone proteins?

Turning the Histone Code Upside Down

Although HP1 and H3-Lys^9 methylation are mainly associated with heterochromatic regions, HP1 also interacts with a variety of transcriptional coactivators involved in gene regulation in euchromatin (*17, 25*). Likewise, whereas histone hypoacetylation correlates most often with transcriptionally silent chromatin domains, acetylation of Lys^{12} in H4 has been reported to be a hallmark property of heterochromatin in organisms ranging from yeast to flies (*7, 55*). Also counterintuitive are the findings that mutations in the HDAC Rpd3 are enhancers rather than suppressors of PEV (*62*). These observations suggest that not all histone methylation marks correspond with gene silencing, and that some histone acetylation events may repress rather than stimulate the readout of the genetic information.

Indeed, methylation of Lys^4 in H3 occurs in transcriptionally active macronuclei of *Tetrahymena* and appears to be a euchromatic imprint in a wide range of organisms (*63*). In addition, several arginine-directed HMTases, such as the steroid receptor coactivators CARM1 and PRMT1, methylate selective arginine positions in H3 and H4 NH_2-termini and induce synergistic transcriptional activation from transiently transfected reporter constructs (*64, 65*). In vivo evidence that histones are physiological targets of these coactivators is beginning to emerge (*66, 67*). Assuming that euchromatic methylation marks exist (Fig. 1C), we predict that chromodomain-containing, positive regulators may be recruited to their target loci in much the same way that Su(var)3-9–catalyzed H3-Lys^9 methylation triggers the recruitment of HP1 to heterochromatin.

There are several intriguing candidates for such positively acting methyl-docking partners. The chromodomain-containing HAT, Esa1, is the only known essential HAT in *S. cerevisiae* (*27*) and represents the catalytic subunit of the NuA4 HAT complex, which has been linked to transcriptional activation and nucleosome remodeling in yeast and flies (*68, 69*). Because Esa1 displays robust in vitro acetylation activity toward Lys^5 in H4 (*70, 71*), it is possible that Arg^3 methylation in H4, catalyzed by the PRMT1 HMTase (*66,*

