CONTROLLING CHROMATIN STRUCTURE

5.23.1 Introduction

Key Terms

Epigenetic changes influence the phenotype without altering the genotype. They consist of changes in the properties of a cell that are inherited but that do not represent a change in genetic information.

A **prion** is a proteinaceous infectious agent, which behaves as an inheritable trait, although it contains no nucleic acid. Examples are PrP^{Sc}, the agent of scrapie in sheep and bovine spongiform encephalopathy, and Psi, which confers an inherited state in yeast.

When transcription is treated in terms of interactions involving DNA and individual transcription factors and RNA polymerases, we get an accurate description of the events that occur *in vitro*, but this lacks an important feature of transcription *in vivo*. The cellular genome is organized as nucleosomes, but initiation of transcription generally is prevented if the promoter region is packaged into nucleosomes. In this sense, histones function as generalized repressors of transcription (a rather old idea), although we see in this Chapter that they are also involved in more specific interactions. Activation of a gene requires changes in the state of chromatin: the essential issue is how the transcription factors gain access to the promoter DNA.

Local chromatin structure is an integral part of controlling gene expression. Genes may exist in either of two structural conditions. Genes are found in an "active" state only in the cells in which they are expressed. The change of structure precedes the act of transcription, and indicates that the gene is "transcribable." This suggests that acquisition of the "active" structure must be the first step in gene expression. Active genes are found in domains of euchromatin with a preferential susceptibility to nucleases (see *Molecular Biology 5.20.16 Domains define regions that contain active genes*). Hypersensitive sites are created at promoters before a gene is activated (see *Molecular Biology 5.20.15 DNAase hypersensitive sites change chromatin structure*).

More recently it has turned out that there is an intimate and continuing connection between initiation of transcription and chromatin structure. Some activators of gene transcription directly modify histones; in particular, acetylation of histones is associated with gene activation. Conversely, some repressors of transcription function by deacetylating histones. So a reversible change in histone structure in the vicinity of the promoter is involved in the control of gene expression. This may be part of the mechanism by which a gene is maintained in an active or inactive state.

The mechanisms by which local regions of chromatin are maintained in an inactive (silent) state are related to the means by which an individual promoter is repressed. The proteins involved in the formation of heterochromatin act on chromatin via the histones, and modifications of the histones may be an important feature in the interaction. Once established, such changes in chromatin may persist through cell divisions, creating an **epigenetic** state in which the properties of a gene are

determined by the self-perpetuating structure of chromatin. The name epigenetic reflects the fact that a gene may have an inherited condition (it may be active or may be inactive) which does not depend on its sequence. Yet a further insight into epigenetic properties is given by the self-perpetuating structures of **prions** (proteinaceous infectious agents).

Last updated on 10-2-2001

CONTROLLING CHROMATIN STRUCTURE

5.23.2 Chromatin can have alternative states

Key Concepts

- Chromatin structure is stable and cannot be changed by altering the equilibrium of transcription factors and histones.

Two types of model have been proposed to explain how the state of expression of DNA is changed: equilibrium and discontinuous change-of-state.

Figure 23.1 shows the equilibrium model. Here the only pertinent factor is the concentration of the repressor or activator protein, which drives an equilibrium between free form and DNA-bound form. When the concentration of the protein is high enough, its DNA-binding site is occupied, and the state of expression of the DNA is affected. (Binding might either repress or activate any particular target sequence.) This type of model explains the regulation of transcription in bacterial cells, where gene expression is determined exclusively by the actions of individual repressor and activator proteins (see *Molecular Biology 3.10 The operon*). Whether a bacterial gene is transcribed can be predicted from the sum of the concentrations of the various factors that either activate or repress the individual gene. Changes in these concentrations *at any time* will change the state of expression accordingly. In most cases, the protein binding is cooperative, so that once the concentration becomes high enough, there is a rapid association with DNA, resulting in a switch in gene expression.

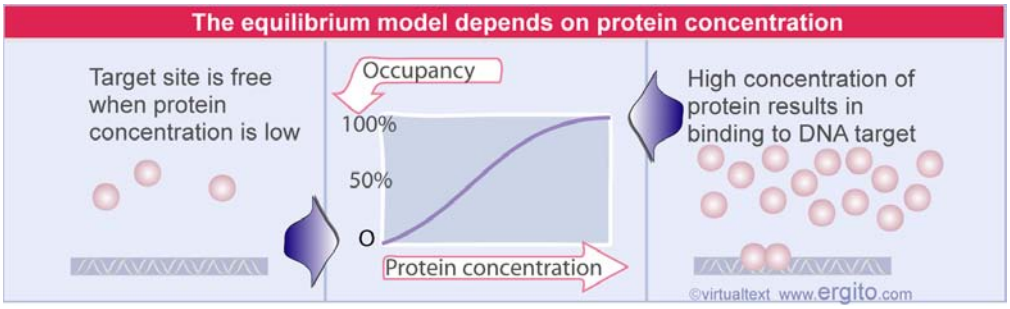


Figure 23.1 In an equilibrium model, the state of a binding site on DNA depends on the concentration of the protein that binds to it.

A different situation applies with eukaryotic chromatin. Early *in vitro* experiments showed that either an active or inactive state can be established, but this is not affected by the subsequent addition of other components. The transcription factor TFIIIA, required for RNA polymerase III to transcribe 5S rRNA genes, cannot activate its target genes *in vitro* if they are complexed with histones. However, if the factor is presented with free DNA, it forms a transcription complex, and then the addition of histones does not prevent the gene from remaining active. Once the factor has bound, it remains at the site, allowing a succession of RNA polymerase molecules to initiate transcription. Whether the factor or histones get to the control

site first may be the critical factor (680; for review see 227; 228).

Figure 23.2 illustrates the two types of condition that can exist at a eukaryotic promoter. In the inactive state, nucleosomes are present, and they prevent basal factors and RNA polymerase from binding. In the active state, the basal apparatus occupies the promoter, and histone octamers cannot bind to it. Each type of state is stable.

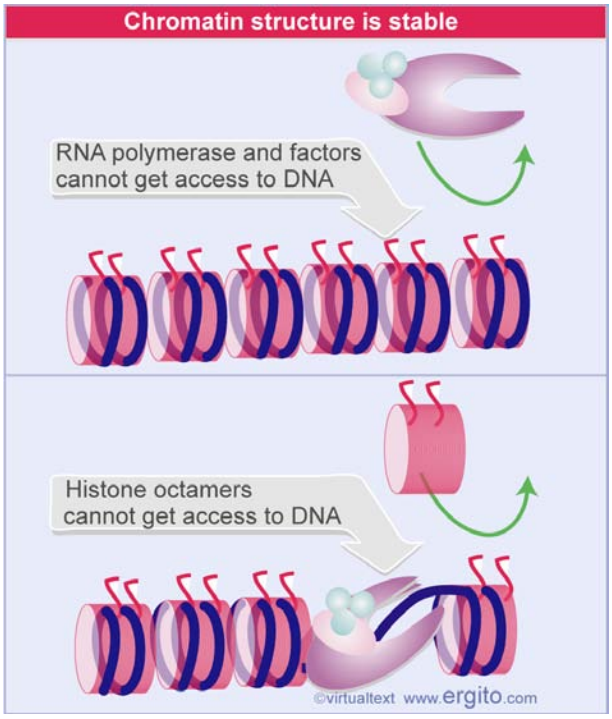


Figure 23.2 If nucleosomes form at a promoter, transcription factors (and RNA polymerase) cannot bind. If transcription factors (and RNA polymerase) bind to the promoter to establish a stable complex for initiation, histones are excluded.

A similar situation is seen with the $TF_{II}D$ complex at promoters for RNA polymerase II. A plasmid containing an adenovirus promoter can be transcribed *in vitro* by RNA polymerase II in a reaction that requires $TF_{II}D$ and other transcription factors. The template can be assembled into nucleosomes by the addition of histones. If the histones are added *before* the $TF_{II}D$, transcription cannot be initiated. But if the $TF_{II}D$ is added first, the template still can be transcribed in its chromatin form. So $TF_{II}D$ can recognize free DNA, but either cannot recognize or cannot function on nucleosomal DNA. Only the $TF_{II}D$ must be added before the histones; the other transcription factors and RNA polymerase can be added later. This suggests that binding of $TF_{II}D$ to the promoter creates a structure to which the other components of the transcription apparatus can bind (681).

It is important to note that these *in vitro* systems use disproportionate quantities of components, which may create unnatural situations. The major importance of these results, therefore, is not that they demonstrate the mechanism used *in vivo*, but that they establish the principle that *transcription factors or nucleosomes may form stable structures that cannot be changed merely by changing the equilibrium with free*

components.

Last updated on 5-12-2001

Reviews

227. Brown, D. D. (1984). *The role of stable complexes that repress and activate eukaryotic genes*. Cell 37, 359-365.

228. Weintraub, H. (1985). *Assembly and propagation of repressed and derepressed chromosomal states*. Cell 42, 705-711.

References

680. Bogenhagen, D. F., Wormington, W. M., and Brown, D. D. (1982). *Stable transcription complexes of Xenopus 5S RNA genes: a means to maintain the differentiated state*. Cell 28, 413-421.

681. Workman, J. L. and Roeder, R. G. (1987). *Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II*. Cell 51, 613-622.

CONTROLLING CHROMATIN STRUCTURE

5.23.3 Chromatin remodeling is an active process

Key Terms

Chromatin remodeling describes the energy-dependent displacement or reorganization of nucleosomes that occurs in conjunction with activation of genes for transcription.

SWI/SNF is a chromatin remodeling complex; it uses hydrolysis of ATP to change the organization of nucleosomes.

Key Concepts

- There are several chromatin remodeling complexes that use energy provided by hydrolysis of ATP.
- The SWI/SNF, RSC, and NURF complexes all are very large; there are some common subunits.
- A remodeling complex does not itself have specificity for any particular target site, but must be recruited by a component of the transcription apparatus.

The general process of inducing changes in chromatin structure is called **chromatin remodeling**. This consists of mechanisms for displacing histones that depend on the input of energy. Many protein-protein and protein-DNA contacts need to be disrupted to release histones from chromatin. There is no free ride: the energy must be provided to disrupt these contacts. **Figure 23.3** illustrates the principle of a *dynamic model* by a factor that hydrolyzes ATP. When the histone octamer is released from DNA, other proteins (in this case transcription factors and RNA polymerase) can bind.

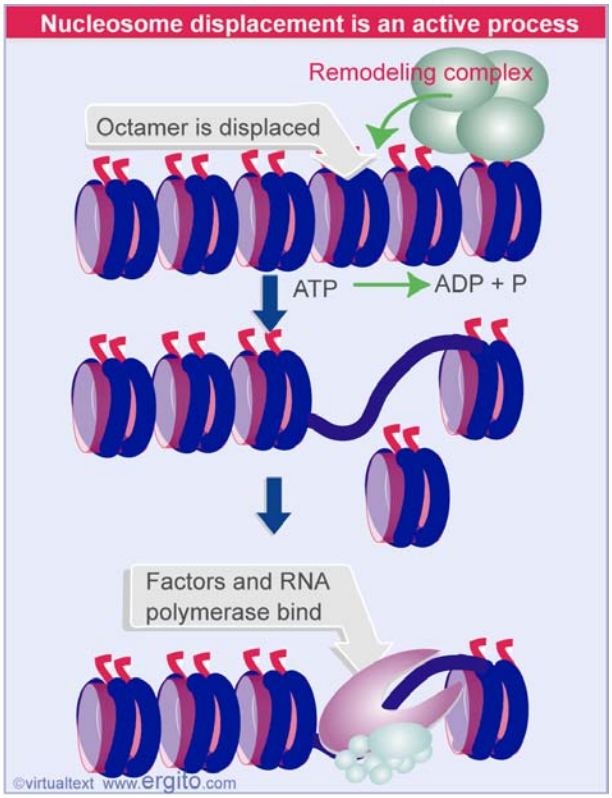


Figure 23.3 The dynamic model for transcription of chromatin relies upon factors that can use energy provided by hydrolysis of ATP to displace nucleosomes from specific DNA sequences.

Figure 23.4 summarizes the types of remodeling changes in chromatin that can be characterized *in vitro*:

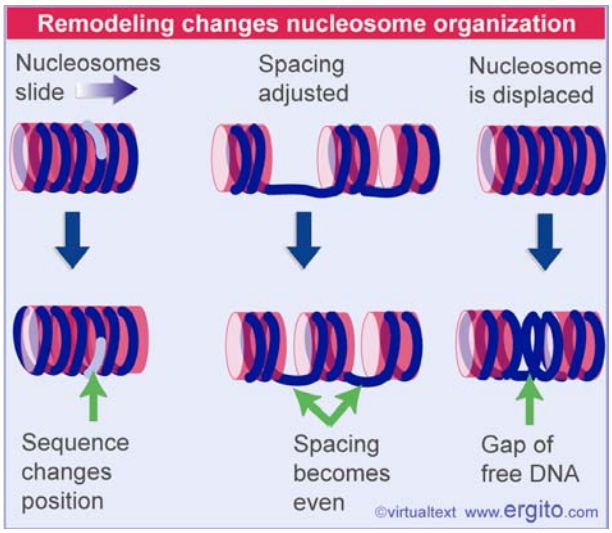
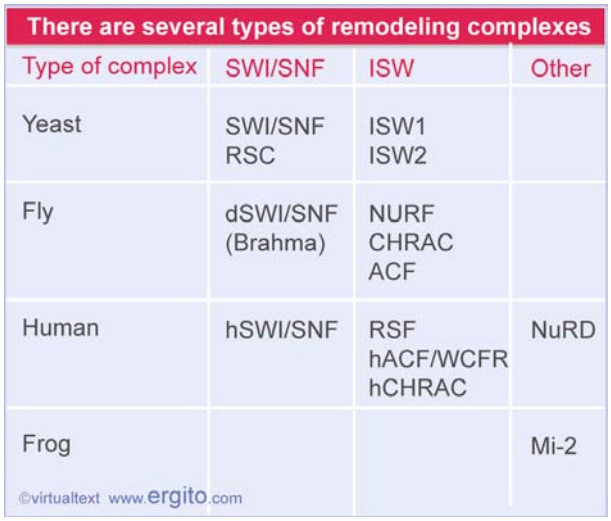


Figure 23.4 Remodeling complexes can cause nucleosomes to slide along DNA, can displace nucleosomes from DNA, or can reorganize the spacing between nucleosomes.

- Histone octamers may slide along DNA, changing the relationship between the nucleic acid and protein. This alters the position of a particular sequence on the nucleosomal surface.
- The spacing between histone octamers may be changed, again with the result that the positions of individual sequences are altered relative to protein.
- And the most extensive change is that an octamer(s) may be displaced entirely from DNA to generate a nucleosome-free gap.

The most common use of chromatin remodeling is to change the organization of nucleosomes at the promoter of a gene that is to be transcribed. This is required to allow the transcription apparatus to gain access to the promoter. The remodeling most often takes the form of displacing one or more histone octamers. This can be detected by a change in the micrococcal nuclease ladder where protection against cleavage has been lost. It often results in the creation of a site that is hypersensitive to cleavage with DNAase I (see *Molecular Biology 5.20.15 DNAase hypersensitive sites change chromatin structure*). Sometimes there are less dramatic changes, for example, involving a change in rotational positioning of a single nucleosome; this may be detected by loss of the DNAaseI 10 base ladder. So changes in chromatin structure may extend from altering the positions of nucleosomes to removing them altogether (for review see 234; 237).

Chromatin remodeling is undertaken by large complexes that use ATP hydrolysis to provide the energy for remodeling. The heart of the remodeling complex is its ATPase subunit. Remodeling complexes are usually classified according to the type of ATPase subunit – those with related ATPase subunits are considered to belong to the same family (usually some other subunits are common also). **Figure 23.5** keeps the names straight (for review see 3007). The two major types of complex are SWI/SNF and ISW (ISW stands for imitation SWI). Yeast has two complexes of each type. Complexes of both types are also found in fly and in Man (for review see 1969; 2413). Each type of complex may undertake a different range of remodeling activities (for review see 3252; 3432).



There are several types of remodeling complexes

Type of complex	SWI/SNF	ISW	Other
Yeast	SWI/SNF RSC	ISW1 ISW2	
Fly	dSWI/SNF (Brahma)	NURF CHRAC ACF	
Human	hSWI/SNF	RSF hACF/WCFR hCHRAC	NuRD
Frog			Mi-2

Figure 23.5 Remodeling complexes can be classified by their ATPase subunits.

SWI/SNF was the first remodeling complex to be identified. Its name reflects the fact that many of its subunits are coded by genes originally identified by *SWI* or *SNF* mutations in *S. cerevisiae*. Mutations in these loci are pleiotropic, and the range of defects is similar to those shown by mutants that have lost the CTD tail of RNA polymerase II. These mutations also show genetic interactions with mutations in genes that code for components of chromatin, in particular *SIN1*, which codes for a nonhistone protein, and *SIN2*, which codes for histone H3. The *SWI* and *SNF* genes are required for expression of a variety of individual loci (~120 or 2% of *S. cerevisiae* genes are affected). Expression of these loci may require the SWI/SNF complex to remodel chromatin at their promoters (684; 682; 685).

SWI/SNF acts catalytically *in vitro* (1971), and there are only ~150 complexes per yeast cell. All of the genes encoding the SWI/SNF subunits are nonessential, which implies that yeast must also have other ways of remodeling chromatin (1974; 1975). The RSC complex is more abundant and also is essential. It acts at ~ 700 target loci (2495).

SWI/SNF complexes can remodel chromatin *in vitro* without overall loss of histones or can displace histone octamers (1972; 1973). Both types of reaction may pass through the same intermediate in which the structure of the target nucleosome is altered, leading either to reformation of a (remodeled) nucleosome on the original DNA or to displacement of the histone octamer to a different DNA molecule. The SWI/SNF complex alters nucleosomal sensitivity to DNAase I at the target site, and induces changes in protein-DNA contacts that persist after it has been released from the nucleosomes (688; 694). The SWI2 subunit is the ATPase that provides the energy for remodeling by SWI/SNF.

There are many contacts between DNA and a histone octamer – 14 are identified in the crystal structure. All of these contacts must be broken for an octamer to be released or for it to move to a new position. How is this achieved? Some obvious mechanisms can be excluded because we know that single-stranded DNA is not generated during remodeling (and there are no helicase activities associated with the

complexes). Present thinking is that remodeling complexes in the SWI and ISW classes use the hydrolysis of ATP to twist DNA on the nucleosomal surface. Indirect evidence suggests that this creates a mechanical force that allows a small region of DNA to be released from the surface and then repositioned (3008).

One important reaction catalyzed by remodeling complexes involves nucleosome sliding. It was first observed that the ISW family affects nucleosome positioning without displacing octamers (1974; 1975). This is achieved by a sliding reaction, in which the octamer moves along DNA. Sliding is prevented if the N-terminal tail of histone H4 is removed, but we do not know exactly how the tail functions in this regard (2218). SWI/SNF complexes have the same capacity; the reaction is prevented by the introduction of a barrier in the DNA, which suggests that a sliding reaction is involved, in which the histone octamer moves more or less continuously along DNA without ever losing contact with it (3253).

One puzzle about the action of the SWI/SNF complex is its sheer size. It has 11 subunits with a combined molecular weight ~2×10^6. It dwarfs RNA polymerase and the nucleosome, making it difficult to understand how all of these components could interact with DNA retained on the nucleosomal surface. However, a transcription complex with full activity, called RNA polymerase II holoenzyme, can be found that contains the RNA polymerase itself, all the TF_{II} factors except TBP and $TF_{II}A$, and the SWI/SNF complex, which is associated with the CTD tail of the polymerase. In fact, virtually all of the SWI/SNF complex may be present in holoenzyme preparations. This suggests that the remodeling of chromatin and recognition of promoters is undertaken in a coordinated manner by a single complex (689).

Last updated on 10-16-2002

Reviews

234. Grunstein, M. (1990). *Histone function in transcription*. Annu. Rev. Cell Biol. 6, 643-678.

237. Felsenfeld, G. (1992). *Chromatin as an essential part of the transcriptional mechanism*. Nature 355, 219-224.

2413. Narlikar, G. J., Fan, H. Y., and Kingston, R. E. (2002). *Cooperation between complexes that regulate chromatin structure and transcription.* Cell 108, 475-487.

3007. Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000). *ATP-dependent chromatin-remodeling complexes.* Mol. Cell Biol. 20, 1899-1910.

3252. Becker, P. B. and Horz, W. (2002). *ATP-dependent nucleosome remodeling.* Annu. Rev. Biochem. 71, 247-273.

3432. Tsukiyama, T. (2002). *The in vivo functions of ATP-dependent chromatin-remodelling factors.* Nat. Rev. Mol. Cell Biol. 3, 422-429.

References

682. Cairns, B. R., Kim, Y.- J., Sayre, M. H., Laurent, B. C., and Kornberg, R. (1994). *A multisubunit complex containing the SWI/ADR6, SWI2/1, SWI3, SNF5, and SNF6 gene products isolated from yeast*. Proc. Natl. Acad. Sci. USA 91, 1950-622.

684. Peterson, C. L. and Herskowitz, I. (1992). *Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription*. Cell 68, 573-583.

685. Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, A. M., Kaufman, T. C., and Kennison, J. A. (1992). *brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2*. Cell 68, 561-572.

688. Cote, J., Quinn, J., Workman, J. L., and Peterson, C. L. (1994). *Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex*. Science 265, 53-60.

689. Kwon, H., Imbaizano, A. N., Khavari, P. A., Kingston, R. E., and Green, M. R. (1994). *Nucleosome disruption and enhancement of activator binding of human SWI/SNF complex*. Nature 370, 477-481.

694. Schnitzler, G., Sif, S., and Kingston, R. E. (1998). *Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state*. Cell 94, 17-27.

1969. Kingston, R. E. and Narlikar, G. J. (1999). *ATP-dependent remodeling and acetylation as regulators of chromatin fluidity*. Genes Dev. 13, 2339-2352.

1971. Logie, C. and Peterson, C. L. (1997). *Catalytic activity of the yeast SWI/SNF complex on reconstituted nucleosome arrays.* EMBO J. 16, 6772-6782.

1972. Lorch, Y., Cairns, B. R., Zhang, M., and Kornberg, R. D. (1998). *Activated RSC-nucleosome complex and persistently altered form of the nucleosome.* Cell 94, 29-34.

1973. Lorch, Y., Zhang, M., and Kornberg, R. D. (1999). *Histone octamer transfer by a chromatin-remodeling complex.* Cell 96, 389-392.

1974. Tsukiyama, T., Daniel, C., Tamkun, J., and Wu, C. (1995). *ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor*. Cell 83, 1021-1026.

1975. Tsukiyama, T., Palmer, J., Landel, C. C., Shiloach, J., and Wu, C. (1999). *Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in S. cerevisiae.* Genes Dev. 13, 686-697.

2218. Hamiche, A., Kang, J. G., Dennis, C., Xiao, H., and Wu, C. (2001). *Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF*. Proc. Natl. Acad. Sci. USA 98, 14316-14321.

2495. Robert, F., Young, R. A., and Struhl, K. (2002). *Genome-wide location and regulated recruitment of the RSC nucleosome remodeling complex.* Genes Dev. 16, 806-819.

3008. Gavin, I., Horn, P. J., and Peterson, C. L. (2001). *SWI/SNF chromatin remodeling requires changes in DNA topology.* Mol. Cell 7, 97-104.

3253. Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999). *Nucleosome mobilization catalysed by the yeast SWI/SNF complex.* Nature 400, 784-787.

CONTROLLING CHROMATIN STRUCTURE

5.23.4 Nucleosome organization may be changed at the promoter

Key Concepts

- Remodeling complexes are recruited to promoters by sequence-specific activators.
- The factor may be released once the remodeling complex has bound.
- The MMTV promoter requires a change in rotational positioning of a nucleosome to allow an activator to bind to DNA on the nucleosome.

How are remodeling complexes targeted to specific sites on chromatin? They do not themselves contain subunits that bind specific DNA sequences. This suggests the model shown in **Figure 23.6** in which they are recruited by activators or (sometimes) by repressors (1864; 1970).