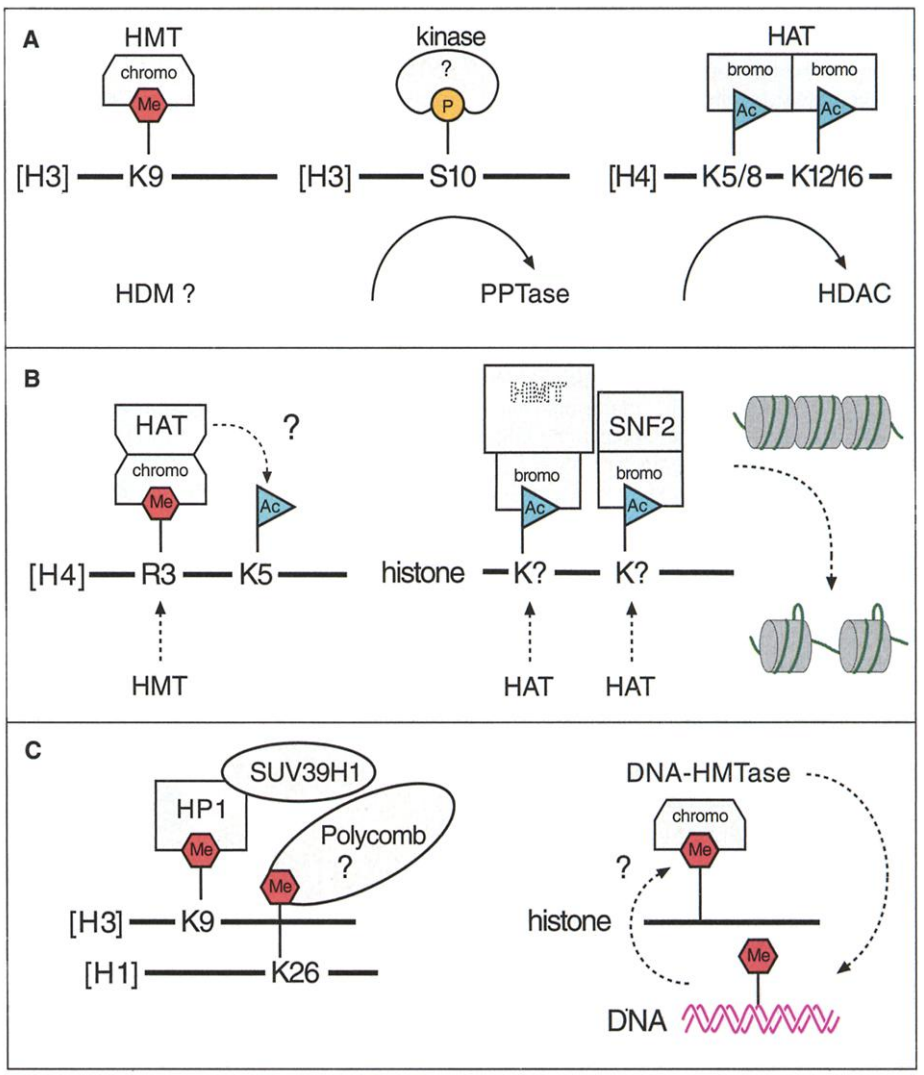

Fig. 2. Translating the "histone code." (**A**) Described protein modules of histone-modifying enzymes that have been shown to interact with site-specific methylation (chromodomain) or acetylation (bromodomain) marks in histone NH_2-termini. A protein module that would selectively recognize phosphorylated positions is currently not known. Abbreviations: HMT, histone methyltransferase; HAT, histone acetyltransferase; HDM, histone demethylase; PPTase, protein phosphatase; HDAC, histone deacetylase. (**B**) Proposed histone tail interactions for a "reversed" histone code, showing a chromodomain-containing HAT (e.g., Esa1) and part of a nucleosome-remodeling complex that may comprise a bromodomain-containing, inactive HMTase (dashed lettering), such as the trx-G protein HRX. (**C**) Possible functional interactions between Su(var) and Pc-G proteins or between histone- and DNA-methylating enzymes that could be induced or stabilized by site-selective combinations of methylation marks.

67), might play a role in recruiting Esa1 to active chromatin regions (Fig. 2B). Another chromodomain-containing HAT, Mof, has been shown to display strong selectivity for acetylation of Lys^{16} in H4, a hallmark modification correlated with the doubling of transcriptional up-regulation observed on the male X chromosome in *Drosophila* (*7*). The chromodomain of Mof has been suggested to bind RNA (*72*), raising the possibility that association with RNA—or even with methylated RNA—may contribute to the recruitment of Mof-containing complexes, which also include another chromodomain component, Msl3. Because Lys^{20} in H4 is a well-documented methylation site (*56*), it is conceivable that this methylation mark may be involved in stabilizing the fly dosage compensation complex, thereby facilitating Mof-dependent acetylation of adjacent Lys^{16}.

According to these views, appropriate methylation mark(s) would dictate the recruitment of different chromodomain-containing complexes, which in turn contribute to gene activation or gene silencing. It remains an intriguing, but undocumented, possibility that distinct histone methylation marks may also interfere with the association of repressive chromatin complexes, in much the same way that nearby modifications may influence bromodomain recognition and binding (*39*). Finally, the molecular function(s) of the bromodomain-containing HRX and SNF2 proteins are characterized by transcriptional stimulation and nucleosome remodeling. HRX also contains a SET domain that appears to be catalytically inactive (*30*) but has been shown to interact with a SWI/SNF subunit (*37*), suggesting that some remodeling complexes could transiently incorporate a "mute" HMTase (Fig. 2B). Thus, intrinsically impaired HMTase function in HRX could preclude methylation-dependent binding of repressor proteins, thereby reinforcing an activated chromatin state. It therefore seems plausible that the activities of several E(var) and trx-G proteins may be facilitated by the recruitment to transcriptionally positive histone tail modifications and by subsequently antagonizing the establishment of negative marks.

Transient Versus "Stable" Epigenetic Imprints

Given that histone methylation is linked with both euchromatic and heterochromatic states, how stable is this histone modification? On the basis of thermodynamic principles alone, methyl groups, in particular methyl-lysine, have a considerably lower turnover than do acetyl or phosphoryl groups. The latter two modifications can be removed from histone tails by the activity of HDACs or phosphatases (*29, 73*), whereas histone demethylases (HDMases) have yet to be characterized. If HDMases do not exist, histone lysine methylation would be a nearly perfect long-term epigenetic mark for maintaining chromatin states. In contrast to DNA methylation—where the methylated imprint can be removed by nucleotide excision followed by repair—DNA replication and semiconservative nucleosome distribution appears as the sole means to "dilute" histone lysine methylation below a critical threshold level.

Another potential mechanism for removing methylation marks from histone tails is proteolytic processing. Histone NH_2-termini are exposed and labile to proteolysis (*56*), and portions of certain histone tails are known to be clipped at precise stages in the cell cycle (*74*) or at specific stages of development (*75*). For example, in *Tetrahymena*, the first six amino acids are removed from the NH_2-terminus of H3 in transcriptionally silent micronuclei, but not in transcriptionally active macronuclei. H3 is ubiquitinated at specific stages of mouse spermatogenesis (*76*), and H3 is also degraded at a low level in many organisms in what is most often assumed to be uncontrolled proteolysis occurring during isolation. Ubiquitin-based protein processing, as opposed to degradation, can occur (*77*). Conserved lysines in the COOH-terminal tails of histones H2A and H2B are also subjected to monoubiquitination in a pathway that seems not to be tied to histone turnover (*78*). Further, the $TAF_{II}250$-mediated monoubiquitination of H1 has been shown to correlate with transcriptional stimulation (*79*). Whether ubiquitination may be linked to the proteolytic removal of more stable methylation marks in histone tails—or whether, in certain cases, it could even represent a synergistic signal for their addition—is not known, but remains an intriguing possibility (Fig. 3). A putative ubiquitin-specific protease is encoded by an E(var) gene in *Drosophila* (*80*), and the DNA repair and histone-ubiquitinating rph6 protein has been implicated in post-replication remodeling of the chromatin structure at the silent mating-type loci in fission yeast (*81*). Similarly, SIR-dependent gene silencing in *S. cerevisiae* also appears to be coregulated by a de-ubiquitinating enzyme (*82*).

The extent to which male versus female genomes are marked differentially by histone methylation is not known, but it seems likely that imprinting mechanisms may well use epigenetic marks outside of DNA methylation. Nearly complete removal of histones from the genome is known to occur during vertebrate spermatogenesis and other specialized developmental situations (*83*). Bulk displacement of histones during spermatogenesis would provide a means to "erase" potential male marks in the germ line, allowing the reprogramming of developmental imprints.