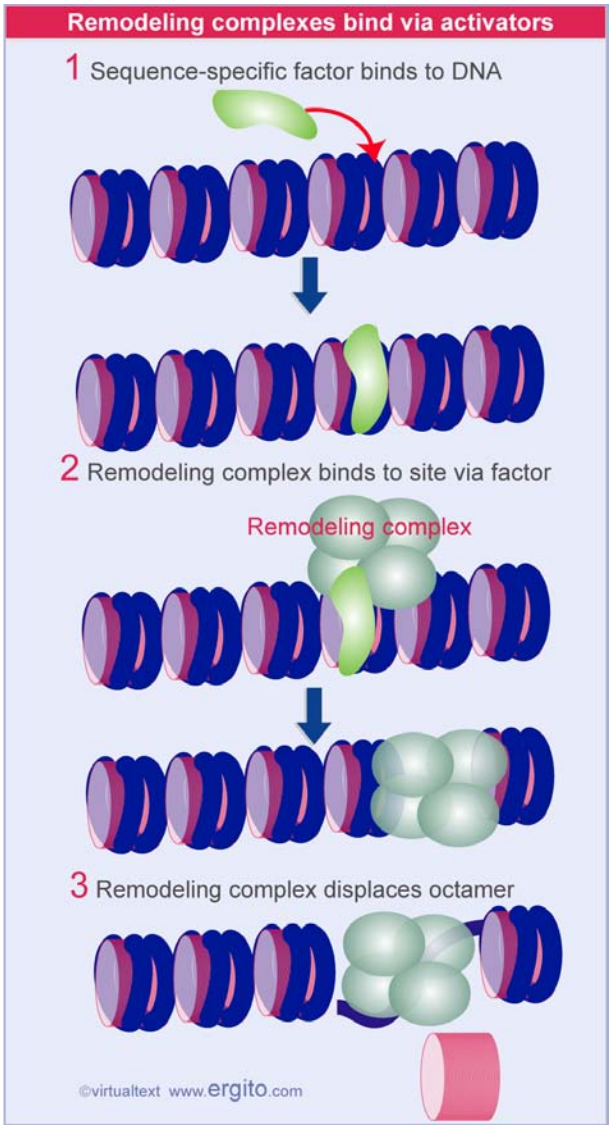


Figure 23.6 A remodeling complex binds to chromatin via an activator (or repressor).

The interaction between transcription factors and remodeling complexes gives a key insight into their modus operandi. The transcription factor Swi5p activates the *HO* locus in yeast. (Note that Swi5p is not a member of the SWI/SNF complex.) Swi5p enters nuclei toward the end of mitosis and binds to the *HO* promoter. It then recruits SWI/SNF to the promoter. Then Swi5p is released, leaving SWI/SNF at the promoter (1966). This means that a transcription factor can activate a promoter by a "hit and run" mechanism, in which its function is fulfilled once the remodeling complex has bound.

The involvement of remodeling complexes in gene activation was discovered because the complexes are necessary for the ability of certain transcription factors to activate their target genes. One of the first examples was the GAGA factor, which activates the *hsp70 Drosophila* promoter *in vitro*. Binding of GAGA to four $(CT)_n$-rich sites on the promoter disrupts the nucleosomes, creates a hypersensitive

region, and causes the adjacent nucleosomes to be rearranged so that they occupy preferential instead of random positions. Disruption is an energy-dependent process that requires the NURF remodeling complex. The organization of nucleosomes is altered so as to create a boundary that determines the positions of the adjacent nucleosomes (691). During this process, GAGA binds to its target sites and DNA, and its presence fixes the remodeled state.

The *PHO* system was one of the first in which it was shown that a change in nucleosome organization is involved in gene activation (for review see 1976). At the *PHO5* promoter, the bHLH regulator PHO4 responds to phosphate starvation by inducing the disruption of four precisely positioned nucleosomes. This event is independent of transcription (it occurs in a $TATA^-$ mutant) and independent of replication. There are two binding sites for PHO4 at the promoter, one located between nucleosomes, which can be bound by the isolated DNA-binding domain of PHO4, and the other within a nucleosome, which cannot be recognized. Disruption of the nucleosome to allow DNA binding at the second site is necessary for gene activation. This action requires the presence of the transcription-activating domain. The activator sequence of VP16 can substitute for the PHO4 activator sequence in nucleosome disruption. This suggests that disruption occurs by protein-protein interactions that involve the same region that makes protein-protein contacts to activate transcription (605). In this case, it is not known which remodeling complex is involved in executing the effects.

It is not always the case, however, that nucleosomes must be excluded in order to permit initiation of transcription. Some activators can bind to DNA on a nucleosomal surface. Nucleosomes appear to be precisely positioned at some steroid hormone response elements in such a way that receptors can bind. Receptor binding may alter the interaction of DNA with histones, and even lead to exposure of new binding sites. The exact positioning of nucleosomes could be required either because the nucleosome "presents" DNA in a particular rotational phase or because there are protein-protein interactions between the activators and histones or other components of chromatin. So we have now moved some way from viewing chromatin exclusively as a repressive structure to considering which interactions between activators and chromatin can be required for activation.

The MMTV promoter presents an example of the need for specific nucleosomal organization. It contains an array of 6 partly palindromic sites, each bound by one dimer of hormone receptor (HR), which constitute the HRE. It also has a single binding site for the factor NF1, and two adjacent sites for the factor OTF. HR and NF1 cannot bind simultaneously to their sites in free DNA. **Figure 23.7** shows how the nucleosomal structure controls binding of the factors.

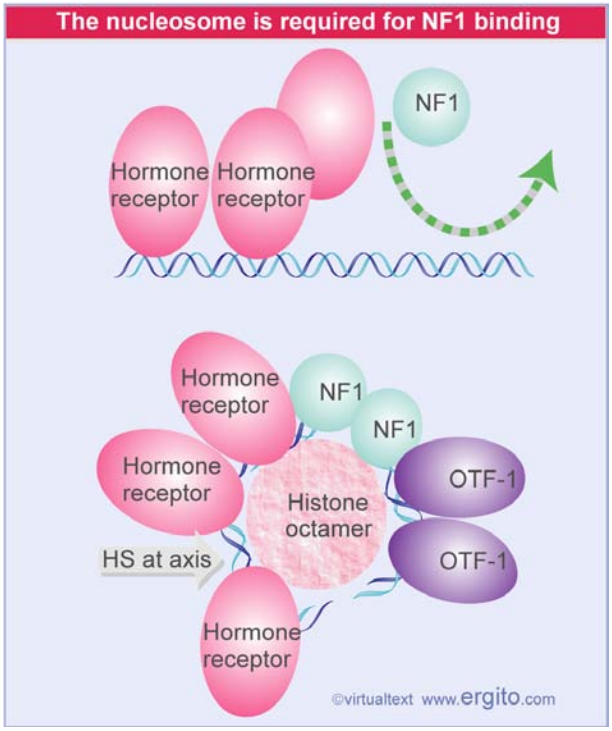


Figure 23.7 Hormone receptor and NF1 cannot bind simultaneously to the MMTV promoter in the form of linear DNA, but can bind when the DNA is presented on a nucleosomal surface.

The HR protects its binding sites at the promoter when hormone is added, but does not affect the micrococcal nuclease-sensitive sites that mark either side of the nucleosome. This suggests that HR is binding to the DNA on the nucleosomal surface. However, the rotational positioning of DNA on the nucleosome prior to hormone addition allows access to only two of the four sites. Binding to the other two sites requires a change in rotational positioning on the nucleosome. This can be detected by the appearance of a sensitive site at the axis of dyad symmetry (which is in the center of the binding sites that constitute the HRE). NF1 can be footprinted on the nucleosome after hormone induction, so these structural changes may be necessary to allow NF1 to bind, perhaps because they expose DNA and abolish the steric hindrance by which HR blocks NF1 binding to free DNA (687; 690).

Last updated on 9-5-2001

Reviews

1976. Lohr, D. (1997). *Nucleosome transactions on the promoters of the yeast GAL and PHO genes.* J. Biol. Chem. 272, 26795-26798.

References

605. Schmid, V. M., Fascher, K.-D., and Horz, W. (1992). *Nucleosome disruption at the yeast PHO5 promoter upon PHO5 induction occurs in the absence of DNA replication*. Cell 71, 853-864.

687. McPherson, C. E., Shim, E.-Y., Friedman, D. S., and Zaret, K. S. (1993). *An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array*. Cell 75, 387-398.

690. Truss, M., Barstch, J., Schelbert, A., Hache, R. J. G., and Beato, M. (1994). *Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vitro*. EMBO J. 14, 1737-1751.

691. Tsukiyama, T., Becker, P. B., and Wu, C. (1994). *ATP-dependent nucleosome disruption at a heat shock promoter mediated by binding of GAGA transcription factor*. Nature 367, 525-532.

1864. Kadam, S., McAlpine, G. S., Phelan, M. L., Kingston, R. E., Jones, K. A., and Emerson, B. M. (2000). *Functional selectivity of recombinant mammalian SWI/SNF subunits.* Genes Dev. 14, 2441-2451.

1966. Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). *Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter.* Cell 97, 299-311.

1970. Yudkovsky, N., Logie, C., Hahn, S., and Peterson, C. L. (1999). *Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators.* Genes Dev. 13, 2369-2374.

CONTROLLING CHROMATIN STRUCTURE

5.23.5 Histone modification is a key event

Key Terms

Silencing describes the repression of gene expression in a localized region, usually as the result of a structural change in chromatin.

Heterochromatin describes regions of the genome that are highly condensed, are not transcribed, and are late-replicating. Heterochromatin is divided into two types, which are called constitutive and facultative.

Whether a gene is expressed depends on the structure of chromatin both locally (at the promoter) and in the surrounding domain. Chromatin structure correspondingly can be regulated by individual activation events or by changes that affect a wide chromosomal region. The most localized events concern an individual target gene, where changes in nucleosomal structure and organization occur in the immediate vicinity of the promoter. More general changes may affect regions as large as a whole chromosome.

Changes that affect large regions control the potential of a gene to be expressed. The term **silencing** is used to refer to repression of gene activity in a local chromosomal region. The term **heterochromatin** is used to describe chromosomal regions that are large enough to be seen to have a physically more compact structure in the microscope. The basis for both types of change is the same: additional proteins bind to chromatin and either directly or indirectly prevent transcription factors and RNA polymerase from activating promoters in the region.

Changes at an individual promoter control whether transcription is initiated for a particular gene. These changes may be either activating or repressing.

All of these events depend on interactions with histones. Changes in chromatin structure are initiated by modifying the N-terminal tails of the histones, especially H3 and H4. The histone tails consist of the N-terminal 20 amino acids, and extend from the nucleosome between the turns of DNA (see **Figure 20.25** in *Molecular Biology 5.20.8 Organization of the histone octamer*). **Figure 23.8** shows that they can be modified at several sites, by methylation, acetylation, or phosphorylation (see *Molecular Biology 5.20.9 The N-terminal tails of histones are modified*). The modifications reduce positive charge. The histone modifications may directly affect nucleosome structure or create binding sites for the attachment of nonhistone proteins that change the properties of chromatin.

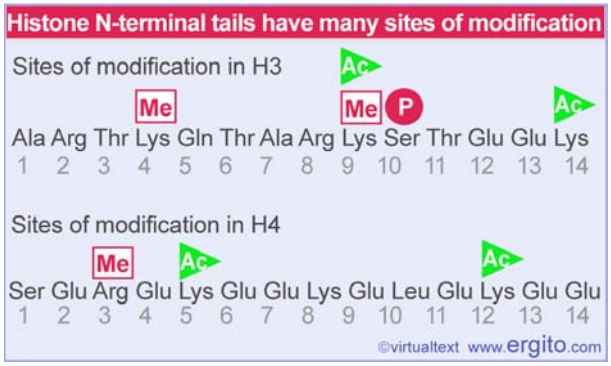


Figure 23.8 The N-terminal tails of histones H3 and H4 can be acetylated, methylated, or phosphorylated at several positions.

The range of nucleosomes that is targeted for modification can vary. Modification can be a local event, for example, restricted to nucleosomes at the promoter. Or it can be a general event, extending for example to an entire chromosome. **Figure 23.9** shows that there is a general correlation in which acetylation is associated with active chromatin while methylation is associated with inactive chromatin. However, this is not a simple rule, and the particular sites that are modified, as well as combinations of specific modifications may be important, so there are certainly exceptions in which (for example) histones methylated at a certain position are found in active chromatin. Mutations in one of the histone acetylase complexes of yeast have the opposite effect from usual (they prevent silencing of some genes), emphasizing the lack of a uniform effect of acetylation (2228).

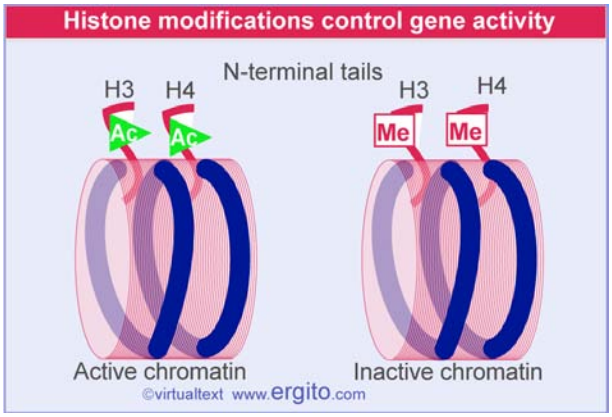


Figure 23.9 Acetylation of H3 and H4 is associated with active chromatin, while methylation is associated with inactive chromatin.

The specificity of the modifications is indicated by the fact that many of the modifying enzymes have individual target sites in specific histones. **Figure 23.10** summarizes the effects of some of the modifications. Most modified sites are subject to only a single type of modification. In some cases, modification of one site may activate or inhibit modification of another site. The idea that combinations of signals may be used to define chromatin types has sometimes been called the *histone code* (for review see 2033).

Histone modification affects the structure and function of chromatin

Histone	Site	Modification	Function
H3	Lys-4	methylation	
H3	Lys-9	methylation	chromatin condensation
	"	"	required for DNA methylation
	"	acetylation	
H3	Ser-10	phosphorylation	
H3	Lys-14	acetylation	prevents methylation at Lys-9
H3	Lys-79	methylation	telomeric silencing
H4	Arg-3	methylation	
H4	Lys-5	acetylation	
H4	Lys-12	acetylation	
H4	Lys-16	acetylation	nucleosome assembly
	"	"	fly X activation

Figure 23.10 Most modified sites in histones have a single, specific type of modification, but some sites can have more than one type of modification. Individual functions can be associated with some of the modifications.

Last updated on 1-3-2003

Reviews

2033. Jenuwein, T. and Allis, C. D. (2001). *Translating the histone code.* Science 293, 1074-1080.

References

2228. Osada, S., Sutton, A., Muster, N., Brown, C. E., Yates, J. R., Sternglanz, R., and Workman, J. L. (2001). *The yeast SAS (something about silencing) protein complex contains a MYST-type putative acetyltransferase and functions with chromatin assembly factor ASF1.* Genes Dev. 15, 3155-3168.

CONTROLLING CHROMATIN STRUCTURE

5.23.6 Histone acetylation occurs in two circumstances

Key Concepts

- Histone acetylation occurs transiently at replication.
- Histone acetylation is associated with activation of gene expression.

All the core histones can be acetylated. The major targets for acetylation are lysines in the N-terminal tails of histones H3 and H4. Acetylation occurs in two different circumstances:

- during DNA replication;
- and when genes are activated.

When chromosomes are replicated, during the S phase of the cell cycle, histones are transiently acetylated (1980; for review see 1978). **Figure 23.11** shows that this acetylation occurs before the histones are incorporated into nucleosomes. We know that histones H3 and H4 are acetylated at the stage when they are associated with one another in the $H3_2 \cdot H4_2$ tetramer. The tetramer is then incorporated into nucleosomes. Quite soon after, the acetyl groups are removed.

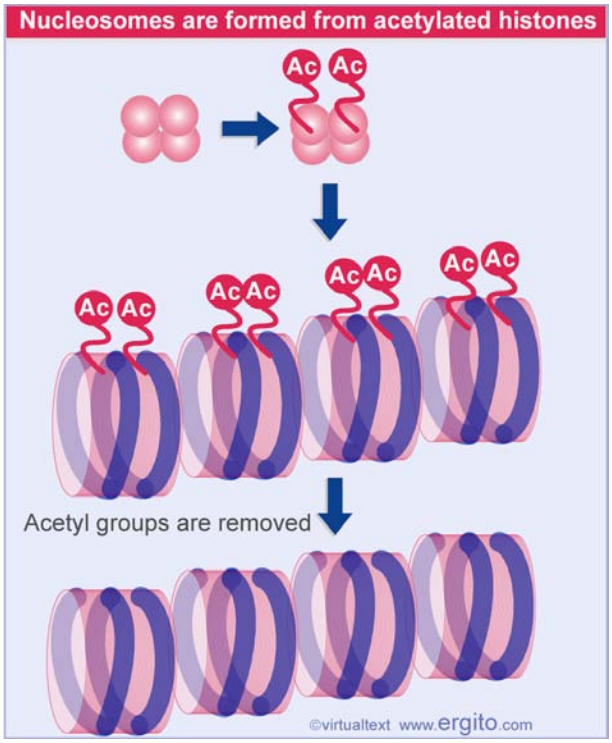


Figure 23.11 Acetylation at replication occurs on histones before they are incorporated into nucleosomes.

The importance of the acetylation is indicated by the fact that preventing acetylation of both histones H3 and H4 during replication causes loss of viability in yeast (1977). The two histones are redundant as substrates, since yeast can manage perfectly well so long as they can acetylate either one of these histones during S phase. There are two possible roles for the acetylation: it could be needed for the histones to be recognized by factors that incorporate them into nucleosomes; or it could be required for the assembly and/or structure of the new nucleosome.

The factors that are known to be involved in chromatin assembly do not distinguish between acetylated and nonacetylated histones, suggesting that the modification is more likely to be required for subsequent interactions (1979). It has been thought for a long time that acetylation might be needed to help control protein-protein interactions that occur as histones are incorporated into nucleosomes. Some evidence for such a role is that the yeast SAS histone acetylase complex binds to chromatin assembly complexes at the replication fork, where it acetylates ^{16}Lys of histone H4 (2008; 2007). This may be part of the system that establishes the histone acetylation patterns after replication.

Outside of S phase, acetylation of histones in chromatin is generally correlated with the state of gene expression. The correlation was first noticed because histone acetylation is increased in a domain containing active genes, and acetylated chromatin is more sensitive to DNAase I and (possibly) to micrococcal nuclease. **Figure 23.12** shows that this involves the acetylation of histone tails in nucleosomes. We now know that this occurs largely because of acetylation of the nucleosomes in the vicinity of the promoter when a gene is activated.

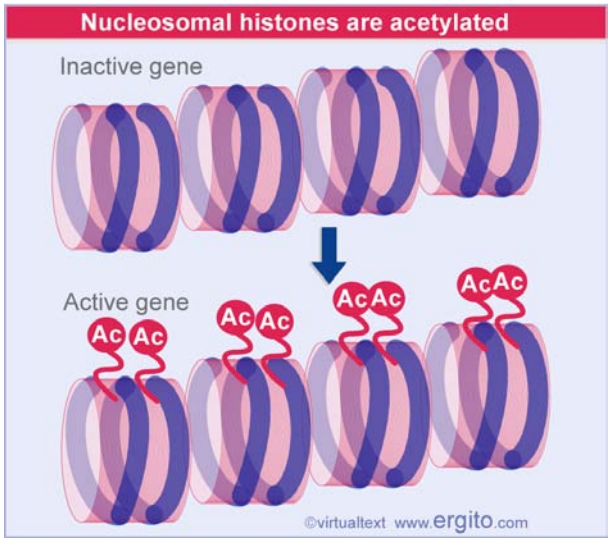


Figure 23.12 Acetylation associated with gene activation occurs by directly modifying histones in nucleosomes.

In addition to events at individual promoters, widescale changes in acetylation occur on sex chromosomes. This is part of the mechanism by which the activities of genes on the X chromosome are altered to compensate for the presence of two X chromosomes in one species but only one X chromosome (in addition to the Y chromosome) in the other species (see *Molecular Biology 5.23.17 X chromosomes undergo global changes*). The inactive X chromosome in female mammals has underacetylated H4. The super-active X chromosome in *Drosophila* males has increased acetylation of H4 (616). This suggests that the presence of acetyl groups may be a prerequisite for a less condensed, active structure. In male *Drosophila*, the X chromosome is acetylated specifically at ^{16}Lys of histone H4. The HAT that is responsible is an enzyme called MOF that is recruited to the chromosome as part of a large protein complex (1229). This "dosage compensation" complex is responsible for introducing general changes in the X chromosome that enable it to be more highly expressed. The increased acetylation is only one of its activities.

Last updated on 12-17-2001

Reviews

1978. Verreault, A. (2000). *De novo nucleosome assembly: new pieces in an old puzzle.* Genes Dev. 14, 1430-1438.

2007. Hirose, Y. and Manley, J. L. (2000). *RNA polymerase II and the integration of nuclear events.* Genes Dev. 14, 1415-1429.

References

616. Turner, B. M., Birley, A. J., and Lavender, J. (1992). *Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei*. Cell 69, 375-384.

1229. Akhtar, A. and Becker, P. B. (2000). *Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila.* Mol. Cell 5, 367-375.

1977. Ling, X., Harkness, T. A., Schultz, M. C., Fisher-Adams, G., and Grunstein, M. (1996). *Yeast histone H3 and H4 amino termini are important for nucleosome assembly in vivo and in vitro: redundant and position-independent functions in assembly but not in gene regulation.* Genes Dev. 10, 686-699.

1979. Shibahara, K., Verreault, A., and Stillman, B. (2000). *The N-terminal domains of histones H3 and H4 are not necessary for chromatin assembly factor-1- mediated nucleosome assembly onto replicated DNA in vitro.* Proc. Natl. Acad. Sci. USA 97, 7766-7771.

1980. Jackson, V., Shires, A., Tanphaichitr, N., and Chalkley, R. (1976). *Modifications to histones immediately after synthesis.* J. Mol. Biol. 104, 471-483.

2008. Alwine, J. C., Kemp, D. J., and Stark, G. R. (1977). *Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes.* Proc. Natl. Acad. Sci. USA 74, 5350-5354.

CONTROLLING CHROMATIN STRUCTURE

5.23.7 Acetylases are associated with activators

Key Terms

Histone acetyltransferase (HAT) enzymes modify histones by addition of acetyl groups; some transcriptional coactivators have HAT activity.

A **deacetylase** is an enzyme that removes acetyl groups from proteins.

Histone deacetyltransferase (HDAC) enzymes remove acetyl groups from histones; they may be associated with repressors of transcription.

Key Concepts

- Deacetylated chromatin may have a more condensed structure.
- Transcription activators are associated with histone acetylase activities in large complexes.
- Histone acetylases vary in their target specificity.
- Acetylation could affect transcription in a quantitative or qualitative way.

Acetylation is reversible. Each direction of the reaction is catalyzed by a specific type of enzyme. Enzymes that can acetylate histones are called **histone acetyltransferases** or **HATs**; the acetyl groups are removed by histone **deacetylases** or **HDACs**. There are two groups of HAT enzymes: group A describes those that are involved with transcription; group B describes those involved with nucleosome assembly.

Two inhibitors have been useful in analyzing acetylation. Trichostatin and butyric acid inhibit histone deacetylases, and cause acetylated nucleosomes to accumulate. The use of these inhibitors has supported the general view that acetylation is associated with gene expression; in fact, the ability of butyric acid to cause changes in chromatin resembling those found upon gene activation was one of the first indications of the connection between acetylation and gene activity.

The breakthrough in analyzing the role of histone acetylation was provided by the characterization of the acetylating and deacetylating enzymes, and their association with other proteins that are involved in specific events of activation and repression. A basic change in our view of histone acetylation was caused by the discovery that HATs are not necessarily dedicated enzymes associated with chromatin: rather it turns out that known activators of transcription have HAT activity.

The connection was established when the catalytic subunit of a group A HAT was identified as a homologue of the yeast regulator protein GCN5. Then it was shown that GCN5 itself has HAT activity (with histones H3 and H4 as substrates). GCN5 is part of an adaptor complex that is necessary for the interaction between certain enhancers and their target promoters. Its HAT activity is required for activation of

the target gene (693).