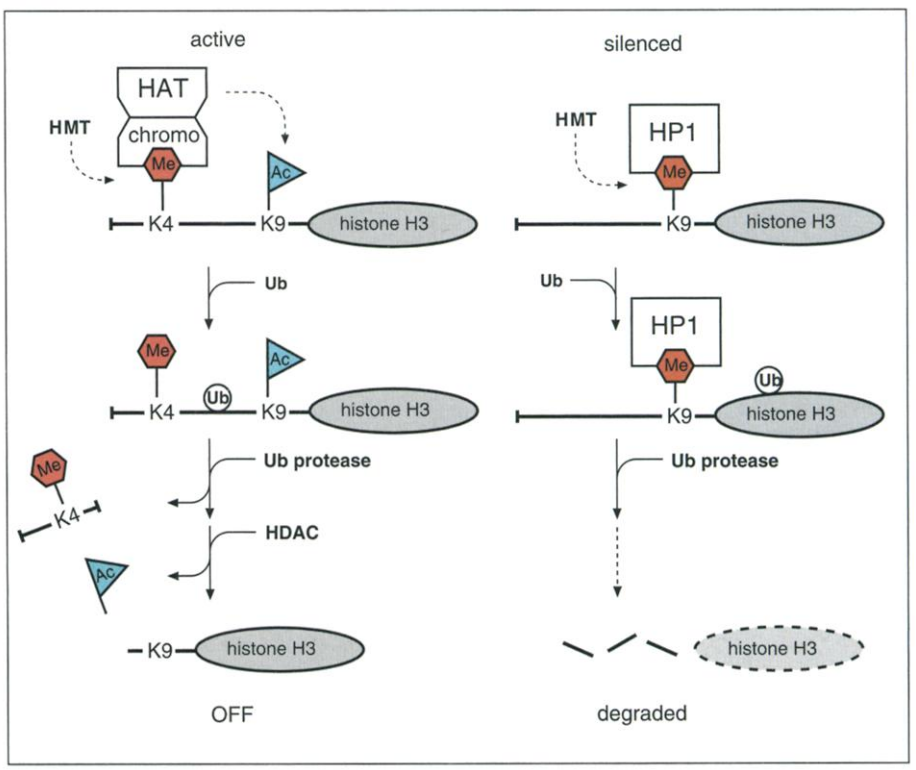

Fig. 3. A proteolytic model to remove "stable" methylation marks from histone H3. Abbreviations: Ub, ubiquitin-conjugating activity; Ub protease, ubiquitin-directed proteolytic activity. Depending on the chromatin environment and/or the nature of the ubiquitin signal, a methylated H3 NH_2-terminus may be removed by proteolytic processing (left; see also Fig. 1D), or the entire H3 molecule may be degraded (right).

Immortal Chromatin

The importance of chromatin in the information storage and decoding processes of the eukaryotic genome is reinforced by the growth in our knowledge about covalent modifications of histone proteins, and about the enzyme systems that transduce or remove these imprints. Moreover, histone modifications may also be a "sensor" of the metabolic state of the cell. For example, the Sir2 enzyme uses an essential metabolic cofactor (nicotinamide adenine dinucleotide) to regulate the activity of a family of silencing-associated HDACs (*84*). Will HDMases be uncovered only when the correct cofactor, itself possibly a direct product from intermediary carbon metabolism, is added to the test reactions? The lessons learned from the Sir2 paradigm lead to an attractive new concept: Because chromatin is the physiological template of eukaryotic cells, are genomes programmed to "open" and "close" on demand by enzyme complexes that evolved to respond directly to metabolic cues? If correct, we anticipate that further insights will be gained as we systematically investigate chromatin changes during different physiological or pathological states.

To what extent does a histone code link directly to our genetic code, or are these codes separate indexing mechanisms? Will we find evidence of interdependence between histone methylation and DNA methylation, similar to the interplay between histone deacetylation and DNA methylation (*44*)? Intriguingly, a "chromo-methylase" has recently been described in *Arabidopsis* that combines a chromodomain with a DNA methylating activity (*85*), and one member of the SET domain family contains a methyl CpG binding motif (*35*) (Fig. 2C). Histone methylation may also help to explain poorly understood chromatin effects where deacetylase inhibitors and/or 5-aza-cytosine fail to cause reversal of previously silent genomic regions (*86*). Indeed, transcription of many genes is regulated by histone acetylation in organisms (e.g., in yeast and flies) that exhibit little DNA modification. Further, X chromosome inactivation in mammals correlates with hypoacetylation of histones, except for a few X-linked loci that escape this silencing mechanism (*87*). In addition, in some branches of mammalian evolution (e.g., marsupials), no allele-specific DNA methylation has been observed. Could histone methylation be one of the conserved mechanisms substituting for the apparent absence of DNA methylation in these organisms, and to what extent is the inactive X chromosome hypoacetylated (*88*) because it may be hypermethylated at distinct histone NH_2-termini?

How far will epigenetics go past transcriptional effects? Emerging evidence indicates that programmed DNA rearrangements (*89*), imprinting phenomena (*90*), germ line silencing (*57*), developmentally cued stem cell divisions (*91*), and overall chromosome stability and identity (*52, 92*) are all influenced by epigenetic alterations of the underlying chromatin structure. In keeping with the distinct qualities of accessible and inaccessible nucleosomal states, could it be that "open" (euchromatic) chromatin represents the underlying principle that is synonymous for the character of progenitor, immortal, and young cells? Conversely, is "closed" (heterochromatic) chromatin the reflection of a developmental "memory" that stabilizes lineage commitment and gradually restricts the self-renewal potential of our somatic cells? As pointed out by others (*93*), epigenetics imparts a fundamental regulatory system beyond the sequence information of our genetic code and emphasizes that "Mendel's gene is more than just a DNA moiety."

References and Notes

1. D. Baltimore, *Nature* **409**, 814 (2001).
2. S. W. Brown, *Science* **151**, 417 (1966).
3. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, T. J. Richmond, *Nature* **389**, 251 (1997).
4. P. Cheung, C. D. Allis, P. Sassone-Corsi, *Cell* **103**, 263 (2000).
5. A. P. Wolffe, D. Pruss, *Trends Genet.* **12**, 58 (1996).
6. B. D. Strahl, C. D. Allis, *Nature* **403**, 41 (2000).
7. B. M. Turner, *Bioessays* **22**, 836 (2000).
8. R. Paro, *Trends Genet.* **11**, 295 (1995).
9. Y. Dou, M. A. Gorovsky, *Mol. Cell* **6**, 255 (2000).
10. A. P. Wolffe, D. Guschin, *J. Struct. Biol.* **129**, 102 (2000).
11. H. J. Muller, *Proc. Int. Congr. Genet.* **1**, 213 (1932).
12. G. Reuter, P. Spierer, *Bioessays* **14**, 605 (1992).
13. G. Thon, A. J. S. Klar, *Genetics* **131**, 287 (1992).
14. R. C. Allshire, J.-P. Javerzat, N. J. Readhead, G. Cranston, *Cell* **76**, 157 (1994).
15. L. Pillus, M. Grunstein, in *Chromatin Structure and Gene Expression*, S. C. R. Elgin, Ed. (IRL Press, New York, 1995), pp. 123–146.
16. S. I. S. Grewal, A. J. S. Klar, *Cell* **86**, 95 (1996).
17. L. L. Wallrath, *Curr. Opin. Genet. Dev.* **8**, 147 (1998).
18. J. C. Eissenberg *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9923 (1990).
19. T. Tsukinama, C. Wu, *Curr. Opin. Genet. Dev.* **7**, 182 (1997).
20. A. J. Kal, T. Mahmoudi, N. B. Zak, C. P. Verrijzer, *Genes Dev.* **14**, 1058 (2000).
21. R. Paro, P. Harte, in *Epigenetic Mechanisms of Gene Regulation*, V. E. A. Russo, R. A. Martienssen, A. D. Riggs, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1996), pp. 507–528.
22. M. van Lohuizen, *Curr. Opin. Genet. Dev.* **9**, 355 (1999).
23. C. Dhalluin *et al.*, *Nature* **399**, 491 (1999).
24. R. Paro, D. S. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 263 (1991).
25. D. A. Jones, F. G. Cowell, P. B. Singh, *Bioessays* **22**, 124 (2000).
26. B. Tschiersch *et al.*, *EMBO J.* **13**, 3822 (1994).
27. S. Y. Roth, J. M. Denu, C. D. Allis, *Annu. Rev. Biochem.* **70**, 81 (2001).
28. J. A. Johnson, B. M. Turner, *Semin. Cell Dev. Biol.* **10**, 179 (1999).
29. J.-Y. Hsu *et al.*, *Cell* **102**, 279 (2000).
30. S. Rea *et al.*, *Nature* **406**, 593 (2000).
31. M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, *Nature* **410**, 116 (2001).
32. A. J. Bannister *et al.*, *Nature* **410**, 120 (2001).
33. J.-I. Nakayama, J. C. Rice, B. D. Strahl, C. D. Allis, S. I. S. Grewal, *Science* **292**, 110 (2001); published online 15 March 2001 (10.1126/science.1060118).
34. J. C. Rice, C. D. Allis, *Curr. Opin. Cell Biol.* **13**, 263 (2001).
35. T. Jenuwein, *Trends Cell Biol.* **11**, 266 (2001).
36. J. van der Vlag, A. P. Otte, *Nature Genet.* **23**, 474 (1999).
37. O. Rozenblatt-Rosen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4152 (1998).
38. F. Winston, C. D. Allis, *Nature Struct. Biol.* **6**, 601 (1999).
39. D. J. Owen *et al.*, *EMBO J.* **19**, 6141 (2000).
40. R. H. Jacobson, A. G. Ladurner, D. S. King, R. Tjian, *Science* **288**, 1422 (2000).
41. J. S. Platero, T. Harnett, J. C. Eissenberg, *EMBO J.* **14**, 3977 (1995).
42. G. Wang *et al.*, *Mol. Cell. Biol.* **20**, 6970 (2000).
43. A. V. Ivanova, M. J. Bonaduce, S. V. Ivanov, A. J. S. Klar, *Nature Genet.* **19**, 192 (1998).