This enables us to redraw our picture for the action of coactivators as shown in **Figure 23.13**, where RNA polymerase is bound at a hypersensitive site and coactivators are acetylating histones on the nucleosomes in the vicinity (692). Many examples are now known of interactions of this type.

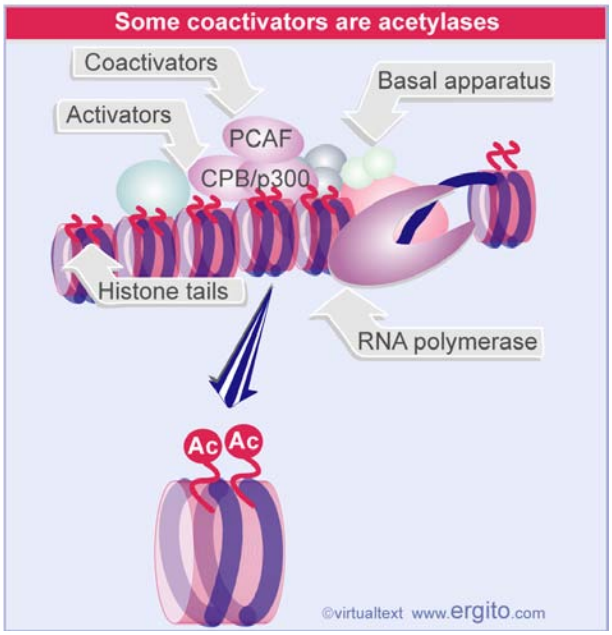


Figure 23.13 Coactivators may have HAT activities that acetylate the tails of nucleosomal histones.

GCN5 leads us into one of the most important acetylase complexes. In yeast, GCN5 is part of the 1.8 MDa SAGA complex, which contains several proteins that are involved in transcription (for review see 1969). Among these proteins are several TAF_{II}s (696). Also, the $TAF_{II}145$ subunit of $TF_{II}D$ is an acetylase. There are some functional overlaps between $TF_{II}D$ and SAGA, most notably that yeast can manage with either $TAF_{II}145$ or GCN5, but is damaged by the deletion of both. This suggests that an acetylase activity is essential for gene expression, but can be provided by either $TF_{II}D$ or SAGA (1062). As might be expected from the size of the SAGA complex, acetylation is only one of its functions, although its other functions in gene activation are less well characterized.

One of the first general activators to be characterized as an HAT was p300/CBP. (Actually, p300 and CBP are different proteins, but they are so closely related that they are often referred to as a single type of activity.) p300/CBP is a coactivator that links an activator to the basal apparatus (see **Figure 22.8**). p300/CBP interacts with various activators, including hormone receptors, AP-1 (c-Jun and c-Fos), and MyoD. The interaction is inhibited by the viral regulator proteins adenovirus E1A and SV40 T antigen, which bind to p300/CBP to prevent the interaction with transcription factors; this explains how these viral proteins inhibit cellular transcription. (This inhibition is important for the ability of the viral proteins to contribute to the tumorigenic state; see *Molecular Biology 6.30.18 Oncoproteins may regulate gene*

expression).

p300/CBP acetylates the N-terminal tails of H4 in nucleosomes. Another coactivator, called PCAF, preferentially acetylates H3 in nucleosomes. p300/CBP and PCAF form a complex that functions in transcriptional activation. In some cases yet another HAT is involved: the coactivator ACTR, which functions with hormone receptors, is itself an HAT that acts on H3 and H4, and also recruits both p300/CBP and PCAF to form a coactivating complex. One explanation for the presence of multiple HAT activities in a coactivating complex is that each HAT has a different specificity, and that multiple different acetylation events are required for activation.

A general feature of acetylation is that an HAT is part of a large complex. **Figure 23.14** shows a simplified model for their behavior. Typically the complex will contain a targeting subunit(s) that determines the binding sites on DNA. This determines the target for the HAT. The complex also contains effector subunits that affect chromatin structure or act directly on transcription. Probably at least some of the effectors require the acetylation event in order to act. Deacetylation, catalyzed by an HDAC, may work in a similar way.

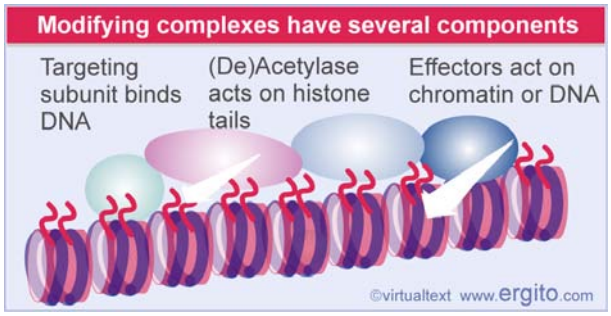


Figure 23.14 Complexes that modify chromatin structure or activity have targeting subunits that determine their sites of action, HAT or HDAC enzymes that acetylate or deacetylate histones, and effector subunits that have other actions on chromatin or DNA.

Acetylation occurs at both replication (when it is transient) and at transcription (when it is maintained while the gene is active). Is it playing the same role in each case? One possibility is that the important effect is on nucleosome structure. Acetylation may be necessary to "loosen" the nucleosome core. At replication, acetylation of histones could be necessary to allow them to be incorporated into new cores more easily. At transcription, a similar effect could be necessary to allow a related change in structure, possibly even to allow the histone core to be displaced from DNA. Alternatively, acetylation could generate binding sites for other proteins that are required for transcription. In either case, deacetylation would reverse the effect.

Is the effect of acetylation quantitative or qualitative? One possibility is that a certain number of acetyl groups are required to have an effect, and the exact positions at which they occur are largely irrelevant. An alternative is that individual acetylation events have specific effects. We might interpret the existence of complexes containing multiple HAT activities in either way – if individual enzymes have different specificities, we may need multiple activities either to acetylate a sufficient

number of different positions or because the individual events are necessary for different effects upon transcription. At replication, it appears, at least with respect to histone H4, that acetylation at any two of three available positions is adequate, favoring a quantitative model in this case. Where chromatin structure is changed to affect transcription, acetylation at specific positions may be important (see *Molecular Biology 5.23.15 Heterochromatin depends on interactions with histones*).

References

692. Chen, H. et al. (1997). *Nuclear receptor coactivator ACTR is a novel histoneacetyltransferase and forms a multimeric activation complex with P/CAF and CP/p300*. Cell 90, 569-580.

693. Brownell, J. E. et al. (1996). *Tetrahymena histone acetyltransferase A: a homologue to yeast Gcn5p linking histone acetylation to gene activation*. Cell 84, 843-851.

696. Grant, P. A. et al. (1998). *A subset of TAF_{II}s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation*. Cell 94, 45-53.

1062. Lee, T. I., Causton, H. C., Holstege, F. C., Shen, W. C., Hannett, N., Jennings, E. G., Winston, F., Green, M. R., and Young, R. A. (2000). *Redundant roles for the TFIID and SAGA complexes in global transcription.* Nature 405, 701-704.

1969. Kingston, R. E. and Narlikar, G. J. (1999). *ATP-dependent remodeling and acetylation as regulators of chromatin fluidity.* Genes Dev. 13, 2339-2352.

CONTROLLING CHROMATIN STRUCTURE

5.23.8 Deacetylases are associated with repressors

Key Concepts

- Deacetylation is associated with repression of gene activity.
- Deacetylases are present in complexes with repressor activity.

In yeast, mutations in *SIN3* and *Rpd3* behave as though these loci repress a variety of genes. The proteins form a complex with the DNA-binding protein Ume6, which binds to the *URS1* element. The complex represses transcription at the promoters containing *URS1*, as illustrated in **Figure 23.15** (3312). Rpd3 has histone deacetylase activity; we do not know whether the function of Sin3 is just to bring Rpd3 to the promoter or whether it has an additional role in repression.

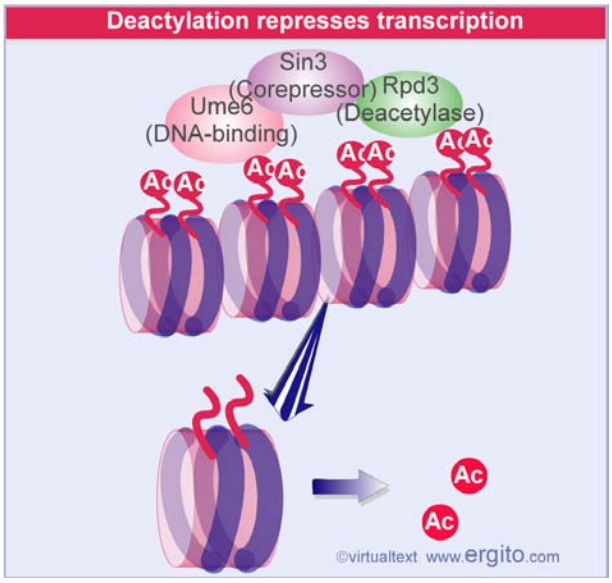


Figure 23.15 A repressor complex contains three components: a DNA binding subunit, a corepressor, and a histone deacetylase.

A similar system for repression is found in mammalian cells (3313; 3314). The bHLH family of transcription regulators includes activators that function as heterodimers, including MyoD (see *Molecular Biology 5.22.15 Helix-loop-helix proteins interact by combinatorial association*). It also includes repressors, in particular the heterodimer Mad:Max, where Mad can be any one of a group of closely related proteins. The Mad:Max heterodimer (which binds to specific DNA sites) interacts with a homologue of Sin3 (called mSin3 in mouse and hSin3 in man). mSin3 is part of a repressive complex that includes histone binding proteins and the histone deacetylases HDAC1 and HDAC2. Deacetylase activity is required for repression. The modular nature of this system is emphasized by other means of

employment: a corepressor (SMRT), which enables retinoid hormone receptors to repress certain target genes, functions by binding mSin3, which in turns brings the HDAC activities to the site. Another means of bringing HDAC activities to the site may be a connection with MeCP2, a protein that binds to methylated cytosines (see *Molecular Biology 5.21.19 CpG islands are regulatory targets*).

Absence of histone acetylation is also a feature of heterochromatin. This is true of both constitutive heterochromatin (typically involving regions of centromeres or telomeres) and facultative heterochromatin (regions that are inactivated in one cell although they may be active in another). Typically the N-terminal tails of histones H3 and H4 are not acetylated in heterochromatic regions (for review see 2414).

Last updated on 9-5-2001

Reviews

2414. Richards, E. J., Elgin, S. C., and Richards, S. C. (2002). *Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects.* Cell 108, 489-500.

References

3312. Kadosh, D. and Struhl, K. (1997). *Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters.* Cell 89, 365-371.

3313. Ayer, D. E., Lawrence, Q. A., and Eisenman, R. N. (1995). *Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3.* Cell 80, 767-776.

3314. Schreiber-Agus, N., Chin, L., Chen, K., Torres, R., Rao, G., Guida, P., Skoultchi, A. I., and DePinho, R. A. (1995). *An amino-terminal domain of Mxi1 mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor SIN3.* Cell 80, 777-786.

CONTROLLING CHROMATIN STRUCTURE

5.23.9 Methylation of histones and DNA is connected

Key Concepts

- Methylation of both DNA and histones is a feature of inactive chromatin.
- The two types of methylation event may be connected.

Methylation of both histones and DNA is associated with inactivity (for review see 2414). Sites that are methylated in histones include two lysines in the tail of H3 and an arginine in the tail of H4.

Methylation of H3 ^{9}Lys is a feature of condensed regions of chromatin, including heterochromatin as seen in bulk and also smaller regions that are known not to be expressed. The histone methyltransferase enzyme that targets this lysine is called SUV39H1 (2014). (We see the origin of this peculiar name in *Molecular Biology 5.23.14 Some common motifs are found in proteins that modify chromatin*). Its catalytic site has a region called the SET domain. Other histone methyltransferases act on arginine (for review see 2237). In addition, methylation may occur on ^{79}Lys in the globular core region of H3; this may be necessary for the formation of heterochromatin at telomeres (3215).

Most of the methylation sites in DNA are CpG islands (see *Molecular Biology 5.21.19 CpG islands are regulatory targets*). CpG sequences in heterochromatin are usually methylated. Conversely, it is necessary for the CpG islands located in promoter regions to be unmethylated in order for a gene to be expressed (see *Molecular Biology 5.21.18 Gene expression is associated with demethylation*).

Methylation of DNA and methylation of histones may be connected. Some histone methyltransferase enzymes contain potential binding sites for the methylated CpG doublet, raising the possibility that a methylated DNA sequence may cause a histone methyltransferase to bind. A possible connection in the opposite direction is indicated by the fact that in the fungus *Neurospora*, the methylation of DNA is prevented by a mutation in a gene coding for a histone methylase that acts on ^{9}Lys of histone H3 (2183). This suggests that methylation of the histone is a signal involved in recruiting the DNA methylase to chromatin. The important point is not the detailed order of events – which remains to be worked out – but the fact that one type of modification can be the trigger for another.

Last updated on 1-3-2003

Reviews

2237. Zhang, Y. and Reinberg, D. (2001). *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev. 15, 2343-2360.

2414. Richards, E. J., Elgin, S. C., and Richards, S. C. (2002). *Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects.* Cell 108, 489-500.

References

2014. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Sun, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature 406, 593-599.

2183. Tamaru, H. and Selker, E. U. (2001). *A histone H3 methyltransferase controls DNA methylation in Neurospora crassa.* Nature 414, 277-283.

3215. Ng, H. H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002). *Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association.* Genes Dev. 16, 1518-1527.

CONTROLLING CHROMATIN STRUCTURE

5.23.10 Chromatin states are interconverted by modification

Key Concepts

- Acetylation of histones is associated with gene activation.
- Methylation of DNA and of histones is associated with heterochromatin.

Figure 23.16 summarizes three types of differences that are found between active chromatin and inactive chromatin:

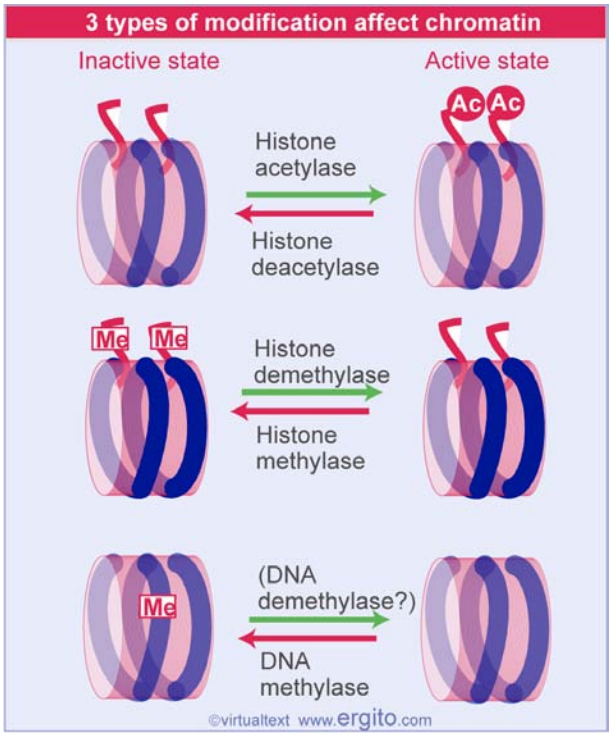


Figure 23.16 Acetylation of histones activates chromatin, and methylation of DNA and histones inactivates chromatin.

- Active chromatin is acetylated on the tails of histones H3 and H4.
- Inactive chromatin is methylated on ^{9}Lys of histone H3.
- Inactive chromatin is methylated on cytosines of CpG doublets.

The reverse types of events occur if we compare the activation of a promoter with the generation of heterochromatin. The actions of the enzymes that modify chromatin

ensure that activating events are mutually exclusive with inactivating events. Methylation of H3 ^{9}Lys and acetylation of H3 ^{14}Lys are mutually antagonistic.

Acetylases and deacetylases may trigger the initiating events. Deacetylation allows methylation to occur, which causes formation of a heterochromatic complex (see *Molecular Biology 5.23.15 Heterochromatin depends on interactions with histones*). Acetylation marks a region as active (see *Molecular Biology 5.23.11 Promoter activation involves an ordered series of events*).

CONTROLLING CHROMATIN STRUCTURE

5.23.11 Promoter activation involves an ordered series of events

Key Concepts

- The remodeling complex may recruit the acetylating complex.
- Acetylation of histones may be the event that maintains the complex in the activated state.

How are acetylases (or deacetylases) recruited to their specific targets? As we have seen with remodeling complexes, the process is likely to be indirect. A sequence-specific activator (or repressor) may interact with a component of the acetylase (or deacetylase) complex to recruit it to a promoter.

There may also be direct interactions between remodeling complexes and histone-modifying complexes. Binding by the SWI/SNF remodeling complex may lead in turn to binding by the SAGA acetylase complex (1966). Acetylation of histones may then in fact stabilize the association with the SWI/SNF complex, making a mutual reinforcement of the changes in the components at the promoter (1967).

We can connect all of the events at the promoter into the series summarized in **Figure 23.17**. The initiating event is binding of a sequence-specific component (which is able to find its target DNA sequence in the context of chromatin). This recruits a remodeling complex. Changes occur in nucleosome structure. An acetylating complex binds, and the acetylation of target histones provides a covalent mark that the locus has been activated.

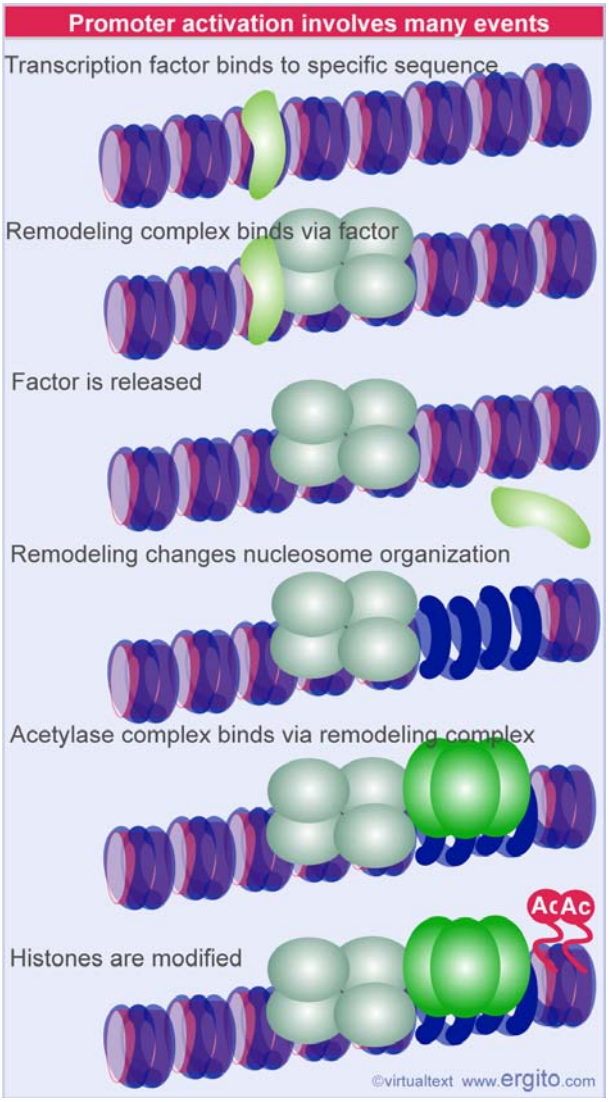


Figure 23.17 Promoter activation involves binding of a sequence-specific activator, recruitment and action of a remodeling complex, and recruitment and action of an acetylating complex.

Modification of DNA also occurs at the promoter. Methylation of cytosine at CpG doublets is associated with gene inactivity (see *Molecular Biology 5.21.18 Gene expression is associated with demethylation*). The basis for recognition of DNA as a target for methylation is not very well established (see *Molecular Biology 5.23.20 DNA methylation is responsible for imprinting*).

It is clear that chromatin remodeling at the promoter requires a variety of changes that affect nucleosomes, including acetylation, but what changes are required within the gene to allow an RNA polymerase to traverse it? We know that RNA polymerase can transcribe DNA *in vitro* at rates comparable to the *in vivo* rate (~25 nucleotides per second) only with template of free DNA. Several proteins have been characterized for their abilities to improve the speed with which RNA polymerase transcribes chromatin *in vivo* (1427; 1428; 1429). The common feature is that they

act on chromatin (for review see 1426). A current model for their action is that they associate with RNA polymerase and travel with it along the template, modifying nucleosome structure by acting on histones. Among these factors are histone acetylases. One possibility is that the first RNA polymerase to transcribe a gene is a pioneer polymerase carrying factors that change the structure of the transcription unit so as to make it easier for subsequent polymerases.

Last updated on 12-5-2001

Reviews

1426. Orphanides, G. and Reinberg, D. (2000). *RNA polymerase II elongation through chromatin*. Nature 407, 471-475.

References

1427. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998). *FACT, a factor that facilitates transcript elongation through nucleosomes.* Cell 92, 105-116.

1428. Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F., Buratowski, S., and Handa, H. (1998). *DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs.* Genes Dev. 12, 343-356.

1429. Bortvin, A. and Winston, F. (1996). *Evidence that Spt6p controls chromatin structure by a direct interaction with histones.* Science 272, 1473-1476.

1966. Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). *Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter.* Cell 97, 299-311.

1967. Hassan, A. H., Neely, K. E., and Workman, J. L. (2001). *Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes.* Cell 104, 817-827.

CONTROLLING CHROMATIN STRUCTURE

5.23.12 Histone phosphorylation affects chromatin structure

Key Concepts

- At least two histones are targets for phosphorylation, possibly with opposing effects.

Histones are phosphorylated in two circumstances:

- cyclically during the cell cycle;
- and in association with chromatin remodeling.

It is has been known for a very long time that histone H1 is phosphorylated at mitosis, and more recently it was discovered that H1 is an extremely good substrate for the Cdc2 kinase that controls cell division. This led to speculations that the phosphorylation might be connected with the condensation of chromatin, but so far no direct effect of this phosphorylation event has been demonstrated, and we do not know whether it plays a role in cell division (see *Molecular Biology 6.29.7 Protein phosphorylation and dephosphorylation control the cell cycle*).

Loss of a kinase that phosphorylates histone H3 on ^{10}Ser has devastating effects on chromatin structure. **Figure 23.18** compares the usual extended structure of the polytene chromosome set of *D. melanogaster* (upper photograph) with the structure that is found in a null mutant that has no JIL-1 kinase (lower photograph). The absence of JIL-1 is lethal, but the chromosomes can be visualized in the larvae before they die (2192).

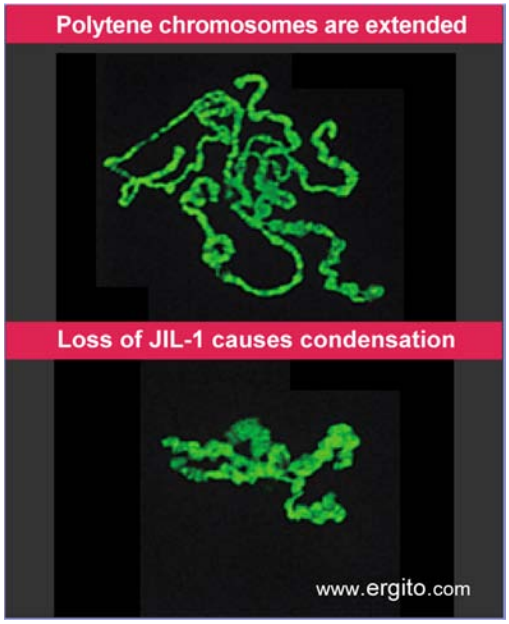


Figure 23.18 Polytene chromosomes of flies that have no JIL-1 kinase have abnormal polytene chromosomes that are condensed instead of extended. Photograph kindly provided by Kristen M. Johansen.

The cause of the disruption of structure is most likely the failure to phosphorylate histone H3 (of course, JIL-1 may also have other targets). This suggests that H3 phosphorylation is required to generate the more extended chromosome structure of euchromatic regions. Evidence supporting the idea that JIL-1 acts directly on chromatin is that it associates with the complex of proteins that binds to the X chromosome to increase its gene expression in males (see *Molecular Biology 5.23.17 X chromosomes undergo global changes*).

This leaves us with somewhat conflicting impressions of the roles of histone phosphorylation. If it is important in the cell cycle, it is likely to be as a signal for condensation. Its effect in chromatin remodeling appears to be the opposite. It is of course possible that phosphorylation of different histones, or even of different amino acid residues in one histone, has opposite effects on chromatin structure.

Last updated on 12-10-2001

References

2192. Wang, Y., Zhang, W., Jin, Y., Johansen, J., and Johansen, K. M. (2001). *The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in Drosophila.* Cell 105, 433-443.

CONTROLLING CHROMATIN STRUCTURE

5.23.13 Heterochromatin propagates from a nucleation event

Key Terms

Epigenetic changes influence the phenotype without altering the genotype. They consist of changes in the properties of a cell that are inherited but that do not represent a change in genetic information.

Position effect variegation (PEV) is silencing of gene expression that occurs as the result of proximity to heterochromatin.

Telomeric silencing describes the repression of gene activity that occurs in the vicinity of a telomere.

Key Concepts

- Heterochromatin is nucleated at a specific sequence and the inactive structure propagates along the chromatin fiber.
- Genes within regions of heterochromatin are inactivated.
- Because the length of the inactive region varies from cell to cell, inactivation of genes in this vicinity causes position effect variegation.
- Similar spreading effects occur at telomeres and at the silent cassettes in yeast mating type.

An interphase nucleus contains both euchromatin and heterochromatin. The condensation state of heterochromatin is close to that of mitotic chromosomes. Heterochromatin is inert. It remains condensed in interphase, is transcriptionally repressed, replicates late in S phase, and may be localized to the nuclear periphery. Centromeric heterochromatin typically consists of satellite DNAs. However, the formation of heterochromatin is not rigorously defined by sequence. When a gene is transferred, either by a chromosomal translocation or by transfection and integration, into a position adjacent to heterochromatin, it may become inactive as the result of its new location, implying that it has become heterochromatic.

Such inactivation is the result of an **epigenetic** effect (see *Molecular Biology 5.23.22 Epigenetic effects can be inherited*). It may differ between individual cells in an animal, and results in the phenomenon of **position effect variegation** (PEV), in which genetically identical cells have different phenotypes. This has been well characterized in *Drosophila*. **Figure 23.19** shows an example of position effect variegation in the fly eye, in which some regions lack color while others are red, because the *white* gene is inactivated by adjacent heterochromatin in some cells, while it remained active in other cells.

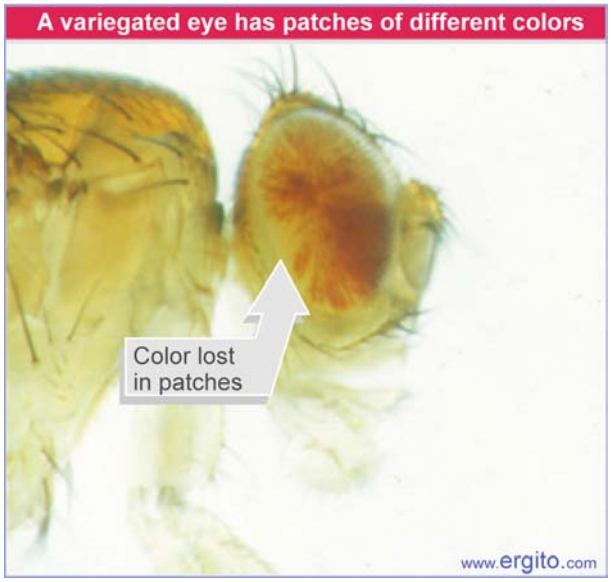


Figure 23.19 Position effect variegation in eye color results when the *white* gene is integrated near heterochromatin. Cells in which *white* is inactive give patches of white eye, while cells in which *white* is active give red patches. The severity of the effect is determined by the closeness of the integrated gene to heterochromatin. Photograph kindly provided by Steve Henikoff.

The explanation for this effect is shown in **Figure 23.20**. Inactivation spreads from heterochromatin into the adjacent region for a variable distance. In some cells it goes far enough to inactivate a nearby gene, but in others it does not. This happens at a certain point in embryonic development, and after that point the state of the gene is inherited by all the progeny cells. Cells descended from an ancestor in which the gene was inactivated form patches corresponding to the phenotype of loss-of-function (in the case of *white*, absence of color).

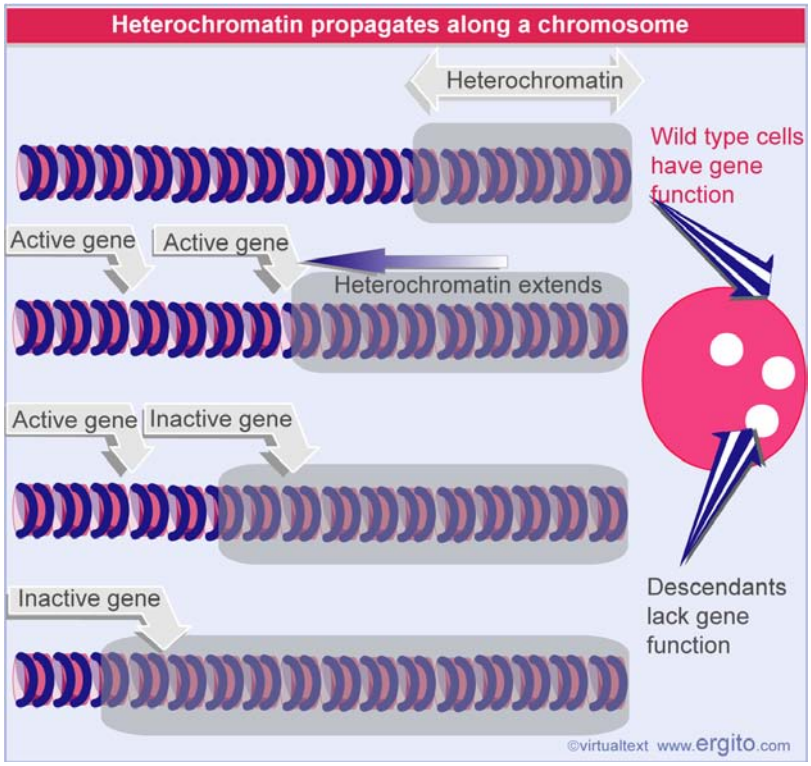


Figure 23.20 Extension of heterochromatin inactivates genes. The probability that a gene will be inactivated depends on its distance from the heterochromatin region.

The closer a gene lies to heterochromatin, the higher the probability that it will be inactivated. This suggests that the formation of heterochromatin may be a two-stage process: a *nucleation* event occurs at a specific sequence; and then the inactive structure *propagates* along the chromatin fiber. The distance for which the inactive structure extends is not precisely determined, and may be stochastic, being influenced by parameters such as the quantities of limiting protein components. One factor that may affect the spreading process is the activation of promoters in the region; an active promoter may inhibit spreading (2191).

Genes that are closer to heterochromatin are more likely to be inactivated, and will therefore be inactive in a greater proportion of cells. On this model, the boundaries of a heterochromatic region might be terminated by exhausting the supply of one of the proteins that is required.

The effect of **telomeric silencing** in yeast is analogous to position effect variegation in *Drosophila;* genes translocated to a telomeric location show the same sort of variable loss of activity. This results from a spreading effect that propagates from the telomeres.

A second form of silencing occurs in yeast. Yeast mating type is determined by the activity of a single active locus *(MAT),* but the genome contains two other copies of the mating type sequences (*HML* and *HMR*), which are maintained in an inactive form. The silent loci *HML* and *HMR* share many properties with heterochromatin, and could be regarded as constituting regions of heterochromatin in miniature (see *Molecular Biology 4.18.7 Silent cassettes at HML and HMR are repressed*).

Last updated on 12-10-2001

References

2191. Ahmad, K. and Henikoff, S. (2001). *Modulation of a transcription factor counteracts heterochromatic gene silencing in Drosophila.* Cell 104, 839-847.

CONTROLLING CHROMATIN STRUCTURE

5.23.14 Some common motifs are found in proteins that modify chromatin

Key Concepts

- The chromo domain is found in several chromatin proteins that have either activating or repressing effects on gene expression
- The SET domain is part of the catalytic site of protein methyltransferases.

Our insights into the molecular mechanisms for controlling the structure of chromatin start with mutants that affect position effect variegation. Some 30 genes have been identified in *Drosophila*. They are named systematically as *Su(var)* for genes whose products act to suppress variegation and *E(var)* for genes whose products enhance variegation. Remember that the genes were named for the behavior of the mutant loci. *Su(var)* mutations lie in genes whose products are needed for the formation of heterochromatin. They include enzymes that act on chromatin, such as histone deacetylases, and proteins that are localized to heterochromatin. *E(var)* mutations lie in genes whose products are needed to activate gene expression. They include members of the SWI/SNF complex. We see immediately from these properties that modification of chromatin structure is important for controlling the formation of heterochromatin. The universality of these mechanisms is indicated by the fact that many of these loci have homologues in yeast that display analogous properties. Some of the homologues in *S. pombe* are *clr* (cryptic loci regulator) genes, in which mutations affect silencing.

Many of the Su(var) and E(var) proteins have a common protein motif of 60 amino acids called the chromo domain. The fact that this domain is found in proteins of both groups suggests that it represents a motif that participates in protein-protein interactions with targets in chromatin (for summary see 2026). Chromo domain(s) are mostly responsible for targeting proteins to heterochromatin. They function by recognizing methylated lysines in histone tails (see *Molecular Biology 5.23.15 Heterochromatin depends on interactions with histones* and *Molecular Biology 5.23.16 Polycomb and trithorax are antagonistic repressors and activators*)

Su(var)3-9 has a chromo domain and also a SET domain, a motif that is found in several Su(var) proteins. Its mammalian homologues localize to centromeric heterochromatin. It is the histone methyltransferase that acts on ^{9}Lys of histone H3 (see *Molecular Biology 5.23.9 Methylation of histones and DNA is connected*). The SET domain is part of the active site, and in fact is a marker for the methylase activity.

The bromo domain is found in a variety of proteins that interact with chromatin, including histone acetylases. The crystal structure shows that it has a binding site for acetylated lysine (2028). The bromo domain itself recognizes only a very short sequence of 4 amino acids including the acetylated lysine, so specificity for target

recognition must depend on interactions involving other regions (2032). Besides the acetylases, the bromo domain is found in a range of proteins that interact with chromatin, including components of the transcription apparatus. This implies that it is used to recognize acetylated histones, which means that it is likely to be found in proteins that are involved with gene activation.

Although there is a general correlation in which active chromatin is acetylated while inactive chromatin is methylated on histones, there are some exceptions to the rule. The best characterized is that acetylation of ^{12}Lys of H4 is associated with heterochromatin (616).

Multiple modifications may occur on the same histone tail, and one modification may influence another. Phosphorylation of a lysine at one position may be necessary for acetylation of a lysine at another position. **Figure 23.21** shows the situation in the tail of H3, which can exist in either of two alternative states. The inactive state has Methyl-^{9}Lys. The active state has Acetyl-^{9}Lys and Phospho-^{10}Ser. These states can be maintained over extended regions of chromatin (2025). The phosphorylation of ^{10}Ser and the methylation of ^{9}Lysare mutually inhibitory, suggesting the order of events shown in the figure. This situation may cause the tail to flip between the active and active states.

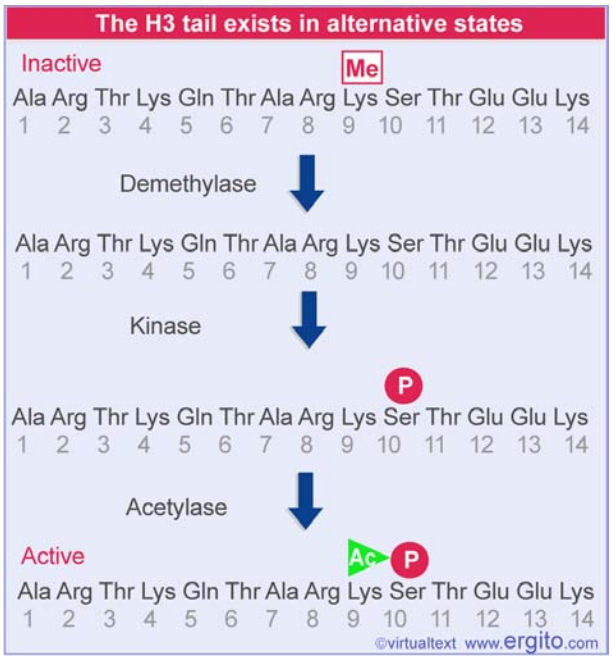


Figure 23.21 Multiple modifications in the H3 tail affect chromatin activity.

Last updated on 9-24-2003

References

616. Turner, B. M., Birley, A. J., and Lavender, J. (1992). *Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei*. Cell 69, 375-384.

2025. Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D., and Felsenfeld, G. (2001). *Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus.* Science 293, 2453-2455.

2026. Koonin, E. V., Zhou, S., and Lucchesi, J. C. (1995). *The chromo superfamily: new members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin.* Nucleic Acids Res. 23, 4229-4233.

2028. Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999). *Structure and ligand of a histone acetyltransferase bromodomain.* Nature 399, 491-496.

2032. Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. (2000). *The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p.* EMBO J. 19, 6141-6149.

CONTROLLING CHROMATIN STRUCTURE

5.23.15 Heterochromatin depends on interactions with histones

Key Concepts

- HP1 is the key protein in forming mammalian heterochromatin, and acts by binding to methylated H3 histone.
- RAP1 initiates formation of heterochromatin in yeast by binding to specific target sequences in DNA.
- The targets of RAP1 include telomeric repeats and silencers at *HML* and *HMR*.
- RAP1 recruits SIR3/SIR4, which interact with the N-terminal tails of H3 and H4.

Inactivation of chromatin occurs by the addition of proteins to the nucleosomal fiber. The inactivation may be due to a variety of effects, including condensation of chromatin to make it inaccessible to the apparatus needed for gene expression, addition of proteins that directly block access to regulatory sites, or proteins that directly inhibit transcription.

Two systems that have been characterized at the molecular level involve HP1 in mammals and the SIR complex in yeast (for review see 2420). Although there are no detailed similarities between the proteins involved in each system, the general mechanism of reaction is similar: the points of contact in chromatin are the N-terminal tails of the histones.

HP1 (heterochromatin protein 1) is one of most important Su(var) proteins. This was originally identified as a protein that is localized to heterochromatin by staining polytene chromosomes with an antibody directed against the protein (2021). It was later shown to be the product of the gene *Su(var)2-5* (2022). Its homologue in the yeast *S. pombe* is coded by *swi6*. The original protein identified as HP1 is now called HP1 α, since two related proteins, HP1 β and HP1 γ, have since been found.

HP1 contains a chromo domain near the N-terminus, and another domain that is related to it, called the chromo-shadow domain, at the C-terminus (see **Figure 23.23**). The importance of the chromo domain is indicated by the fact that it is the location of many of the mutations in HP1 (2023).

Mutation of a deacetylase that acts on the H3 Ac-^{14}Lys prevents the methylation at ^{9}Lys (2238). H3 that is methylated at ^{9}Lys binds the protein HP1 via the chromo domain (2016, 2015, (4178, 4180). This suggests the model for initiating formation of heterochromatin shown in **Figure 23.22**. First the deacetylase acts to remove the modification at ^{14}Lys. Then the SUV39H1 methylase acts on the histone H3 tail to create the methylated signal to which HP1 will bind. **Figure 23.23** expands the reaction to show that the interaction occurs between the chromo domain and the methylated lysine. This is a trigger for forming inactive chromatin. **Figure 23.24**

shows that the inactive region may then be extended by the ability of further HP1 molecules to interact with one another (for review see 2237).

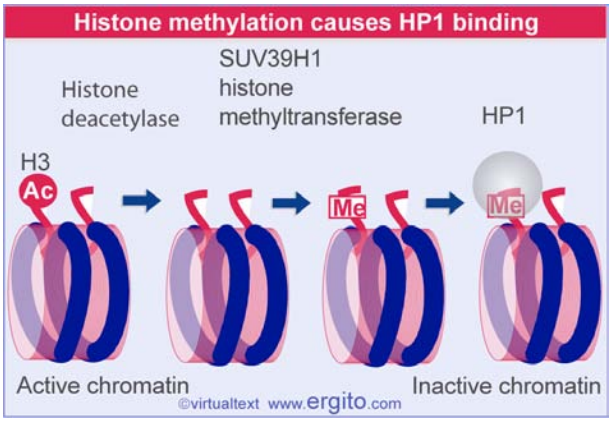


Figure 23.22 SUV39H1 is a histone methyltransferase that acts on ^{9}Lys of histone H3. HP1 binds to the methylated histone.

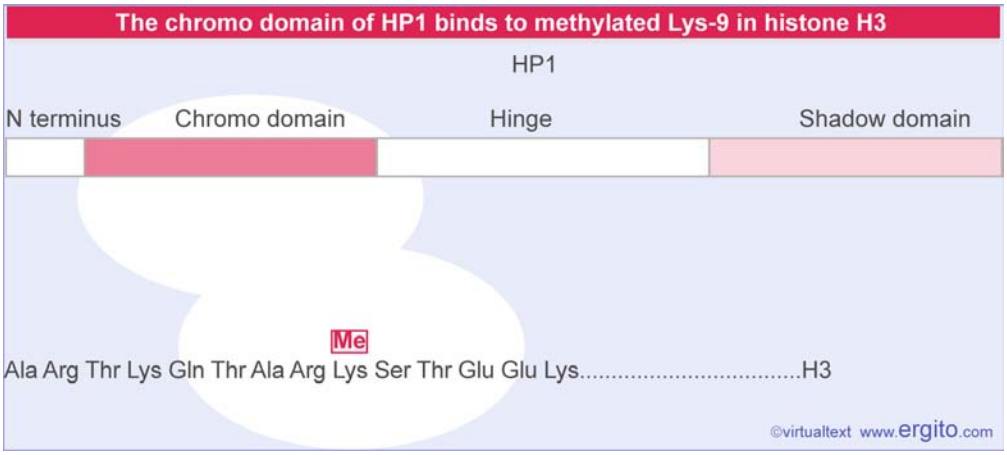


Figure 23.23 Methylation of histone H3 creates a binding site for HP1.

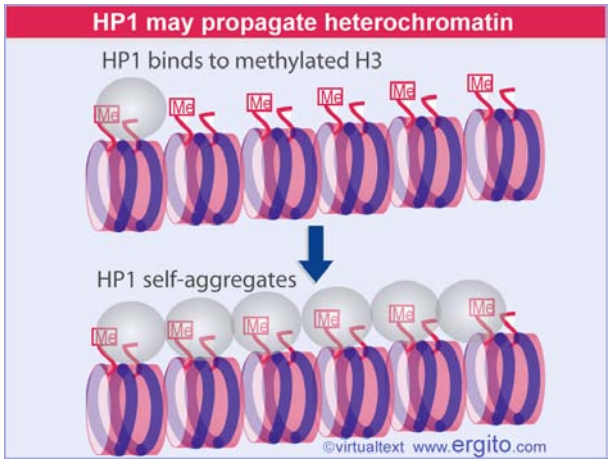


Figure 23.24 Binding of HP1 to methylated histone H3 forms a trigger for silencing because further molecules of HP1 aggregate on the nucleosome chain.

The existence of a common basis for silencing in yeast is suggested by its reliance on a common set of genetic loci. Mutations in any one of a number of genes cause *HML* and *HMR* to become activated, and also relieve the inactivation of genes that have been integrated near telomeric heterochromatin. The products of these loci therefore function to maintain the inactive state of both types of heterochromatin (for review see 210, 4530).

Figure 23.25 proposes a model for actions of these proteins. Only one of them is a sequence-specific DNA-binding protein. This is RAP1, which binds to the $C_{1-3}A$ repeats at the telomeres, and also binds to the *cis*-acting silencer elements that are needed for repression of *HML* and *HMR* (619). The proteins SIR3 and SIR4 interact with RAP1 and also with one another (they may function as a heteromultimer). SIR3/SIR4 interact with the N-terminal tails of the histones H3 and H4. [In fact, the first evidence that histones might be involved directly in formation of heterochromatin was provided by the discovery that mutations abolishing silencing at *HML/HMR* map to genes coding for H3 and H4 (620; for review see 209)].

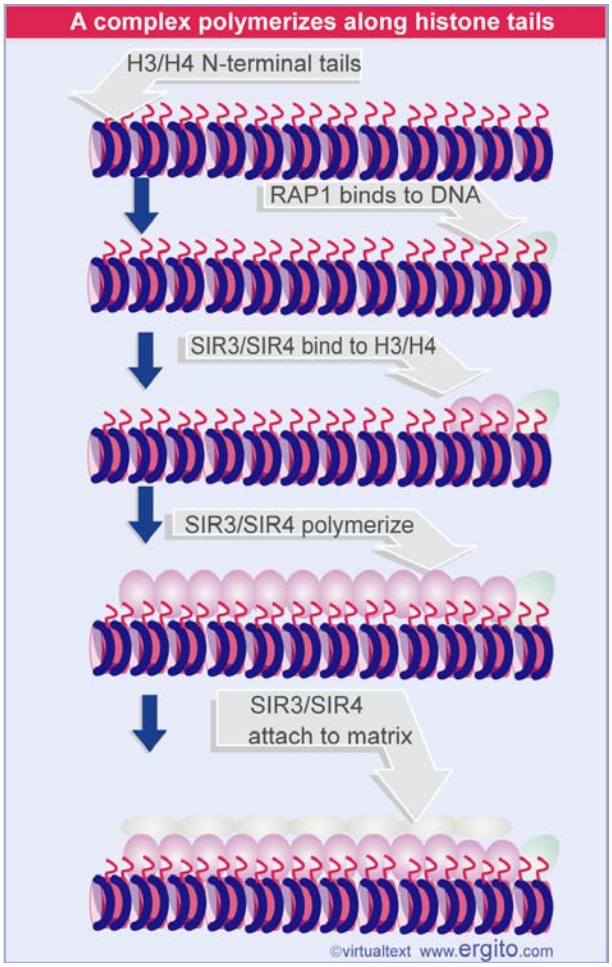


Figure 23.25 Formation of heterochromatin is initiated when RAP1 binds to DNA. SIR3/4 bind to RAP1 and also to histones H3/H4. The complex polymerizes along chromatin and may connect telomeres to the nuclear matrix.

RAP1 has the crucial role of identifying the DNA sequences at which heterochromatin forms. It recruits SIR3/SIR4, and they interact directly with the histones H3/H4 (622; 1214). Once SIR3/SIR4 have bound to histones H3/H4, the complex may polymerize further, and spread along the chromatin fiber. This may inactivate the region, either because coating with SIR3/SIR4 itself has an inhibitory effect, or because binding to histones H3/H4 induces some further change in structure. We do not know what limits the spreading of the complex. The C-terminus of SIR3 has a similarity to nuclear lamin proteins (constituents of the nuclear matrix) and may be responsible for tethering heterochromatin to the nuclear periphery (621; 623).

A similar series of events forms the silenced regions at HMR and HML (see also *Molecular Biology 4.18.7 Silent cassettes at HML and HMR are repressed*). Three sequence-specific factors are involved in triggering formation of the complex: RAP1, ABF1 (a transcription factor), and ORC (the origin replication complex). In this case, SIR1 binds to a sequence-specific factor and recruits SIR2,3,4 to form the repressive structure. SIR2 is a histone deacetylase (2416; 2418; 2419). The deacetylation

reaction is necessary to maintain binding of the SIR complex to chromatin (for review see 2420).

How does a silencing complex repress chromatin activity? It could condense chromatin so that regulator proteins cannot find their targets. The simplest case would be to suppose that the presence of a silencing complex is mutually incompatible with the presence of transcription factors and RNA polymerase. The cause could be that silencing complexes block remodeling (and thus indirectly prevent factors from binding) or that they directly obscure the binding sites on DNA for the transcription factors. However, the situation may not be this simple, because transcription factors and RNA polymerase can be found at promoters in silenced chromatin (1963). This could mean that the silencing complex prevents the factors from working rather than from binding as such. In fact, there may be competition between gene activators and the repressing effects of chromatin, so that activation of a promoter inhibits spread of the silencing complex (2191).