44. J. Ahringer, *Trends Genet.* **16**, 351 (2000).
45. M. Tachibana, K. Sugimoto, T. Fukushima, Y. Shinkai, *J. Biol. Chem.* **276**, 25309 (2001).
46. S. V. Brasher *et al.*, *EMBO J.* **19**, 1587 (2000).
47. N. P. Cowieson, J. F. Partridge, R. C. Allshire, P. J. McLaughlin, *Curr. Biol.* **10**, 517 (2000).
48. J. C. Eissenberg, S. C. R. Elgin, *Curr. Opin. Genet. Dev.* **10**, 204 (2000).
49. P. Cheung *et al.*, *Mol. Cell* **5**, 905 (2000).
50. W. S. Lo *et al.*, *Mol. Cell* **5**, 917 (2000).
51. A. L. Clayton, S. Rose, M. J. Barratt, L. C. Mahadevan, *EMBO J.* **19**, 3714 (2000).
52. Y. Wei, L. Yu, J. Bowen, M. A. Gorovsky, C. D. Allis, *Cell* **97**, 99 (1999).
53. S. G. Zeitlin, C. M. Barber, C. D. Allis, K. E. Sullivan, *J. Cell Sci.* **114**, 653 (2001).
54. M. Melcher *et al.*, *Mol. Cell. Biol.* **20**, 3728 (2000).
55. B. M. Turner, A. J. Birley, J. Lavender, *Cell* **69**, 375 (1992).
56. K. E. van Holde, *Chromatin* (Springer-Verlag, New York, 1989).
57. M. A. Jedrusik, E. Schulze, *Development* **128**, 1069 (2001).
58. R. Holdeman, S. Nehrt, S. Strome, *Development* **125**, 2457 (1998).
59. G. Reuter, personal communication.
60. K. Stankunas *et al.*, *Development* **125**, 4055 (1998).
61. T. Jenuwein, G. Laible, R. Dorn, G. Reuter, *Cell. Mol. Life Sci.* **54**, 80 (1998).
62. F. De Rubertis *et al.*, *Nature* **384**, 589 (1996).
63. B. D. Strahl, R. Ohba, R. G. Cook, C. D. Allis, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14967 (1999).
64. D. Chen *et al.*, *Science* **284**, 2174 (1999).
65. M. R. Stallcup, *Oncogene* **20**, 3014 (2001).
66. B. D. Strahl *et al.*, *Curr. Biol*, **11**, 996 (2001).
67. H. Wang *et al.*, *Science* **293**, 853 (2001).
68. S. Allard *et al.*, *EMBO J.* **18**, 5108 (1999).
69. A. Eisen *et al.*, *J. Biol. Chem.* **276**, 3484 (2001).
70. E. R. Smith *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3561 (1998).
71. A. S. Clarke, J. E. Lowell, S. J. Jacobson, L. Pillus, *Mol. Cell. Biol.* **19**, 2515 (1999).
72. A. Akhtar, D. Zink, P. B. Becker, *Nature* **407**, 405 (2000).
73. J. Taunton, C. A. Hassig, S. L. Schreiber, *Science* **272**, 408 (1996).
74. C. D. Allis, J. K. Bowen, G. N. Abraham, C. V. Glover, M. A. Gorovsky, *Cell* **20**, 55 (1980).
75. R. Lin, R. G. Cook, C. D. Allis, *Genes Dev.* **5**, 1601 (1991).
76. W. M. Baarends *et al.*, *Dev. Biol.* **207**, 322 (1999).
77. V. J. Palombella, O. J. Rando, A. L. Goldberg, T. Maniatis, *Cell* **78**, 773 (1994).
78. K. Robzyk, J. Recht, M. A. Osley, *Science* **287**, 501 (2000).
79. A.-D. Pham, F. Sauer, *Science* **289**, 2357 (2000).
80. S. Henchoz, F. De Rubertis, D. Pauli, P. Spierer, *Mol. Cell. Biol.* **16**, 5717 (1996).
81. J. Singh, V. Goel, A. J. S. Klar, *Mol. Cell. Biol.* **18**, 5511 (1998).
82. D. Moazed, A. Johnson, *Cell* **86**, 667 (1996).
83. R. Ballhorn, S. Weston, C. Thomas, A. J. Wyrobek, *Exp. Cell Res.* **150**, 298 (1984).
84. S. Imai, C. M. Armstrong, M. Kaeberlein, L. Guarente, *Nature* **403**, 795 (2000).
85. A. M. Lindroth *et al.*, *Science* **292**, 2077 (2001); published online 10 May 2001 (10.1126/science.1059745).
86. A. El-Osta, A. P. Wolffe, *Gene Expr.* **9**, 63 (2000).