Another specialized chromatin structure forms at the centromere. Its nature is suggested by the properties of an *S. cerevisiae* mutation, *cse4*, that disrupts the structure of the centromere. Cse4p is a protein that is related to histone H3. A mammalian centromeric protein, CENP-A, has a related sequence. Genetic interactions between *cse4* and CDE-II, and between *cse4* and a mutation in the H4 histone gene, suggest that a histone octamer may form around a core of Cse4p-H4, and then the centromeric complexes CBF1 and CBF3 may attach to form the centromere (591; 624).

The centromere may then be associated with the formation of heterochromatin in the region. In human cells, the centromere-specific protein CENP-B is required to initiate modifications of histone H3 (deacetylation of ^{9}Lys and ^{14}Lys, followed by methylation of ^{9}Lys) that trigger an association with the protein Swi6 that leads to the formation of heterochromatin in the region (3218).

Last updated on 9-24-2003

Reviews

209. Thompson, J. S., Hecht, A., and Grunstein, M. (1993). *Histones and the regulation of heterochromatin in yeast.* Cold Spring Harbor Symp. Quant. Biol. 58, 247-256.

210. Loo, S. and Rine, J. (1995). *Silencing and heritable domains of gene expression*. Annu. Rev. Cell Dev. Biol. 11, 519-548.

2237. Zhang, Y. and Reinberg, D. (2001). *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev. 15, 2343-2360.

2420. Moazed, D. (2001). *Common themes in mechanisms of gene silencing.* Mol. Cell 8, 489-498.

4530. Rusche, L. N., Kirchmaier, A. L., and Rine, J. (2003). *The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae.* Annu. Rev. Biochem. 72, 481-516.

References

591. Bloom, K. S. and Carbon, J. (1982). *Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes*. Cell 29, 305-317.

619. Shore, D. and Nasmyth, K. (1987). *Purification and cloning of a DNA-binding protein from yeast that binds to both silencer and activator elements*. Cell 51, 721-732.

620. Kayne, P. S., Kim, U. J., Han. M., Mullen, R. J., Yoshizaki, F., and Grunstein, M. (1988). *Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast*. Cell 55, 27-39.

621. Palladino, F., Laroche, T., Gilson, E., Axelrod, A., Pillus, L., and Gasser, S. M. (1993). *SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres*. Cell 75, 543-555.

622. Moretti, P., Freeman, K., Coodly, L., and Shore, D. (1994). *Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1*. Genes Dev. 8, 2257-2269.

623. Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S. M., and Grunstein, M. (1995). *Histone H3 and H4 N-termini interact with the silent information regulators SIR3 and SIR4: a molecular model for the formation of heterochromatin in yeast*. Cell 80, 583-592.

624. Meluh, P. B. et al. (1998). *Cse4p is a component of the core centromere of S. cerevisiae.* Cell 94, 607-613.

1214. Manis, J. P., Gu, Y., Lansford, R., Sonoda, E., Ferrini, R., Davidson, L., Rajewsky, K., and Alt, F. W. (1998). *Ku70 is required for late B cell development and immunoglobulin heavy chain class switching.* J. Exp. Med. 187, 2081-2089.

1963. Sekinger, E. A. and Gross, D. S. (2001). *Silenced chromatin is permissive to activator binding and PIC recruitment.* Cell 105, 403-414.

2015. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.* Nature 410, 120-124.

2016. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.* Nature 410, 116-120.

2021. James, T. C. and Elgin, S. C. (1986). *Identification of a nonhistone chromosomal protein associated with heterochromatin in D. melanogaster and its gene.* Mol. Cell Biol. 6, 3862-3872.

2022. Eissenberg, J. C., Morris, G. D., Reuter, G., and Hartnett, T. (1992). *The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation.* Genetics 131, 345-352.

2023. Platero, J. S., Hartnett, T., and Eissenberg, J. C. (1995). *Functional analysis of the chromo domain of HP1*. EMBO J. 14, 3977-3986.

2191. Ahmad, K. and Henikoff, S. (2001). *Modulation of a transcription factor counteracts heterochromatic gene silencing in Drosophila.* Cell 104, 839-847.

2238. Nakayama , J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001). *Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly.* Science 292, 110-113.

2416. Imai, S., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000). *Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase.* Nature 403, 795-800.

2418. Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C., and Boeke, J. D. (2000). *A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein*

family. Proc. Natl. Acad. Sci. USA 97, 6658-6663.

2419. Landry, J., Sutton, A., Tafrov, S. T., Heller, R. C., Stebbins, J., Pillus, L., and Sternglanz, R. (2000). *The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases.* Proc. Natl. Acad. Sci. USA 97, 5807-5811.

3218. Nakagawa, H., Lee, J. K., Hurwitz, J., Allshire, R. C., Nakayama, J., Grewal, S. I., Tanaka, K., and Murakami, Y. (2002). *Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications.* Genes Dev. 16, 1766-1778.

4178. Cheutin, T., McNairn, A. J., Jenuwein, T., Gilbert, D. M., Singh, P. B., and Misteli, T. (2003). *Maintenance of stable heterochromatin domains by dynamic HP1 binding.* Science 299, 721-725.

4180. Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). *Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing.* EMBO J. 21, 1121-1131.

CONTROLLING CHROMATIN STRUCTURE

5.23.16 Polycomb and trithorax are antagonistic repressors and activators

Key Concepts

- Polycomb group proteins (Pc-G) perpetuate a state of repression through cell divisions.
- The PRE is a DNA sequence that is required for the action of Pc-G.
- The PRE provides a nucleation center from which Pc-G proteins propagate an inactive structure.
- No individual Pc-G protein has yet been found that can bind the PRE.
- Trithorax group proteins antagonize the actions of the Pc-G.

Heterochromatin provides one example of the specific repression of chromatin. Another is provided by the genetics of homeotic genes in *Drosophila,* which have led to the identification of a protein complex that may maintain certain genes in a repressed state. *Pc* mutants show transformations of cell type that are equivalent to gain-of-function mutations in the genes *Antennapedia* (*Antp*) or *Ultrabithorax*, because these genes are expressed in tissues in which usually they are repressed. This implicates *Pc* in regulating transcription. Furthermore, *Pc* is the prototype for a class of loci called the *Pc* group (*Pc-G*); mutations in these genes generally have the same result of derepressing homeotic genes, suggesting the possibility that the group of proteins has some common regulatory role. A connection between chromatin remodeling and repression is indicated by the properties of *brahma*, a fly counterpart to *SWI2*, which codes for component of the SWI/SNF remodeling complex. Loss of *brahma* function suppresses mutations in *Polycomb.*

Consistent with the pleiotropy of *Pc* mutations, Pc is a nuclear protein that can be visualized at ~80 sites on polytene chromosomes. These sites include the *Antp* gene. Another member of the *Pc-G*, *polyhomeotic*, is visualized at a set of polytene chromosome bands that are identical with those bound by Pc. The two proteins coimmunoprecipitate in a complex of ~2.5 × 10^6 D that contains 10-15 polypeptides. The relationship between these proteins and the products of the ~30 *Pc-G* genes remains to be established. One possibility is that some of these gene products form a general repressive complex,and then some of the other proteins associate with it to determine its specificity (683; 697).

The Pc-G proteins are not conventional repressors. They are not responsible for determining the initial pattern of expression of the genes on which they act. In the absence of Pc-G proteins, these genes are initially repressed as usual, but later in development the repression is lost without Pc-G group functions. This suggests that *the Pc-G proteins in some way recognize the state of repression when it is established, and they then act to perpetuate it through cell division of the daughter*

cells. **Figure 23.26** shows a model in which Pc-G proteins bind in conjunction with a repressor, but the Pc-G proteins remain bound after the repressor is no longer available. This is necessary to maintain repression, so that if Pc-G proteins are absent, the gene becomes activated (698).

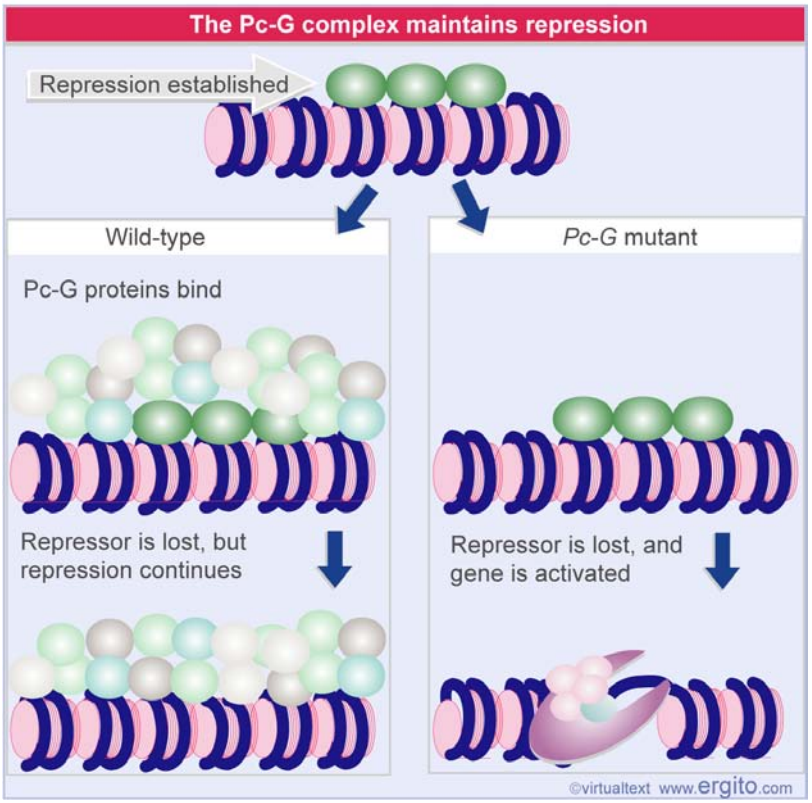


Figure 23.26 Pc-G proteins do not initiate repression, but are responsible for maintaining it.

A region of DNA that is sufficient to enable the response to the *Pc-G* genes is called a PRE (*Polycomb* response element). It can be defined operationally by the property that it maintains repression in its vicinity throughout development. The assay for a PRE is to insert it close to a reporter gene that is controlled by an enhancer that is repressed in early development, and then to determine whether the reporter becomes expressed subsequently in the descendants. An effective PRE will prevent such re-expression (700).

The PRE is a complex structure, ~10 kb. No individual member of the Pc-G proteins has yet been shown to bind to specific sequences in the PRE, so the basis for the assembly of the complex is still unknown. When a locus is repressed by Pc-G proteins, however, the proteins appear to be present over a much larger length of DNA than the PRE itself. Polycomb is found locally over a few kilobases of DNA surrounding a PRE.

This suggests that the PRE may provide a nucleation center, from which a structural state depending on Pc-G proteins may propagate. This model is supported by the observation of effects related to position effect variegation (see **Figure 23.20**), that

is, a gene near to a locus whose repression is maintained by Pc-G may become heritably inactivated in some cells but not others. In one typical situation, crosslinking experiments *in vivo* showed that Pc protein is found over large regions of the *bithorax* complex that are inactive, but the protein is excluded from regions that contain active genes. The idea that this could be due to cooperative interactions within a multimeric complex is supported by the existence of mutations in *Pc* that change its nuclear distribution and abolish the ability of other *Pc-G* members to localize in the nucleus. The role of Pc-G proteins in maintaining, as opposed to establishing, repression must mean that the formation of the complex at the PRE also depends on the local state of gene expression (699).

A connection between the Pc-G complex and more general structural changes in chromatin is suggested by the inclusion of a chromo domain in Pc. (In fact, the chromo domain was first identified as a region of homology between Pc and the protein HP1 found in heterochromatin.) Since variegation is caused by the spreading of inactivity from constitutive heterochromatin, it is likely that the chromo domain is used by Pc and HP1 in a similar way to induce the formation of heterochromatic or inactive structures (see *Molecular Biology 5.23.14 Some common motifs are found in proteins that modify chromatin*). The chromo domain of Pc binds to ^{27}Lys on H3 (4179) (analogous to HP1's use of its chromo domain to bind to ^{9}Lys). This model implies that similar mechanisms are used to repress individual loci or to create heterochromatin. This is probably how the PRC-1 (Polycomb-repressive complex)works. Another complex that contains Polycomb, the Esc-E(z) complex, has a histone methyltransferase activity, carried by E(z), which may target the complex by methylating H3 in the appropriate chromatin locations.

The *trithorax* group (*trxG*) of proteins have the opposite effect to the Pc-G proteins: they act to maintain genes in an active state. There may be some similarities in the actions of the two groups: mutations in some loci prevent both Pc-G and trx from functioning, suggesting that they could rely on common components. A factor coded by the *trithorax-like* gene, called GAGA because it binds to GA-rich consensus sequences, has binding sites in the PRE. In fact, the sites where Pc binds to DNA coincide with the sites where GAGA factor binds (704, 702).

What does this mean? GAGA is probably needed for activating factors, including trxG members, to bind to DNA. Is it also needed for PcG proteins to bind and exercise repression? This is not yet clear, but such a model would demand that something other than GAGA determines which of the alternative types of complex subsequently assemble at the site.

Last updated on 9-24-2003

References

683. Franke, A., DeCamillis, M., Zink, D., Cheng, N., Brock, H. W., and Paro, R. (1992). *Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin ofD. melanogaster*. EMBO J. 11, 2941-29.

697. Zink, B. and Paro, R. (1989). *In vivo binding patterns of a trans-regulator of the homeotic genes in D. melanogaster*. Nature 337, 468-471.

698. Eissenberg, J. C., James, T. C., Fister-Hartnett, D. M., Hartnett, T., Ngan, V., and Elgin, S. C. R. (1990). *Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in D. melanogaster*. Proc. Natl. Acad. Sci. USA 87, 9923-9927.

699. Orlando, V. and Paro, R. (1993). *Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin*. Cell 75, 1187-1198.

700. Chan, C.-S., Rastelli, L., and Pirrotta, V. (1994). *A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression*. EMBO J. 13, 2553-2564.

702. Strutt, H., Cavalli, G., and Paro, R. (1997). *Colocalization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression*. EMBO J. 16, 3621-3632.

704. Geyer, P. K. and Corces, V. G. (1992). *DNA position-specific repression of transcription by a Drosophila zinc finger protein*. Genes Dev. 6, 1865-1873.

4179. Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003). *Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains.* Genes Dev. 17, 1870-1881.

CONTROLLING CHROMATIN STRUCTURE

5.23.17 X chromosomes undergo global changes

Key Terms

Dosage compensation describes mechanisms employed to compensate for the discrepancy between the presence of two X chromosomes in one sex but only one X chromosome in the other sex.

Constitutive heterochromatin describes the inert state of permanently nonexpressed sequences, usually satellite DNA.

Facultative heterochromatin describes the inert state of sequences that also exist in active copies-for example, one mammalian X chromosome in females.

The **single X hypothesis** describes the inactivation of one X chromosome in female mammals.

The **n-1 rule** states that only one X chromosome is active in female mammalian cells; any other(s) are inactivated.

Key Concepts

- One of the two X chromosomes is inactivated at random in each cell during embryogenesis of eutherian mammals.
- In exceptional cases where there are >2 X chromosomes, all but one are inactivated.
- The *Xic* (X inactivation center) is a *cis*-acting region on the X chromosome that is necessary and sufficient to ensure that only one X chromosome remains active.
- *Xic* includes the *Xist* gene which codes for an RNA that is found only on inactive X chromosomes.
- The mechanism that is responsible for preventing *Xist* RNA from accumulating on the active chromosome is unknown.

Sex presents an interesting problem for gene regulation, because of the variation in the number of X chromosomes. If X-linked genes were expressed equally well in each sex, females would have twice as much of each product as males. The importance of avoiding this situation is shown by the existence of **dosage compensation**, which equalizes the level of expression of X-linked genes in the two sexes. Mechanisms used in different species are summarized in **Figure 23.27**:

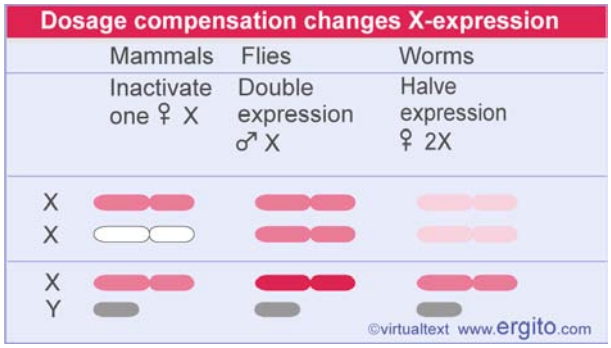


Figure 23.27 Different means of dosage compensation are used to equalize X chromosome expression in male and female.

- In mammals, one of the two female X chromosomes is inactivated completely. The result is that females have only one active X chromosome, which is the same situation found in males. The active X chromosome of females and the single X chromosome of males are expressed at the same level.
- In *Drosophila*, the expression of the single male X chromosome is doubled relative to the expression of each female X chromosome.
- In *C. elegans*, the expression of each female X chromosome is halved relative to the expression of the single male X chromosome.

The common feature in all these mechanisms of dosage compensation is that *the entire chromosome is the target for regulation*. A global change occurs that quantitatively affects all of the promoters on the chromosome. We know most about the inactivation of the X chromosome in mammalian females, where the entire chromosome becomes heterochromatic.

The twin properties of heterochromatin are its condensed state and associated inactivity. It can be divided into two types:

- **Constitutive heterochromatin** contains specific sequences that have no coding function. Typically these include satellite DNAs, and are often found at the centromeres. These regions are invariably heterochromatic because of their intrinsic nature.
- **Facultative heterochromatin** takes the form of entire chromosomes that are inactive in one cell lineage, although they can be expressed in other lineages. The example *par excellence* is the mammalian X chromosome. The inactive X chromosome is perpetuated in a heterochromatic state, while the active X chromosome is part of the euchromatin. So *identical DNA sequences are involved in both states*. Once the inactive state has been established, it is inherited by descendant cells. This is an example of epigenetic inheritance, because it does not depend on the DNA sequence.

Our basic view of the situation of the female mammalian X chromosomes was formed by the **single X hypothesis** in 1961 (see *Great Experiments 11.10 The*

discovery of X-chromosome inactivation). Female mice that are heterozygous for X-linked coat color mutations have a variegated phenotype in which some areas of the coat are wild-type, but others are mutant. **Figure 23.28** shows that this can be explained *if one of the two X chromosomes is inactivated at random in each cell of a small precursor population.* Cells in which the X chromosome carrying the wild-type gene is inactivated give rise to progeny that express only the mutant allele on the active chromosome. Cells derived from a precursor where the other chromosome was inactivated have an active wild-type gene. In the case of coat color, cells descended from a particular precursor stay together and thus form a patch of the same color, creating the pattern of visible variegation. In other cases, individual cells in a population will express one or the other of X-linked alleles; for example, in heterozygotes for the X-linked locus G6PD, any particular red blood cell will express only one of the two allelic forms. [Random inactivation of one X chromosome occurs in eutherian mammals. In marsupials, the choice is directed: it is always the X chromosome inherited from the father that is inactivated (625).]

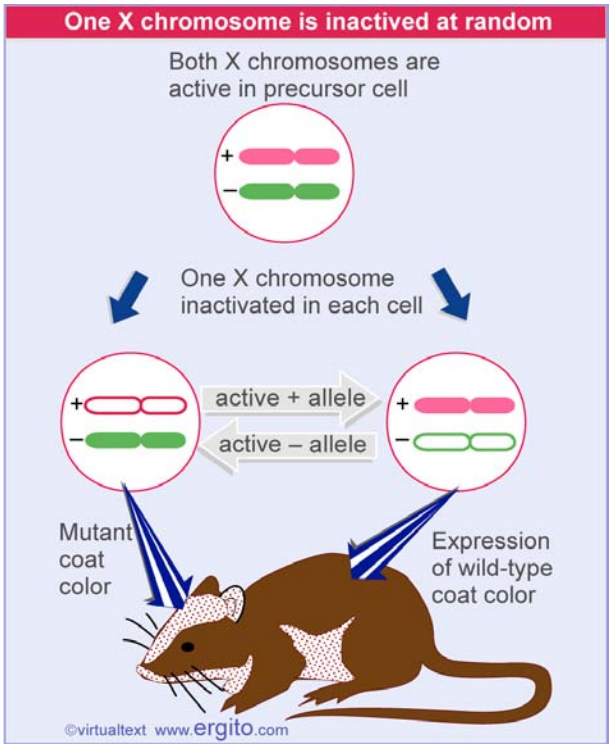


Figure 23.28 X-linked variegation is caused by the random inactivation of one X chromosome in each precursor cell. Cells in which the + allele is on the active chromosome have wild phenotype; but cells in which the – allele is on the active chromosome have mutant phenotype.

Inactivation of the X chromosome in females is governed by the **n-1 rule**: however many X chromosomes are present, all but one will be inactivated. In normal females there are of course 2 X chromosomes, but in rare cases where nondisjunction has generated a 3X or greater genotype, only one X chromosome remains active. This suggests a general model in which a specific event is limited to one X chromosome and protects it from an inactivation mechanism that applies to all the others.

A single locus on the X chromosome is sufficient for inactivation. When a translocation occurs between the X chromosome and an autosome, this locus is present on only one of the reciprocal products, and only that product can be inactivated. By comparing different translocations, it is possible to map this locus, which is called the *Xic* (X-inactivation center). A cloned region of 450 kb contains all the properties of the *Xic*. When this sequence is inserted as a transgene on to an autosome, the autosome becomes subject to inactivation (in a cell culture system) (627).

Xic is a *cis*-acting locus that contains the information necessary to count X chromosomes and inactivate all copies but one. Inactivation spreads from *Xic* along the entire X chromosome. When *Xic* is present on an X chromosome-autosome translocation, inactivation spreads into the autosomal regions (although the effect is not always complete).

Xic contains a gene, called *Xist*, that is expressed only on the *inactive* X chromosome. The behavior of this gene is effectively the opposite from all other loci on the chromosome, which are turned off. Deletion of *Xist* prevents an X chromosome from being inactivated. However, it does not interfere with the counting mechanism (because other X chromosomes can be inactivated). So we can distinguish two features of *Xic*: an unidentified element(s) required for counting; and the *Xist* gene required for inactivation.

Figure 23.29 illustrates the role of *Xist* RNA in X-inactivation (for review see 3237). *Xist* codes for an RNA that lacks open reading frames. The *Xist* RNA "coats" the X chromosome from which it is synthesized, suggesting that it has a structural role. Prior to X-inactivation, it is synthesized by both female X chromosomes. Following inactivation, the RNA is found only on the inactive X chromosome. The transcription rate remains the same before and after inactivation, so the transition depends on post-transcriptional events (626).

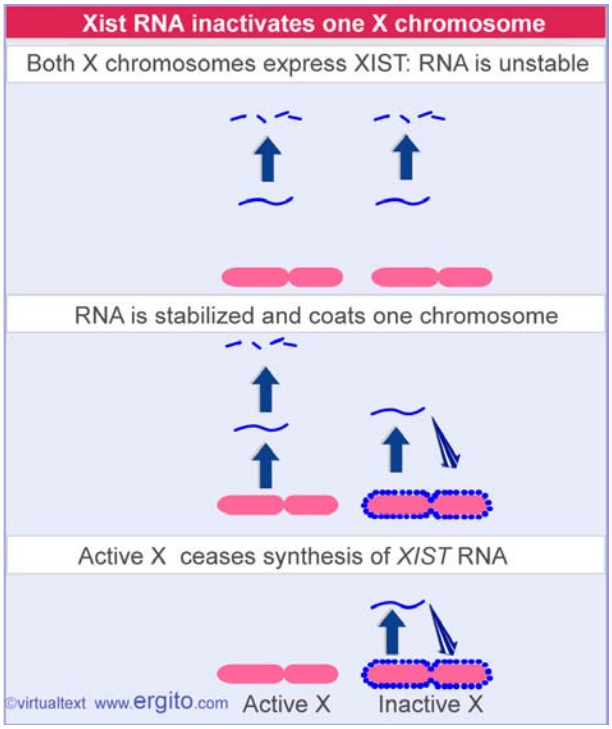


Figure 23.29 X-inactivation involves stabilization of *Xist* RNA, which coats the inactive chromosome.

Prior to X-inactivation, *Xist* RNA decays with a half life of ~2 hr. X-inactivation is mediated by stabilizing the *Xist* RNA on the inactive X chromosome. The *Xist* RNA shows a punctate distribution along the X chromosome, suggesting that association with proteins to form particulate structures may be the means of stabilization. We do not know yet what other factors may be involved in this reaction and how the *Xist* RNA is limited to spreading in *cis* along the chromosome. The characteristic features of the inactive X chromosome, which include a lack of acetylation of histone H4, and methylation of CpG sequences (see *Molecular Biology 5.21.19 CpG islands are regulatory targets*), presumably occur later as part of the mechanism of inactivation (628; 617).

The n–1 rule suggests that stabilization of *Xist* RNA is the "default," and that some blocking mechanism prevents stabilization at one X chromosome (which will be the active X). This means that, although *Xic* is necessary and sufficient for a chromosome to be *inactivated*, the products of other loci may be necessary for the establishment of an *active* X chromosome.

Silencing of *Xist* expression is necessary for the active X. Deletion of the gene for DNA methyltransferase prevents silencing of *Xist*, probably because methylation at the *Xist* promoter is necessary for cessation of transcription.

Reviews

3237. Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A., and Panning, B. (2002). *Xist RNA and the mechanism of x chromosome inactivation.* Annu. Rev. Genet. 36, 233-278.

References

617. Jeppesen, P. and Turner, B. M. (1993). *The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression*. Cell 74, 281-289.

625. Lyon, M. F. (1961). *Gene action in the X chromosome of the mouse*. Nature 190, 372-373.

626. Penny, G. D. et al. (1996). *Requirement for Xist in X chromosome inactivation*. Nature 379, 131-137.

627. Lee, J. T. et al. (1996). *A 450 kb transgene displays properties of the mammalian X-inactivation center*. Cell 86, 83-94.

628. Panning, B., Dausman, J., and Jaenisch, R. (1997). *X chromosome inactivation is mediated by Xist RNA stabilization*. Cell 90, 907-916.

CONTROLLING CHROMATIN STRUCTURE

5.23.18 Chromosome condensation is caused by condensins

Key Terms

Structural maintenance of chromosomes (SMC) describes a group of proteins that include the cohesins, which hold sister chromatids together, and the condensins, which are involved in chromosome condensation.

Condensin proteins are components of a complex that binds to chromosomes to cause condensation for meiosis or mitosis. They are members of the SMC family of proteins.

Cohesin proteins form a complex that holds sister chromatids together. They include some SMC proteins.

Key Concepts

- SMC proteins are ATPases that include the condensins and the cohesins.
- A heterodimer of SMC proteins associates with other subunits.
- The condensins cause chromatin to be more tightly coiled by introducing positive supercoils into DNA.
- Condensins are responsible for condensing chromosomes at mitosis.
- Chromosome-specific condensins are responsible for condensing inactive X chromosomes in *C. elegans*.

The structures of entire chromosomes are influenced by interactions with proteins of the **SMC** (structural maintenance of chromosome) family. They are ATPases that fall into two functional groups (for review see 3445). **Condensins** are involved with the control of overall structure, and are responsible for the condensation into compact chromosomes at mitosis. **Cohesins** are concerned with connections between sister chromatids that must be released at mitosis (see *Molecular Biology 6.29.19 Cohesins hold sister chromatids together*). Both consist of dimers formed by SMC proteins. Condensins form complexes that have a core of the heterodimer SMC2-SMC4 associated with other (non SMC) proteins. Cohesins have a similar organization based on the heterodimeric core of SMC1-SMC3.

Figure 23.31 shows that an SMC protein has a coiled-coil structure in its center, interrupted by a flexible hinge region. Both the amino and carboxyl termini have ATP- and DNA-binding motifs. Different models have been proposed for the actions of these proteins depending on whether they dimerize by intra- or inter-molecular interactions.

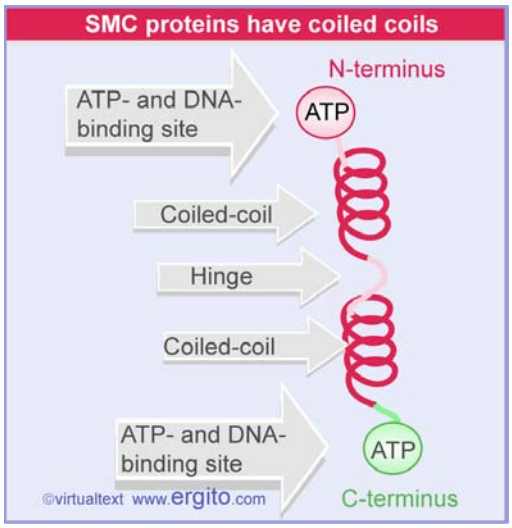


Figure 23.31 An SMP protein has a "Walker module" with an ATP-binding motif and DNA-binding site at each end, connected by coiled coils that are linked by a hinge region.

Experiments with the bacterial homologues of the SMC proteins suggest that a dimer is formed by an antiparallel interaction between the coiled coils, so that the N-terminus of one subunit bonds to the C-terminus of the other subunit. The existence of a flexible hinge region could allow cohesins and condensins to depend on a different mode of action by the dimer. **Figure 23.33** shows that cohesins have a V-shaped structure, with the arms separated by an 86° angle, whereas condensins are more sharply bent back, with only 6° between the arms. This enables cohesins to hold sister chromatids together, while condensins instead condense an individual chromosome (for review see 2323; 2262; 2375). **Figure 23.30** shows that a cohesin could take the form of an extended dimer that cross-links two DNA molecules. **Figure 23.32** shows that a condensin could take the form of a V-shaped dimer – essentially bent at the hinge – that pulls together distant sites on the same DNA molecule, causing it to condense.

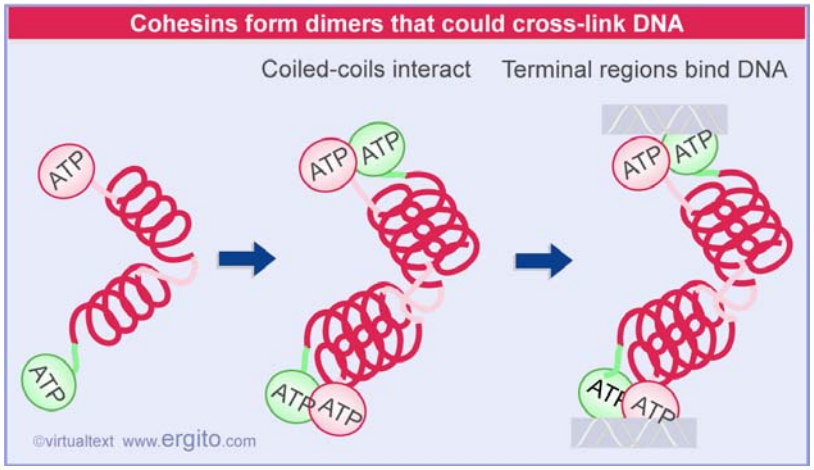


Figure 23.30 SMC proteins dimerize by anti-parallel interactions between the central coiled coils. Both terminal regions of each subunit have ATP- and DNA-binding motifs. Cohesins may form an extended structure that allows two different DNA molecules to be linked.

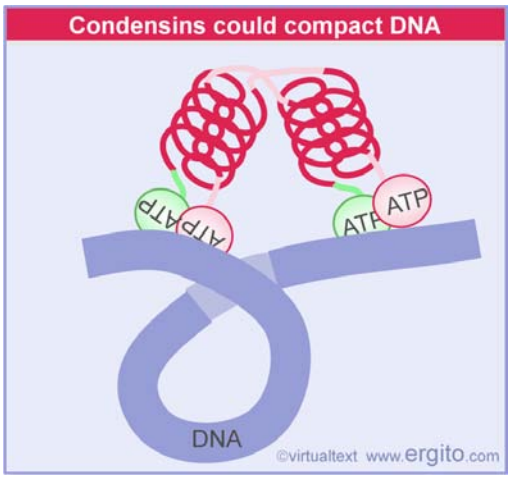


Figure 23.32 Condensins may form a compact structure by bending at the hinge, causing DNA to become compacted.

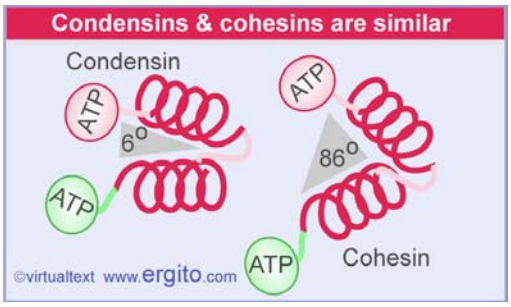


Figure 23.33 The two halves of a condensin are folded back at an angle of 6°. Cohesins have a more open conformation with an angle of 86° between the two halves.

An alternative model is suggested by experiments to suggest that the yeast proteins dimerize by intramolecular interactions, that is, a homodimer is formed solely by interaction between two identical subunits (2860). Dimers of two different proteins (in this case, SMC1 and SMC3) may then interact at both their head and hinge regions to form a circular structure as illustrated in **Figure 23.35**. Instead of binding directly to DNA, a structure of this type could hold DNA molecules together by encircling them (for review see 2859).

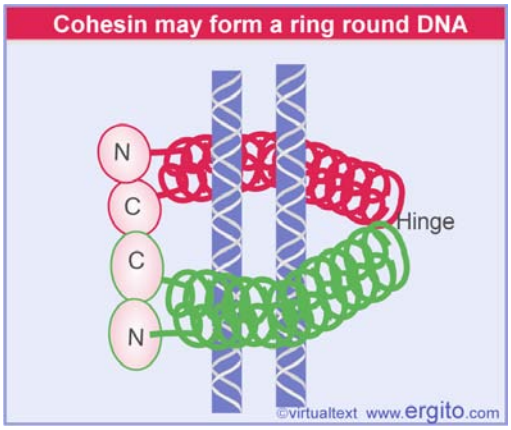


Figure 23.35 Cohesins may dimerize by intramolecular connections, then forming multimers that are connected at the heads and at the hinge. Such a structure could hold two molecules of DNA together by surrounding them.

Visualization of mitotic chromosomes shows that condensins are located all along the length of the chromosome, as can be seen in **Figure 23.34**. (By contrast, cohesins are found at discrete locations; see **Figure 29.34**).

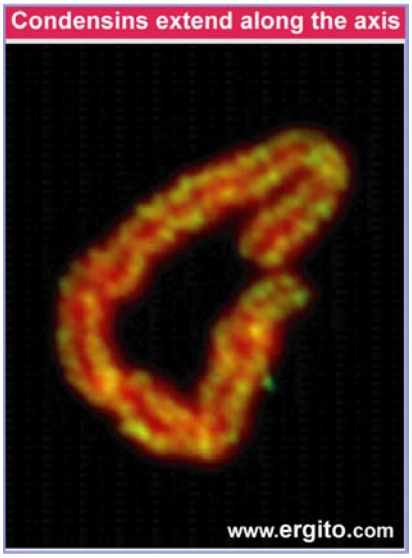


Figure 23.34 Condensins are located along the entire length of a mitotic chromosome. DNA is red; condensins are yellow. Photograph kindly provided by Ana Losada and Tatsuya Hirano.

The condensin complex was named for its ability to cause chromatin to condense *in vitro* (2463). It has an ability to introduce positive supercoils into DNA in an action that uses hydrolysis of ATP and depends on the presence of topoisomerase I. This ability is controlled by the phosphorylation of the non-SMC subunits, which occurs at mitosis. We do not know yet how this connects with other modifications of chromatin, for example, the phosphorylation of histones. The activation of the condensin complex specifically at mitosis makes it questionable whether it is also involved in the formation of interphase heterochromatin.

Global changes occur in other types of dosage compensation. In *Drosophila*, a complex of proteins is found in males, where it localizes on the X chromosome. In *C. elegans*, a protein complex associates with both X chromosomes in XX embryos, but the protein components remain diffusely distributed in the nuclei of XO embryos. The protein complex contains an SMC core, and is similar to the condensin complexes that are associated with mitotic chromosomes in other species. This suggests that it has a structural role in causing the chromosome to take up a more condensed, inactive state. Multiple sites on the X chromosome may be needed for the complex to be fully distributed along it. The complex binds to these sites, and then spreads along the vhromosome to cover it more thoroughly (4844).

Changes affecting all the genes on a chromosome, either negatively (mammals and *C. elegans*) or positively (*Drosophila*) are therefore a common feature of dosage compensation. However, the components of the dosage compensation apparatus may vary as well as the means by which it is localized to the chromosome, and of course its mechanism of action is different in each case.

Last updated on March 10, 2004

Reviews

2262. Hirano, T. (2000). *Chromosome cohesion, condensation, and separation.* Annu. Rev. Biochem. 69, 115-144.

2323. Hirano, T. (1999). *SMC-mediated chromosome mechanics: a conserved scheme from bacteria to vertebrates?* Genes Dev. 13, 11-19.

2375. Hirano, T. (2002). *The ABCs of SMC proteins: two-armed ATPases for chromosome condensation, cohesion, and repair.* Genes Dev. 16, 399-414.

2859. Nasmyth, K. (2002). *Segregating sister genomes: the molecular biology of chromosome separation.* Science 297, 559-565.

3445. Jessberger, R. (2002). *The many functions of SMC proteins in chromosome dynamics.* Nat. Rev. Mol. Cell Biol. 3, 767-778.

References

2463. Kimura, K., Rybenkov, V. V., Crisona, N. J., Hirano, T., and Cozzarelli, N. R. (1999). *13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation.* Cell 98, 239-248.

2860. Haering, C. H., Lowe, J.,Hochwage, A., and Nasmyth, K. (2002). *Molecular architecture of SMC proteins and the yeast cohesin complex.* Mol. Cell 9, 773-788.

4844. Csankovszki, G., McDonel, P., and Meyer, B. J. (2004). *Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes.* Science 303, 1182-1185.

CONTROLLING CHROMATIN STRUCTURE

5.23.19 DNA methylation is perpetuated by a maintenance methylase

Key Terms

A **fully methylated** site is a palindromic sequence that is methylated on both strands of DNA.

A **hemi-methylated** site is a palindromic sequence that is methylated on only one strand of DNA.

A **demethylase** is a casual name for an enzyme that removes a methyl group, typically from DNA, RNA, or protein.

A **methyltransferase (Methylase)** is an enzyme that adds a methyl group to a substrate, which can be a small molecule, a protein, or a nucleic acid.

A **de novo methylase** adds a methyl group to an unmethylated target sequence on DNA.

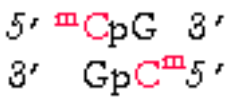
A **maintenance methylase** adds a methyl group to a target site that is already hemimethylated.

Key Concepts

- Most methyl groups in DNA are found on cytosine on both strands of the CpG doublet.
- Replication converts a fully methylated site to a hemi-methylated site.
- Hemi-methylated sites are converted to fully methylated sites by a maintenance methylase.

Methylation of DNA occurs at specific sites. In bacteria, it is associated with identifying the particular bacterial strain, and also with distinguishing replicated and nonreplicated DNA (see *Molecular Biology 4.15.24 Controlling the direction of mismatch repair*). In eukaryotes, its principal known function is connected with the control of transcription; methylation is associated with gene inactivation (see *Molecular Biology 5.21.18 Gene expression is associated with demethylation*).

From 2-7% of the cytosines of animal cell DNA are methylated (the value varies with the species). Most of the methyl groups are found in CG "doublets," and, in fact, the majority of the CG sequences are methylated. Usually the C residues on both strands of this short palindromic sequence are methylated, giving the structure



Such a site is described as **fully methylated**. But consider the consequences of

replicating this site. **Figure 23.36** shows that each daughter duplex has one methylated strand and one unmethylated strand. Such a site is called **hemi-methylated** (for review see 217).

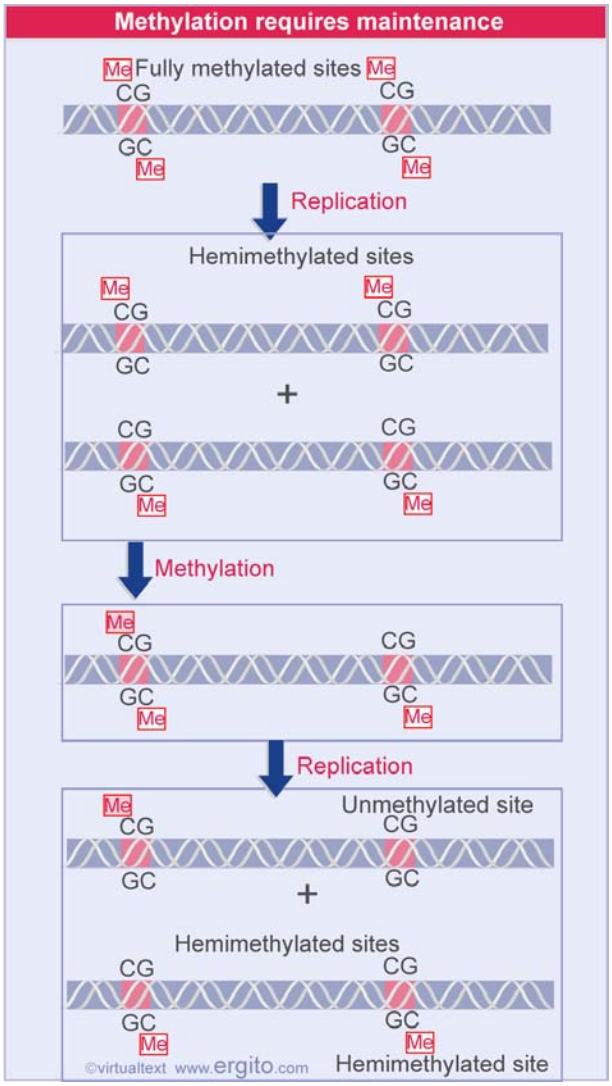


Figure 23.36 The state of methylated sites could be perpetuated by an enzyme that recognizes only hemimethylated sites as substrates.

The perpetuation of the methylated site now depends on what happens to hemimethylated DNA. If methylation of the unmethylated strand occurs, the site is restored to the fully methylated condition. However, if replication occurs first, the hemimethylated condition will be perpetuated on one daughter duplex, but the site will become unmethylated on the other daughter duplex. **Figure 23.37** shows that the state of methylation of DNA is controlled by **methylases**, which add methyl groups to the 5 position of cytosine, and **demethylases**, which remove the methyl groups. (The more formal name for the enzymes uses **methyltransferase** as the description.)

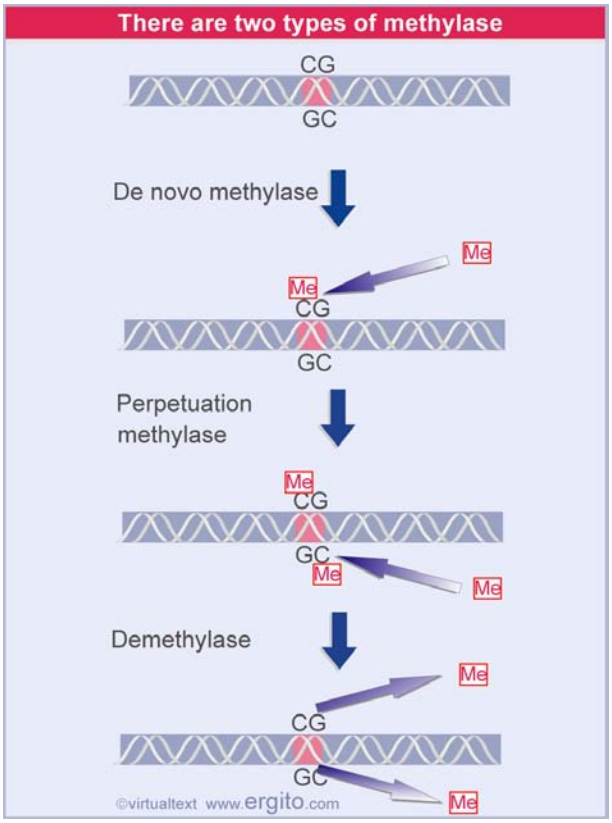


Figure 23.37 The state of methylation is controlled by three types of enzyme. *De novo* and perpetuation methylases are known, but demethylases have not been identified.

There are two types of DNA methylase, whose actions are distinguished by the state of the methylated DNA. To modify DNA at a new position requires the action of the **de novo methylase**, which recognizes DNA by virtue of a specific sequence. It acts *only* on nonmethylated DNA, to add a methyl group to one strand. There are two *de novo* methylases (Dnmt3A and Dnmt3B) in mouse; they have different target sites, and both are essential for development (941).

A **maintenance methylase** acts constitutively *only on hemimethylated sites* to convert them to fully methylated sites. Its existence means that any methylated site is perpetuated after replication. There is one maintenance methylase (Dnmt1) in mouse, and it is essential: mouse embryos in which its gene has been disrupted do not survive past early embryogenesis (942).

Maintenance methylation is virtually 100% efficient, ensuring that the situation shown on the left of **Figure 23.36** usually prevails *in vivo*. The result is that, if a *de novo* methylation occurs on one allele but not on the other, this difference will be perpetuated through ensuing cell divisions, maintaining a difference between the alleles that does not depend on their sequences.

Methylation has various types of targets. Gene promoters are the most common target. The promoters are methylated when the gene is inactive, but unmethylated when it is active. The absence of Dnmt1 in mouse causes widespread demethylation

at promoters, and we assume this is lethal because of the uncontrolled gene expression. Satellite DNA is another target. Mutations in Dnmt3B prevent methylation of satellite DNA, which causes centromere instability at the cellular level. Mutations in the corresponding human gene cause a disease (1964). The importance of methylation is emphasized by another human disease, which is caused by mutation of the gene for the protein McCp2 that binds methylated CpG sequences (1965).

The methylases are conventional enzymes that act on a DNA target. However, there may also be a methylation system that uses a short RNA sequence to target a corresponding DNA sequence for methylation (see *Molecular Biology 3.11.18 Antisense RNA can be used to inactivate gene expression*) Nothing is known about the mechanism of operation of this system (for review see 2077; 2078).

How are demethylated regions established and maintained? If a DNA site has not been methylated, a protein that recognizes the unmethylated sequence could protect it against methylation. Once a site has been methylated, there are two possible ways to generate demethylated sites. One is to block the maintenance methylase from acting on the site when it is replicated. After a second replication cycle, one of the daughter duplexes will be unmethylated (as shown on the right side of **Figure 23.36**). The other is actively to demethylate the site, as shown in **Figure 23.38**, either by removing the methyl group directly from cytosine, or by excising the methylated cytosine or cytidine from DNA for replacement by a repair system. We know that active demethylation can occur to the paternal genome soon after fertilization, but we do not know what mechanism is used (for review see 2424).

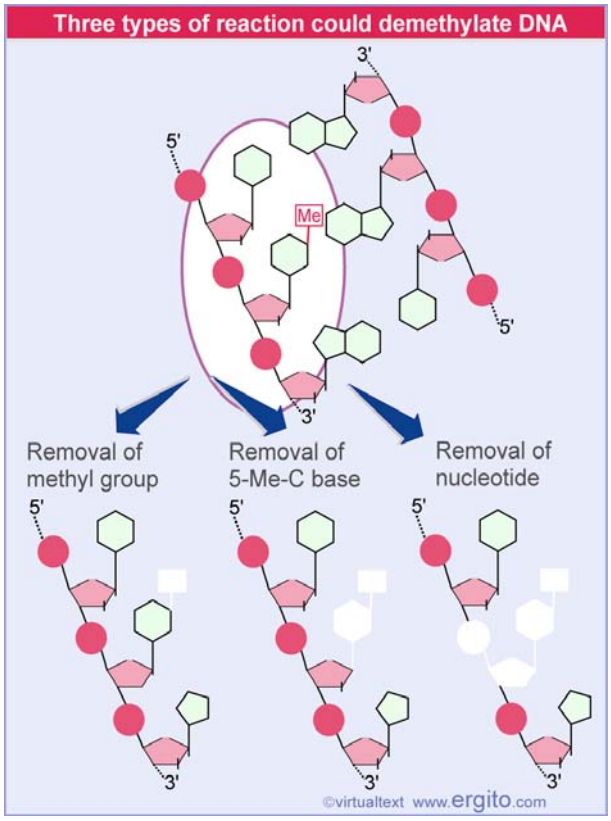


Figure 23.38 DNA could be demethylated by removing the methyl group, the base, or the nucleotide. Removal of the base or nucleotide would require its replacement by a repair system.