87. C. M. Disteche, *Trends Genet.* **11**, 17 (1995).
88. P. Jeppesen, B. M. Turner, *Cell* **74**, 281 (1993).
89. D. R. Roth, S. Y. Roth, *Cell* **103**, 699 (2000).
90. R. I. Gregory *et al.*, *Mol. Cell. Biol.* **21**, 5426 (2001).
91. G. Cavalli, R. Paro, *Science* **286**, 955 (1999).
92. A. H. F. M. Peters, D. O'Carroll, T. Jenuwein, unpublished data.
93. A. S. J. Klar, *Trends Genet.* **14**, 299 (1998).
94. We thank G. Reuter and members of our laboratories for allowing us to cite unpublished observations. We are particularly grateful to S. Rea for his assistance in preparing the figures. Supported by the IMP through Boehringer Ingelheim and by grants from the Austrian Research Promotion Fund and the Vienna Economy Promotion Fund (T.J.), and by NIH grant GM53512 and an NIH MERIT award (C.D.A.).

This article is dedicated to the memory of Alan Wolffe, an inspirational leader to all of us who have pondered the mysteries of chromatin and gene regulation.

VIEWPOINT

RNA: Guiding Gene Silencing

Marjori Matzke,[1]* Antonius J. M. Matzke,[1] Jan M. Kooter[2]

In diverse organisms, small RNAs derived from cleavage of double-stranded RNA can trigger epigenetic gene silencing in the cytoplasm and at the genome level. Small RNAs can guide posttranscriptional degradation of complementary messenger RNAs and, in plants, transcriptional gene silencing by methylation of homologous DNA sequences. RNA silencing is a potent means to counteract foreign sequences and could play an important role in plant and animal development.

RNA silencing is a new field of research that has coalesced during the last decade from independent studies on various organisms. Scientists who study plants and fungi have known since the late 1980s that interactions between homologous DNA and/or RNA sequences can silence genes and induce DNA methylation (*1*). The discovery of RNA interference (RNAi) in *Caenorhabditis elegans* in 1998 (*2*) focused attention on double-stranded RNA (dsRNA) as an elicitor of gene silencing, and indeed, many gene-silencing effects in plants are now known to be mediated by dsRNA (*3*). RNAi is usually described as a posttranscriptional gene-silencing phenomenon in which dsRNA triggers degradation of homologous mRNA in the cytoplasm (*4*). However, the potential for nuclear dsRNA to enter a pathway leading to epigenetic modifications of homologous DNA sequences and silencing at the transcriptional level should not be discounted. Although the nuclear aspects of RNA silencing have been studied primarily in plants, there are hints that similar RNA-directed DNA or chromatin modifications might occur in other organisms as well. Here we adopt a broad definition of RNA silencing that encompasses effects in the cytoplasm and the nucleus, and consider their possible developmental roles and evolutionary origins.

RNA Guiding Homologous RNA Degradation

Although they may differ in detail, RNAi in animals and the related phenomena of posttranscriptional gene silencing (PTGS) in plants and quelling in *Neurospora crassa* result from the same highly conserved mechanism, indicating an ancient origin (*5–10*). The basic process involves a dsRNA that is processed into shorter units that guide recognition and targeted cleavage of homologous mRNA. dsRNAs that trigger PTGS/RNAi can be made in the nucleus or cytoplasm in a number of ways, including transcription through inverted DNA repeats, simultaneous synthesis of sense and antisense RNAs, viral replication, and the activity of cellular or viral RNA–dependent RNA polymerases (RdRP) on single-stranded RNA templates (Fig. 1). In *C. elegans*, dsRNAs can be injected or introduced simply by soaking the worms in a solution containing dsRNA or feeding them bacteria expressing sense and antisense RNA (*10*).