Last updated on 4-24-2002

Reviews

217. Bird, A. P. (1986). *A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA.* Nature 321, 209-213.

2077. Sharp, P. A. (2001). *RNA interference--2001.* Genes Dev. 15, 485-490.

2078. Matzke, M., Matzke, A. J., and Kooter, J. M. (2001). *RNA: guiding gene silencing.* Science 293, 1080-1083.

2424. Bird, A. (2002). *DNA methylation patterns and epigenetic memory.* Genes Dev. 16, 6-21.

References

941. Okano, M., Bell, D. W., Haber, D. A., and Li. E. (1999). *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell 99, 247-257.

942. Li, E., Bestor, T. H., and Jaenisch, R. (1992). *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell 69, 915-926.

1964. Xu, G. L., Bestor, T. H., Bourc'his, D., Hsieh, C. L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J., and Viegas-Paquignot, E. (1999). *Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene.* Nature 402, 187-191.

1965. Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999). *Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2.* Nat. Genet. 23, 185-188.

CONTROLLING CHROMATIN STRUCTURE

5.23.20 DNA methylation is responsible for imprinting

Key Terms

Imprinting describes a change in a gene that occurs during passage through the sperm or egg with the result that the paternal and maternal alleles have different properties in the very early embryo. May be caused by methylation of DNA.

Key Concepts

- Paternal and maternal alleles may have different patterns of methylation at fertilization.
- Methylation is usually associated with inactivation of the gene.
- When genes are differentially imprinted, survival of the embryo may require that the functional allele is provided by the parent with the unmethylated allele.
- Survival of heterozygotes for imprinted genes is different depending on the direction of the cross.
- Imprinted genes occur in clusters and may depend on a local control site where *de novo* methylation occurs unless specifically prevented.

The pattern of methylation of germ cells is established in each sex during gametogenesis by a two stage process: first the existing pattern is erased by a genome-wide demethylation; then the pattern specific for each sex is imposed.

All allelic differences are lost when primordial germ cells develop in the embryo; irrespective of sex, the previous patterns of methylation are erased, and a typical gene is then unmethylated. In males, the pattern develops in two stages. The methylation pattern that is characteristic of mature sperm is established in the spermatocyte. But further changes are made in this pattern after fertilization. In females, the maternal pattern is imposed during oogenesis, when oocytes mature through meiosis after birth.

As may be expected from the inactivity of genes in gametes, the typical state is to be methylated. However, there are cases of differences between the two sexes, where a locus is unmethylated in one sex. A major question is how the specificity of methylation is determined in the male and female gametes.

Systematic changes occur in early embryogenesis. Some sites will continue to be methylated, but others will be specifically unmethylated in cells in which a gene is expressed. From the pattern of changes, we may infer that individual sequence-specific demethylation events occur during somatic development of the organism as particular genes are activated (629).

The specific pattern of methyl groups in germ cells is responsible for the phenomenon of **imprinting**, which describes a difference in behavior between the alleles inherited from each parent. The expression of certain genes in mouse embryos depends upon the sex of the parent from which they were inherited. For example, the allele coding for IGF-II (insulin-like growth factor II) that is inherited from the father is expressed, but the allele that is inherited from the mother is not expressed. The IGF-II gene of oocytes is methylated, but the IGF-II gene of sperm is not methylated, so that the two alleles behave differently in the zygote. This is the most common pattern, but the dependence on sex is reversed for some genes. In fact, the opposite pattern (expression of maternal copy) is shown for IGF-IIR, the receptor for IGF-II (for review see 215).

This sex-specific mode of inheritance requires that the pattern of methylation is established specifically during each gametogenesis. The fate of a hypothetical locus in a mouse is illustrated in **Figure 23.39**. In the early embryo, the paternal allele is nonmethylated and expressed, and the maternal allele is methylated and silent. What happens when this mouse itself forms gametes? If it is a male, the allele contributed to the sperm must be nonmethylated, irrespective of whether it was originally methylated or not. So when the maternal allele finds itself in a sperm, it must be demethylated. If the mouse is a female, the allele contributed to the egg must be methylated; so if it was originally the paternal allele, methyl groups must be added.

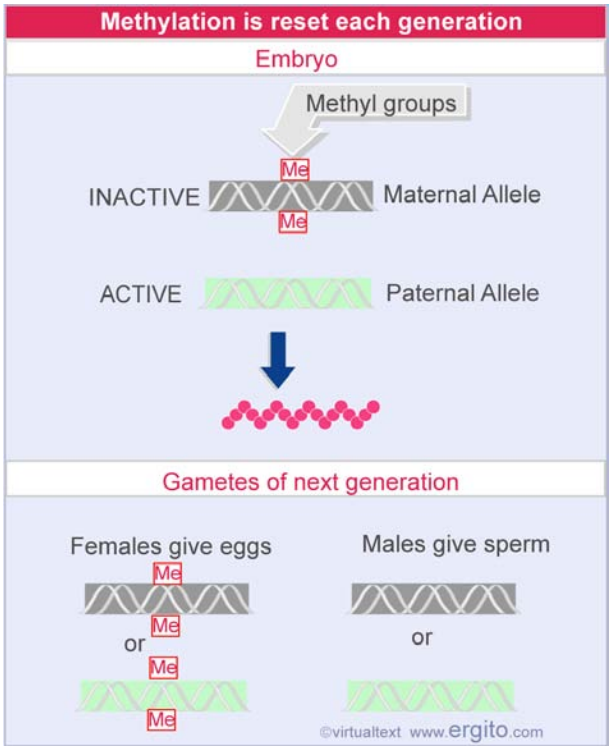


Figure 23.39 The typical pattern for imprinting is that a methylated locus is inactive. If this is the maternal allele, only the paternal allele is active, and will be essential for viability. The methylation pattern is reset when gametes are formed, so that all sperm have the paternal type, and all oocytes have the maternal type.

The consequence of imprinting is that an embryo requires a paternal allele for this gene. So in the case of a heterozygous cross where the allele of one parent has an inactivating mutation, the embryo will survive if the wild-type allele comes from the father, but will die if the wild-type allele is from the mother. This type of dependence on the directionality of the cross (in contrast with Mendelian genetics) is an example of epigenetic inheritance, where some factor other than the sequences of the genes themselves influences their effects (see *Molecular Biology 5.23.22 Epigenetic effects can be inherited*). Although the paternal and maternal alleles have identical sequences, they display different properties, depending on which parent provided them. These properties are inherited through meiosis and the subsequent somatic mitoses.

Imprinted genes are sometimes clustered. More than half of the 17 known imprinted genes in mouse are contained in two particular regions, each containing both maternally and paternally expressed genes. This suggests the possibility that imprinting mechanisms may function over long distances. Some insights into this possibility come from deletions in the human population that cause the Prader-Willi and Angelman diseases. Most cases are caused by the same 4 Mb deletion, but the syndromes are different, depending on which parent contributed the deletion. The reason is that the deleted region contains at least one gene that is paternally imprinted and at least one that is maternally imprinted. There are some rare cases, however, with much smaller deletions. Prader-Willi syndrome can be caused by a 20 kb deletion that silences genes that are distant on either side of it. The basic effect of the deletion is to prevent a father from resetting the paternal mode to a chromosome inherited from his mother. The result is that these genes remain in maternal mode, so that the paternal as well as maternal alleles are silent in the offspring. The inverse effect is found in some small deletions that cause Angelman's syndrome. The implication is that this region comprises some sort of "imprint center" that acts at a distance to switch one parental type to the other.

Last updated on 4-24-2002

Reviews

215. Bartolomei, M. S. and Tilghman, S. (1997). *Genomic imprinting in mammals.* Annu. Rev. Genet. 31, 493-525.

References

629. Chaillet, J. R., Vogt, T. F., Beier, D. R., and Leder, P. (1991). *Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis.* Cell 66, 77-83.

CONTROLLING CHROMATIN STRUCTURE

5.23.21 Oppositely imprinted genes can be controlled by a single center

Key Concepts

- Imprinted genes are controlled by methylation of *cis*-acting sites.
- Methylation may be responsible for either inactivating or activating a gene.

Imprinting is determined by the state of methylation of a *cis*-acting site near a target gene or genes. These regulatory sites are known as DMDs (differentially methylated domains) or ICRs (imprinting control regions). Deletion of these sites removes imprinting, and the target loci then behave the same in both maternal and paternal genomes.

The behavior of a region containing two genes, *Igf2* and *H19*, illustrates the ways in which methylation can control gene activity. **Figure 23.40** shows that these two genes react oppositely to the state of methylation at a site located between them, called the ICR. The ICR is methylated on the paternal allele. *H19* shows the typical response of inactivation. However, *Igf2* is expressed. The reverse situation is found on a maternal allele, where the ICR is not methylated. *H19* now becomes expressed, but *Igf2* is inactivated.

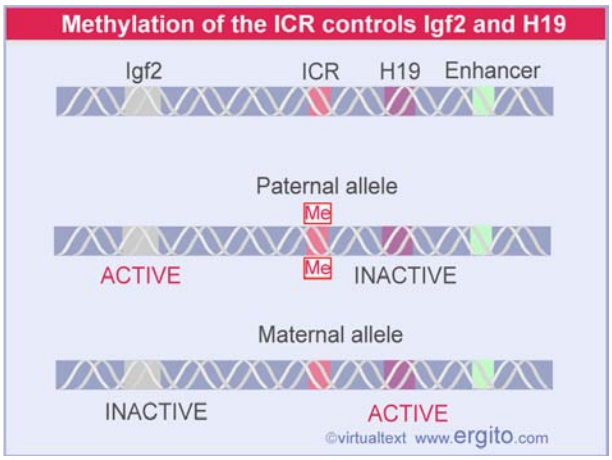


Figure 23.40 ICR is methylated on the paternal allele, where *Igf2* is active and *H19* is inactive. ICR is unmethylated on the maternal allele, where *Igf2* is inactive and *H19* is active.

The control of *Igf2* is exercised by an insulator function of the ICR. **Figure 23.41** shows that when the ICR is unmethylated, it binds the protein CTCF. This creates an insulator function that blocks an enhancer from activating the *Igf2* promoter (2048; 2052). This is an unusual effect in which methylation indirectly activates a gene by blocking an insulator.

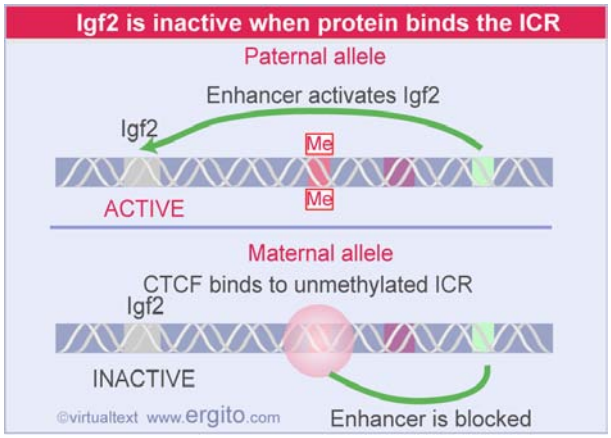


Figure 23.41 The ICR is an insulator that prevents an enhancer from activating Igf2. The insulator functions only when it binds CTCF to unmethylated DNA.

The regulation of *H19* shows the more usual direction of control in which methylation creates an inactive imprinted state. This could reflect a direct effect of methylation on promoter activity.

Last updated on 4-24-2002

References

2048. Bell, A. C. and Felsenfeld, G. (2000). *Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene.* Nature 405, 482-485.

2052. Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000). *CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.* Nature 405, 486-489.

CONTROLLING CHROMATIN STRUCTURE

5.23.22 Epigenetic effects can be inherited

Key Terms

Epigenetic changes influence the phenotype without altering the genotype. They consist of changes in the properties of a cell that are inherited but that do not represent a change in genetic information.

Key Concepts

- Epigenetic effects can result from modification of a nucleic acid after it has been synthesized or by the perpetuation of protein structures.

Epigenetic inheritance describes the ability of different states, which may have different phenotypic consequences, to be inherited without any change in the sequence of DNA. How can this occur?

We can divide epigenetic mechanisms into two general classes:

- DNA may be modified by the covalent attachment of a moiety that is then perpetuated. Two alleles with the same sequence may have different states of methylation that confer different properties.
- Or a self perpetuating protein state may be established. This might involve assembly of a protein complex, modification of specific protein(s), or establishment of an alternative protein conformation.

Methylation establishes epigenetic inheritance so long as the maintenance methylase acts constitutively to restore the methylated state after each cycle of replication, as shown in **Figure 23.36**. A state of methylation can be perpetuated through an indefinite series of somatic mitoses. This is probably the "default" situation. Methylation can also be perpetuated through meiosis: for example, in the fungus *Ascobolus* there are epigenetic effects that can be transmitted through both mitosis and meiosis by maintaining the state of methylation. In mammalian cells, epigenetic effects are created by resetting the state of methylation differently in male and female meioses.

Situations in which epigenetic effects appear to be maintained by means of protein states are less well understood in molecular terms. Position effect variegation shows that constitutive heterochromatin may extend for a variable distance, and the structure is then perpetuated through somatic divisions. Since there is no methylation of DNA in *Saccharomyces* and a vanishingly small amount in *Drosophila*, the inheritance of epigenetic states of position effect variegation or telomeric silencing in these organisms is likely to be due to the perpetuation of protein structures.

Figure 23.42 considers two extreme possibilities for the fate of a protein complex at

replication.

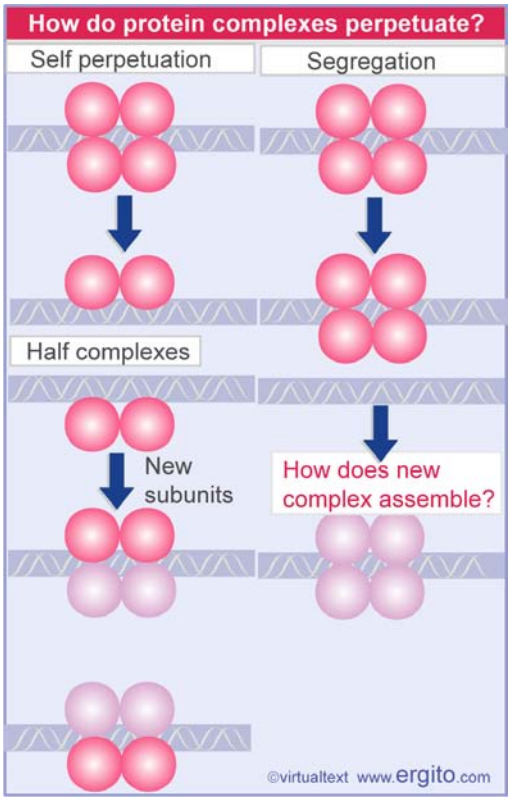


Figure 23.42 What happens to protein complexes on chromatin during replication?

- A complex could perpetuate itself if it splits symmetrically, so that half complexes associate with each daughter duplex. If the half complexes have the capacity to nucleate formation of full complexes, the original state will be restored. This is basically analogous to the maintenance of methylation. The problem with this model is that there is no evident reason why protein complexes should behave in this way.

- A complex could be maintained as a unit and segregate to one of the two daughter duplexes. The problem with this model is that it requires a new complex to be assembled *de novo* on the other daughter duplex, and it is not evident why this should happen.

Consider now the need to perpetuate a heterochromatic structure consisting of protein complexes. Suppose that a protein is distributed more or less continuously along a stretch of heterochromatin, as implied in **Figure 23.20**. If individual subunits are distributed at random to each daughter duplex at replication, the two daughters will continue to be marked by the protein, although its density will be reduced to half of the level before replication. If the protein has a self-assembling property that causes new subunits to associate with it, the original situation may be restored. *Basically, the existence of epigenetic effects forces us to the view that a protein responsible for such a situation must have some sort of self-templating or*

self-assembling capacity.

In some cases, it may be the state of protein modification, rather than the presence of the protein *per se*, that is responsible for an epigenetic effect. There is a general correlation between the activity of chromatin and the state of acetylation of the histones, in particular the acetylation of histones H3 and H4, which occurs on their N-terminal tails. Activation of transcription is associated with acetylation in the vicinity of the promoter; and repression of transcription is associated with deacetylation (see *Molecular Biology 5.23.7 Acetylases are associated with activators*). The most dramatic correlation is that the inactive X chromosome in mammalian female cells is underacetylated on histone H4.

The inactivity of constitutive heterochromatin may require that the histones are not acetylated. If a histone acetyltransferase is tethered to a region of telomeric heterochromatin in yeast, silenced genes become active. When yeast is exposed to trichostatin (an inhibitor of deacetylation), centromeric heterochromatin becomes acetylated, and silenced genes in centromeric regions may become active. *The effect may persist even after trichostatin has been removed.* In fact, it may be perpetuated through mitosis and meiosis. This suggests that an epigenetic effect has been created by changing the state of histone acetylation.

How might the state of acetylation be perpetuated? Suppose that the $H3_2 \cdot H4_2$ tetramer is distributed at random to the two daughter duplexes. This creates the situation shown in **Figure 23.43**, in which each daughter duplex contains some histone octamers that are fully acetylated on the H3 and H4 tails, while others are completely unacetylated. To account for the epigenetic effect, we could suppose that the presence of some fully acetylated histone octamers provides a signal that causes the unacetylated octamers to be acetylated.

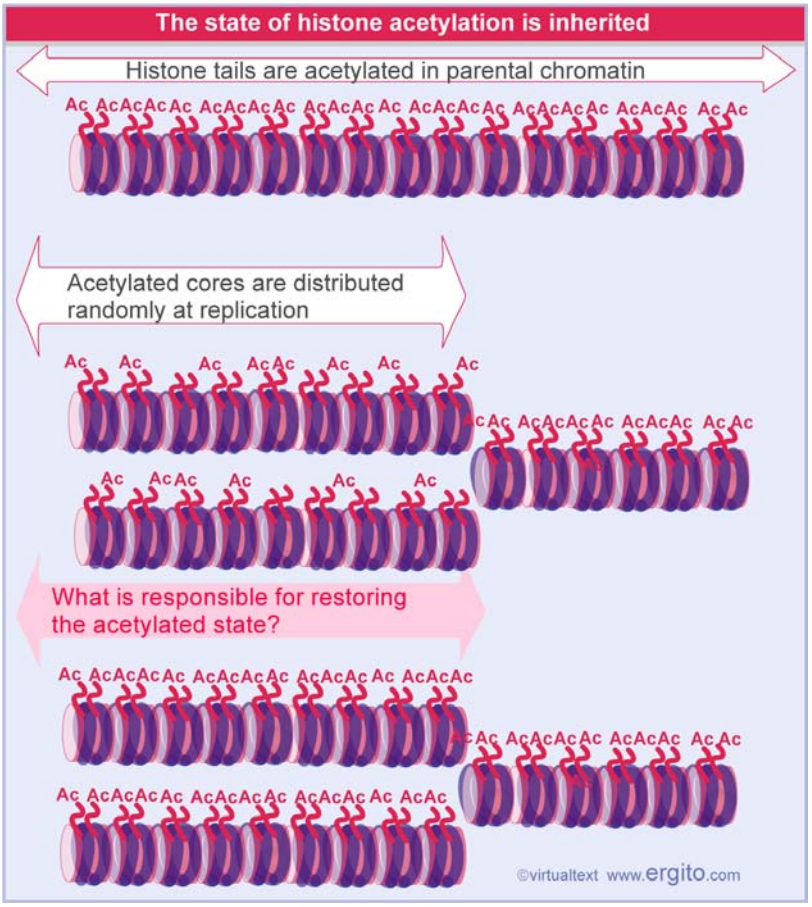


Figure 23.43 Acetylated cores are conserved and distributed at random to the daughter chromatin fibers at replication. Each daughter fiber has a mixture of old (acetylated) cores and new (unacetylated) cores.

(The actual situation is probably more complicated than shown in the figure, because transient acetylations occur during replication. If they are simply reversed following deposition of histones into nucleosomes, they may be irrelevant. An alternative possibility is that the usual deacetylation is prevented, instead of, or as well as, inducing acetylation.)

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.5.23.22

CONTROLLING CHROMATIN STRUCTURE

5.23.23 Yeast prions show unusual inheritance

Key Terms

A **prion** is a proteinaceous infectious agent, which behaves as an inheritable trait, although it contains no nucleic acid. Examples are PrP^{Sc}, the agent of scrapie in sheep and bovine spongiform encephalopathy, and Psi, which confers an inherited state in yeast.

Key Concepts

- The Sup35 protein in its wild-type soluble form is a termination factor for translation.
- It can also exist in an alternative form of oligomeric aggregates, in which it is not active in protein synthesis.
- The presence of the oligomeric form causes newly synthesized protein to acquire the inactive structure.
- Conversion between the two forms is influenced by chaperones.
- The wild-type form has the recessive genetic state *psi*$^-$ and the mutant form has the dominant genetic state *PSI*$^+$.

One of the clearest cases of the dependence of epigenetic inheritance on the condition of a protein is provided by the behavior of **prions** – proteinaceous infectious agents. They have been characterized in two circumstances: by genetic effects in yeast; and as the causative agents of neurological diseases in mammals, including man. A striking epigenetic effect is found in yeast, where two different states can be inherited that map to a single genetic locus, *although the sequence of the gene is the same in both states*. The two different states are [*psi*$^-$] and [*PSI*$^+$]. A switch in condition occurs at a low frequency as the result of a spontaneous transition between the states (for review see 211; 212).

The *psi* genotype maps to the locus *sup35*, which codes for a translation termination factor. **Figure 23.44** summarizes the effects of the Sup35 protein in yeast. In wild-type cells, which are characterized as [*psi*$^-$], the gene is active, and Sup35 protein terminates protein synthesis. In cells of the mutant [*PSI*$^+$] type, the factor does not function, causing a failure to terminate protein synthesis properly. (This was originally detected by the lethal effects of the enhanced efficiency of suppressors of ochre codons in [*PSI*$^+$] strains.)

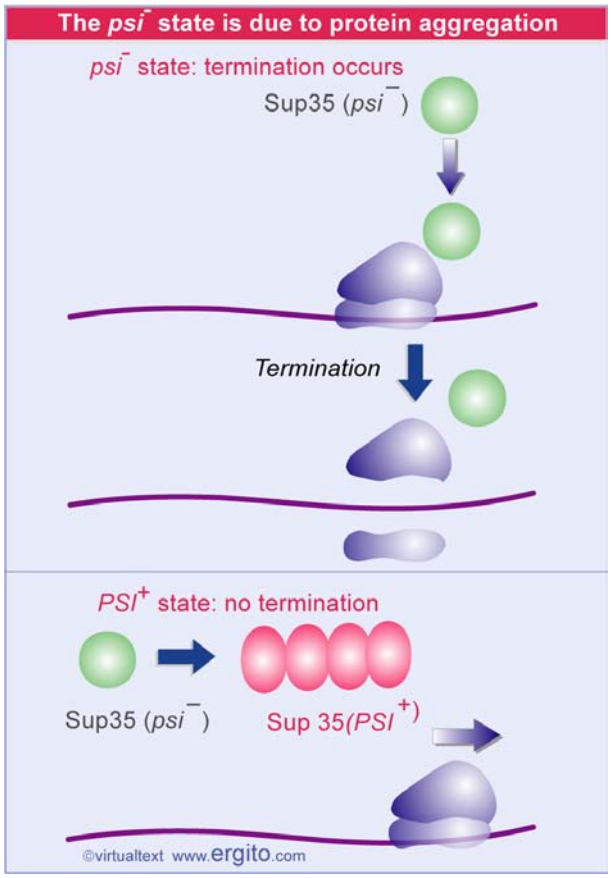


Figure 23.44 The state of the Sup35 protein determines whether termination of translation occurs.

[*PSI*$^+$] strains have unusual genetic properties. When a [*psi*$^-$] strain is crossed with a [*PSI*$^+$] strain, *all of the progeny are [PSI*$^+$*]*. This is a pattern of inheritance that would be expected of an extrachromosomal agent, but the [*PSI*$^+$] trait cannot be mapped to any such nucleic acid. The [*PSI*$^+$] trait is metastable, which means that, although it is inherited by most progeny, it is lost at a higher rate than is consistent with mutation. Similar behavior is shown also by the locus *URE2*, which codes for a protein required for nitrogen-mediated repression of certain catabolic enzymes. When a yeast strain is converted into an alternative state, called [*URE3*], the Ure2 protein is no longer functional (637).

The [*PSI*$^+$] state is determined by the conformation of the Sup35 protein. In a wild-type [*psi*$^-$] cell, the protein displays its normal function. But in a [*PSI*$^+$] cell, the protein is present in an alternative conformation in which its normal function has been lost. To explain the unilateral dominance of [*PSI*$^+$] over [*psi*$^-$] in genetic crosses, we must suppose that *the presence of protein in the [PSI*$^+$*] state causes all the protein in the cell to enter this state.* This requires an interaction between the [*PSI*$^+$] protein and newly synthesized protein, probably reflecting the generation of an oligomeric state in which the [*PSI*$^+$] protein has a nucleating role, as illustrated in **Figure 23.45** (for review see 997).

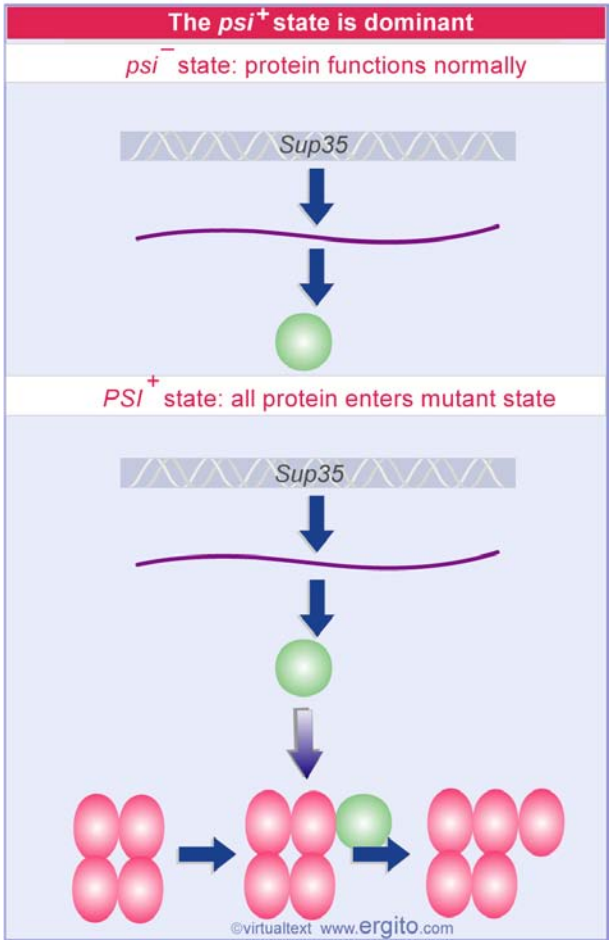


Figure 23.45 Newly synthesized Sup35 protein is converted into the [PSI$^+$] state by the presence of pre-existing [PSI$^+$] protein.

A feature common to both the Sup35 and Ure2 proteins is that each consists of two domains that function independently. The C-terminal domain is sufficient for the activity of the protein. The N-terminal domain is sufficient for formation of the structures that make the protein inactive. So yeast in which the N-terminal domain of Sup35 has been deleted cannot acquire the [*PSI*$^+$] state; and the presence of an [*PSI*$^+$] N-terminal domain is sufficient to maintain Sup35 protein in the [*PSI*$^+$] condition (639). The critical feature of the N-terminal domain is that it is rich in glutamine and asparagine residues.

Loss of function in the [*PSI*$^+$] state is due to the sequestration of the protein in an oligomeric complex. Sup35 protein in [*PSI*$^+$] cells is clustered in discrete foci, whereas the protein in [*psi*$^-$] cells is diffused in the cytosol. Sup35 protein from [*PSI*$^+$] cells forms amyloid fibers *in vitro* – these have a characteristic high content of β sheet structures (640).

The involvement of protein conformation (rather than covalent modification) is suggested by the effects of conditions that affect protein structure. Denaturing treatments cause loss of the [*PSI*$^+$] state. And in particular, the chaperone Hsp104 is

involved in inheritance of [*PSI*$^+$]. Its effects are paradoxical. Deletion of *HSP104* prevents maintenance of the [*PSI*$^+$] state. And overexpression of Hsp104 also causes loss of the [*PSI*$^+$] state. This suggests that Hsp104 is required for some change in the structure of Sup35 that is necessary for acquisition of the [*PSI*$^+$] state, but that must be transitory (638; for review see 213).

Using the ability of Sup35 to form the inactive structure *in vitro*, it is possible to provide biochemical proof for the role of the protein. **Figure 23.46** illustrates a striking experiment in which the protein was converted to the inactive form *in vitro*, put into liposomes (when in effect the protein is surrounded by an artificial membrane), and then introduced directly into cells by fusing the liposomes with [*psi*$^-$] yeast (1069). The yeast cells were converted to [*PSI*$^+$]! This experiment refutes all of the objections that were raised to the conclusion that the protein has the ability to confer the epigenetic state. Experiments in which cells are mated, or in which extracts are taken from one cell to treat another cell, always are susceptible to the possibility that a nucleic acid has been transferred. But when the protein by itself does not convert target cells, but protein converted to the inactive state can do so, the only difference is the treatment of the protein – which must therefore be responsible for the conversion.

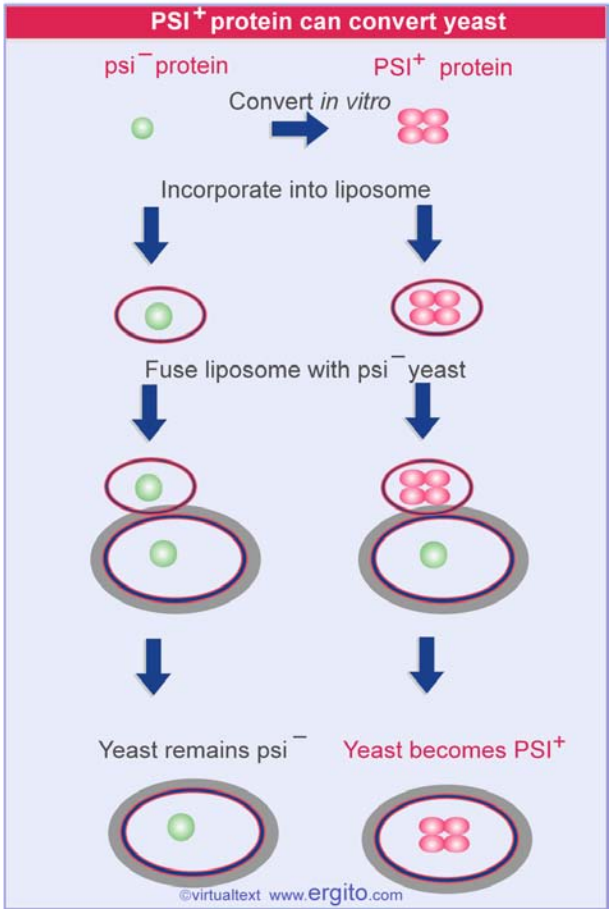


Figure 23.46 Purified protein can convert the[[*psi*$^-$] state of yeast to [*PSI*$^+$].

The ability of yeast to form the [*PSI*$^+$] prion state depends on the genetic background. The yeast must be [*PIN*$^+$] in order for the [*PSI*$^+$] state to form. The [*PIN*$^+$] condition itself is an epigenetic state (1953). It can be created by the formation of prions from any one of several different proteins (1954). These proteins share the characteristic of Sup35 that they have Gln/Asn-rich domains. Overexpression of these domains in yeast stimulates formation of the [*PSI*$^+$] state (1955). This suggests that there is a common model for the formation of the prion state that involves aggregation of the Gln/Asn domains.

How does the presence of one Gln/Asn protein influence the formation of prions by another? We know that the formation of Sup35 prions is specific to Sup35 protein, that is, it does not occur by cross-aggregation with other proteins. This suggests that the yeast cell may contain soluble proteins that antagonize prion formation. These proteins are not specific for any one prion. As a result, the introduction of any Gln/Asn domain protein that interacts with these proteins will reduce the concentration. This will allow other Gln/Asn proteins to aggregate more easily.

Last updated on 8-17-2001

Reviews

211. Wickner, R. B. (1996). *Prions and RNA viruses of S. cerevisiae.* Annu. Rev. Genet. 30, 109-139.

212. Lindquist, S. (1997). *Mad cows meet psi-chotic yeast: the expansion of the prion hypothesis.* Cell 89, 495-498.

213. Horwich, A. L. and Weissman, J. S. (1997). *Deadly conformations: protein misfolding in prion disease.* Cell 89, 499-510.

997. Serio, T. R. and Lindquist, S. L. (1999). *[PSI$^+$]: an epigenetic modulator of translation termination efficiency.* Annu. Rev. Cell Dev. Biol. 15, 661-703.

References

637. Wickner, R. B. (1994). *[URE3] as an altered URE2 protein: evidence for a prion analog in S. cerevisiae.* Science 264, 566-569.

638. Chernoff, Y. O. et al. (1995). *Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [PSI$^+$].* Science 268, 880-884.

639. Masison, D. C. and Wickner, R. B. (1995). *Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells.* Science 270, 93-95.

640. Glover, J. R. et al. (1997). *Self-seeded fibers formed by Sup35, the protein determinant of [PSI$^+$], a heritable prion-like factor of S. cerevisiae.* Cell 89, 811-819.

1069. Sparrer, H E, Santoso, A, Szoka, F C, and Weissman, J S (2000). *Evidence for the prion hypothesis: induction of the yeast [PSI+] factor by in vitro-converted Sup35 protein.* Science 289, 595-599.

1953. Derkatch, I. L., Bradley, M.E., Masse, S. V., Zadorsky, S.P., Polozkov, G. V., Inge-Vechtomov, S. G., Liebman S. W. (2000). *Dependence and independence of [PSI(+)] and [PIN(+)]: a two-prion system in yeast?* EMBO J. 19, 1942-1952.

1954. Derkatch, I. L., Bradley, M. E., Hong, J. Y., and Liebman, S. W. (2001). *Prions affect the appearance of other prions: the story of [PIN(+)].* Cell 106, 171-182.

1955. Osherovich, L. Z. and Weissman, J. S. (2001). *Multiple gln/asn-rich prion domains confer susceptibility to induction of the yeast.* Cell 106, 183-194.

CONTROLLING CHROMATIN STRUCTURE

5.23.24 Prions cause diseases in mammals

Key Terms

Scrapie is a infective agent made of protein.

Kuru is a human neurological disease caused by prions. It may be caused by eating infected brains.

Key Concepts

- The protein responsible for scrapie exists in two forms, the wild-type noninfectious form PrP^{C} which is susceptible to proteases, and the disease-causing form PrP^{Sc} which is resistant to proteases.
- The neurological disease can be transmitted to mice by injecting the purified PrP^{Sc} protein into mice.
- The recipient mouse must have a copy of the *PrP* gene coding for the mouse protein.
- The PrP^{Sc} protein can perpetuate itself by causing the newly synthesized PrP protein to take up the PrP^{Sc} form instead of the PrP^{C} form.
- Multiple strains of PrP^{Sc} may have different conformations of the protein.

Prion diseases have been found in sheep and Man, and, more recently, in cows. The basic phenotype is an ataxia – a neurodegenerative disorder that is manifested by an inability to remain upright. The name of the disease in sheep, **scrapie**, reflects the phenotype: the sheep rub against walls in order to stay upright. Scrapie can be perpetuated by inoculating sheep with tissue extracts from infected animals. The disease **kuru** was found in New Guinea, where it appeared to be perpetuated by cannibalism, in particular the eating of brains. Related diseases in Western populations with a pattern of genetic transmission include Gerstmann-Straussler syndrome; and the related Creutzfeldt-Jakob disease (CJD) occurs sporadically. Most recently, a disease resembling CJD appears to have been transmitted by consumption of meat from cows suffering from "mad cow" disease (634).

When tissue from scrapie-infected sheep is inoculated into mice, the disease occurs in a period ranging from 75-150 days. The active component is a protease-resistant protein. The protein is coded by a gene that is normally expressed in brain. The form of the protein in normal brain, called PrP^{C}, is sensitive to proteases. Its conversion to the resistant form, called Prp^{Sc}, is associated with occurrence of the disease. The infectious preparation has no detectable nucleic acid, is sensitive to UV irradiation at wave lengths that damage protein, and has a low infectivity (1 infectious unit / 10^{5} PrP^{Sc} proteins). This corresponds to an epigenetic inheritance in which there is no change in genetic information, because normal and diseased cells have the same *PrP* gene sequence, but the PrP^{Sc} form of the protein is the infectious agent, whereas PrP^{C} is harmless (383; 384; 385; for review see 203).

The basis for the difference between the PrP^{Sc} and Prp^{C} forms appears to lie with a change in conformation rather than with any covalent alteration. Both proteins are glycosylated and linked to the membrane by a GPI-linkage. No changes in these modifications have been found. The PrP^{Sc} form has a high content of β sheets, which is absent from the PrP^{C} form.

The assay for infectivity in mice allows the dependence on protein sequence to be tested. **Figure 23.47** illustrates the results of some critical experiments. In the normal situation, PrP^{Sc} protein extracted from an infected mouse will induce disease (and ultimately kill) when it is injected into a recipient mouse. If the *PrP* gene is "knocked out", a mouse becomes resistant to infection. This experiment demonstrates two things. First, the endogenous protein is necessary for an infection, presumably because it provides the raw material that is converted into the infectious agent. Second, the cause of disease is not the removal of the PrP^{C} form of the protein, because a mouse with no PrP^{C} survives normally: the disease is caused by a gain-of-function in PrP^{Sc} (386).

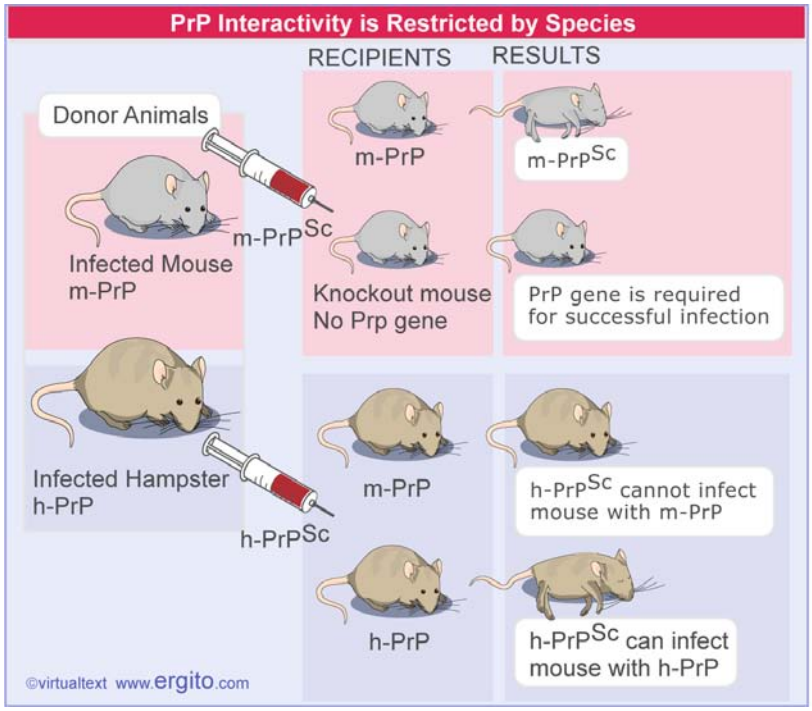


Figure 23.47 A PrpSc protein can only infect an animal that has the same type of endogenous PrPC protein.

The existence of species barriers allows hybrid proteins to be constructed to delineate the features required for infectivity. The original preparations of scrapie were perpetuated in several types of animal, but these cannot always be transferred readily. For example, mice are resistant to infection from prions of hamsters. This means that hamster-PrP^{Sc} cannot convert mouse-PrP^{C} to PrP^{Sc}. However, the situation changes if the mouse *PrP* gene is replaced by a hamster *PrP* gene. (This can be done by introducing the hamster *PrP* gene into the *PrP* knockout mouse.) A mouse with a hamster *PrP* gene is sensitive to infection by hamster PrP^{Sc}. This suggests that the conversion of cellular PrP^{C} protein into the Sc state requires that the PrP^{Sc} and PrP^{C} proteins have matched sequences.

There are different "strains" of PrP^{Sc}, which are distinguished by characteristic incubation periods upon inoculation into mice. This implies that the protein is not restricted solely to alternative states of PrP^{C} and PrP^{Sc}, but that there may be multiple Sc states. These differences must depend on some self-propagating property of the protein other than its sequence. If conformation is the feature that distinguishes PrP^{Sc} from PrP^{C}, then there must be multiple conformations, each of which has a self-templating property when it converts PrP^{C} (636; for review see 214).

The probability of conversion from PrP^{C} to PrP^{Sc} is affected by the sequence of PrP. Gerstmann-Straussler syndrome in man is caused by a single amino acid change in PrP. This is inherited as a dominant trait. If the same change is made in the mouse PrP gene, mice develop the disease. This suggests that the mutant protein has an increased probability of spontaneous conversion into the Sc state. Similarly, the sequence of the PrP gene determines the susceptibility of sheep to develop the disease spontaneously; the combination of amino acids at three positions (codons 136, 154, and 171) determines susceptibility.

The prion offers an extreme case of epigenetic inheritance, in which the infectious agent is a protein that can adopt multiple conformations, each of which has a self-templating property. This property is likely to involve the state of aggregation of the protein.

Reviews

203. Prusiner, S. (1982). *Novel proteinaceous infectious particles cause scrapie*. Science 216, 136-144.

214. Prusiner, S. B. and Scott, M. R. (1997). *Genetics of prions*. Annu. Rev. Genet. 31, 139-175.

References

383. McKinley, M. P., Bolton, D. C., and Prusiner, S. B. (1983). *A protease-resistant protein is a structural component of the scrapie prion*. Cell 35, 57-62.

384. Oesch, B. et al. (1985). *A cellular gene encodes scrapie PrP27-30 protein*. Cell 40, 735-746.

385. Basler, K., Oesch, B., Scott, M., Westaway, D., Walchli, M., Groth, D. F., McKinley, M. P., Prusiner, S. B., and Weissmann, C. (1986). *Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene*. Cell 46, 417-428.

386. Bueler, H. et al. (1993). *Mice devoid of PrP are resistant to scrapie*. Cell 73, 1339-1347.

634. Hsiao, K. et al. (1989). *Linkage of a prion protein missense variant to Gerstmann-Straussler syndrome*. Nature 338, 342-345.

636. Scott, M. et al. (1993). *Propagation of prions with artificial properties in transgenic mice expressing chimeric PrP genes.* Cell 73, 979-988.

CONTROLLING CHROMATIN STRUCTURE

5.23.25 Summary

The existence of a preinitiation complex signals that the gene is in an "active" state, ready to be transcribed. The complex is stable, and may remain in existence through many cycles of replication. The ability to form a preinitiation complex could be a general regulatory mechanism. By binding to a promoter to make it possible for RNA polymerase in turn to bind, the factor in effect switches the gene on.

The variety of situations in which hypersensitive sites occur suggests that their existence reflects a general principle. *Sites at which the double helix initiates an activity are kept free of nucleosomes.* A transcription factor, or some other nonhistone protein concerned with the particular function of the site, modifies the properties of a short region of DNA so that nucleosomes are excluded. The structures formed in each situation need not necessarily be similar (except that each, by definition, creates a site hypersensitive to DNAase I).

Genes whose control regions are organized in nucleosomes usually are not expressed. In the absence of specific regulatory proteins, promoters and other regulatory regions are organized by histone octamers into a state in which they cannot be activated. This may explain the need for nucleosomes to be precisely positioned in the vicinity of a promoter, so that essential regulatory sites are appropriately exposed. Some transcription factors have the capacity to recognize DNA on the nucleosomal surface, and a particular positioning of DNA may be required for initiation of transcription.

Active chromatin and inactive chromatin are not in equilibrium. Sudden, disruptive events are needed to convert one to the other. Chromatin remodeling complexes have the ability to displace histone octamers by a mechanism that involves hydrolysis of ATP. Remodeling complexes are large and are classified according to the type of the ATPase subunit. Two common types are SWI/SNF and ISW. A typical form of this chromatin remodeling is to displace one or more histone octamers from specific sequences of DNA, creating a boundary that results in the precise or preferential positioning of adjacent nucleosomes. Chromatin remodeling may also involve changes in the positions of nucleosomes, sometimes involving sliding of histone octamers along DNA.

Acetylation of histones occurs at both replication and transcription and could be necessary to form a less compact chromatin structure. Some coactivators, which connect transcription factors to the basal apparatus, have histone acetylase activity. Conversely, repressors may be associated with deacetylases. The modifying enzymes are usually specific for particular amino acids in particular histones. The most common sites for modification are located in the N-terminal tails of histones H3 and H4, which extrude from nucleosomes between the turns of DNA. The activating (or repressing) complexes are usually large and often contain several activities that undertake different modifications of chromatin. Some common motifs found in proteins that modify chromatin are the chromo domain (concerned with protein-protein interactions). the bromo domain (which targets acetylated lysine), and the SET domain (part of the active sites of histone methyltransferases).

The formation of heterochromatin occurs by proteins that bind to specific chromosomal regions (such as telomeres) and that interact with histones. The formation of an inactive structure may propagate along the chromatin thread from an initiation center. Similar events occur in silencing of the inactive yeast mating type loci. Repressive structures that are required to maintain the inactive states of particular genes are formed by the Pc-G protein complex in *Drosophila*. They share with heterochromatin the property of propagating from an initiation center.

Formation of heterochromatin may be initiated at certain sites and then propagated for a distance that is not precisely determined. When a heterochromatic state has been established, it is inherited through subsequent cell divisions. This gives rise to a pattern of epigenetic inheritance, in which two identical sequences of DNA may be associated with different protein structures, and therefore have different abilities to be expressed. This explains the occurrence of position effect variegation in *Drosophila*.

Modification of histone tails is a trigger for chromatin reorganization. Acetylation is generally associated with gene activation. Histones acetylases are found in activating complexes, and histone deacetylases are found in inactivating complexes. Histone methylation is associated with gene inactivation. Some histone modifications may be exclusive or synergistic with others.

Inactive chromatin at yeast telomeres and silent mating type loci appears to have a common cause, and involves the interaction of certain proteins with the N-terminal tails of histones H3 and H4. Formation of the inactive complex may be initiated by binding of one protein to a specific sequence of DNA; the other components may then polymerize in a cooperative manner along the chromosome.

Inactivation of one X chromosome in female (eutherian) mammals occurs at random. The *Xic* locus is necessary and sufficient to count the number of X chromosomes. The n-1 rule ensures that all but one X chromosome are inactivated. *Xic* contains the gene *Xist*, which codes for an RNA that is expressed only on the inactive X chromosome. Stabilization of *Xist* RNA is the mechanism by which the inactive X chromosome is distinguished.

Methylation of DNA is inherited epigenetically. Replication of DNA creates hemimethylated products, and a maintenance methylase restores the fully methylated state. Some methylation events depend on parental origin. Sperm and eggs contain specific and different patterns of methylation, with the result that paternal and maternal alleles are differently expressed in the embryo. This is responsible for imprinting, in which the nonmethylated allele inherited from one parent is essential because it is the only active allele; the allele inherited from the other parent is silent. Patterns of methylation are reset during gamete formation in every generation.

Prions are proteinaceous infectious agents that are responsible for the disease of scrapie in sheep and for related diseases in man. The infectious agent is a variant of a normal cellular protein. The PrP^{Sc} form has an altered conformation that is self-templating: the normal PrP^{C} form does not usually take up this conformation, but does so in the presence of PrP^{Sc}. A similar effect is responsible for inheritance of the *PSI* element in yeast.