Genetic and biochemical approaches are being used to dissect the mechanism of PTGS/RNAi. Putative RdRPs, putative helicases, and members of the PAZ/Piwi family are some of the common proteins identified in genetic screens in *N. crassa*, *C. elegans*, and *Arabidopsis* (*3*, *5*, *8*, *10*). Although these proteins provide clues about dsRNA synthesis and processing, the most detailed insight into the two-step RNA degradation process has come from biochemical experiments with cytoplasmic extracts from *Drosophila* (*11–15*) (Fig. 1). The first step involves a dsRNA endonuclease [ribonuclease III (RNase III)–like] activity that processes dsRNA into sense and antisense RNAs 21 to 25 nucleotides (nt) long. These small interfering RNAs (siRNAs), which were first described in a plant system (*16*), are generated in *Drosophila* by an RNase III–type protein termed Dicer. Orthologs of Dicer, which contains a helicase, dsRNA binding domains, and a PAZ domain, have been identified in *Arabidopsis*, *C. elegans*, mammals, and *Schizosaccharomyces pombe* (*15*). In the second step, the antisense siRNAs produced by Dicer serve as guides for a different ribonuclease complex, RISC (RNA-induced silencing complex), which cleaves the homologous single-stranded mRNAs. RISC from *Drosophila* extracts cofractionates with siRNAs that guide sequence-specific mRNA cleavage (*12*). RISC cuts the mRNA approximately in the middle of the region paired with antisense siRNA (*14*) (Fig. 1), after which the mRNA is further degraded. Although most protein components of RISC have not yet been identified, they might include an endonuclease, an exonuclease, a helicase, and a homology-searching activity (*6*, *10*). A candidate for a 3′,5′-exonuclease is *C. elegans* MUT7, an RNase D–like protein recovered in a screen for RNAi mutants (*10*). Another component of RISC is a protein of the PAZ/Piwi family (*17*), which could interact with Dicer through their common PAZ domains (*18*) to incorporate the siRNA into RISC (*17*). Genes encoding members of the PAZ/Piwi family (*Arabidopsis*: AGO1; *N. crassa*: QDE2; *C. elegans*: RDE1), which are homologous to the translation factor eIF2C, have been shown to be required for PTGS/RNAi in several mutant screens (*3*, *5*, *8*, *10*).

A putative RdRP was the first cellular protein shown to be required for PTGS/RNAi in genetic screens (*N. crassa*: QDE1; *C. elegans*: Ego1; *Arabidopsis*: SGS2/SDE1) (*3*, *5*, *8*, *10*), but its exact role is unclear and the predicted enzyme activity remains to be established. This protein might be dispensible when large amounts of dsRNA are produced from transgenes or when viral RdRPs are present (*5*). RdRP might be needed only when dsRNA is synthesized to initiate silencing—for example, from "aberrant" sense RNAs that are prematurely terminated or processed improperly (*19*). RISC-cleaved mRNAs may also be used as templates and converted into dsRNA, increasing the level of siRNAs and enhancing PTGS/RNAi (Fig. 1).

Putative helicases are another class of enzyme found repeatedly in mutant screens (*N. crassa*: QDE3; *C. elegans*: SMG-2; *Chlamydomas*: MUT6; *Arabidopsis*: SDE3) (*3*, *5*, *8*, *10*). Those recovered so far are not highly related and have not yet been characterized biochemically. A DNA helicase (QDE3) and members of two RNA helicase superfamilies (MUT6 and SMG2/SDE3, respectively) have

[1]Institute of Molecular Biology, Austrian Academy of Sciences, A-5020 Salzburg, Austria. [2]Department of Developmental Genetics, Vrije Universiteit, Amsterdam, Netherlands.

*To whom correspondence should be addressed. E-mail: mmatzke@imb.oeaw.ac.at