Chromatin modification and disease

Colin A Johnson

"Physicians consider that when they have discovered the cause of disease, they have also discovered the method of treating it." Cicero, Tusculan Disputations, III.x.23.

In the last few years, the exciting realisation in the field of gene regulation is that transcription factors can function by recruiting large, multiprotein complexes which mediate several types of chromatin modification and remodelling events that alter the structure of chromatin. Chromatin structure changes include post-translational modifications of histones, DNA methylation, remodelling of the chromatin, and the maintenance of a heterochromatic or euchromatic state. Most of these events are brought about by enzymatic mechanisms. In general, the catalytic subunits are only one component of the complexes, with the distribution and localisation of the structural changes dependent on targeting components. Many of the catalytic components (sometimes called coactivators and corepressors) interact with the activator and repressor proteins that mediate the actual process of transcriptional regulation. Transcriptional dysregulation can therefore arise from mutations that cause the loss or perturbation of chromatin modification or remodelling, which are now known to have an important role in the pathogenesis of cancer and other genetic diseases. Some of the proteins that mediate these events are therefore novel molecular targets for future treatments.

In eukaryotes, DNA is packaged by histone proteins into nucleosomes, the fundamental repeating structural unit of chromatin. The nucleosomal core particle consists of an octomeric complex of core histones (two each of H2A, H2B, H3, and H4) around which 147 bp of DNA is wrapped in 1.65 turns of a left handed superhelix. The minor and major grooves of adjacent turns of the DNA superhelix line up and form channels through which the histone N-termini domains protrude from the core. These regions are in the form of “tails” that appear to lack secondary structure and are subject to various enzyme catalysed, post-translational modifications which affect their charge and can influence the degree of chromatin compaction. The tightness with which DNA is packaged into chromatin will limit the binding and function of proteins that mediate transcriptional regulation, and this will therefore influence the transcriptional competence of any given gene in such a chromatin environment.

Covalent post-translational acetylation and deacetylation of specific lysine residues in the histone N-termini is one of the most widely studied chromatin modifications. In the past four years there have been rapid advances in identifying the enzymes and multiprotein complexes that bring about histone acetylation (the family of histone acetyltransferases or HAT coactivators) and deacetylation (the histone deacetylases or HDAC corepressors). This review will focus on some of the clinical aspects of this recent work on acetylation and the intimate connection that it is now known to have with the methylation of cytosine residues in DNA. A third type of chromatin remodelling is the direct physical repositioning or disruption of nucleosomes mediated by a family of DNA dependent ATPases. The connection between this latter type of remodelling and either histone acetylation or DNA methylation is complicated, but progress is being made. For example, the NuRD multiprotein complex (see below, fig 1C) contains histone deacetylase and chromatin remodelling activities, as well as the methyl DNA binding protein MBD3, which suggests that a profound interplay between these modifications is required during gene regulation. Therefore, it is probable that a particular pathogenesis may be caused by defects in more than one type of chromatin modification. Relevant pathologies and syndromes are discussed in following sections and are summarised in table 1.

Histone acetylation, protein acetylation, and gene regulation

HISTONE DEACETYLASES AND COREPRESSOR COMPLEXES

Deacetylation of histones is, in general, associated with repression of gene transcription, presumably because the highly positively charged N-terminal tails of the core histones can now interact with DNA on the nucleosome surface and in the linker DNA. In addition, the positively charged lysines in the H4 tail may interact with the negative face of an H2A-H2B dimer from a neighbouring nucleosome, and hence bring about further compaction of the chromatin. Deacetylation is brought about by the action of the histone deacetylases (HDACs), which would therefore enhance histone-histone interactions by maintaining the positively charged (unmodified) state of lysines in the histone tails. HDACs are now known to be corepressor components of many multiprotein complexes that modify and remodel chromatin.

Targeting of complexes containing HDAC1 and HDAC2 is achieved by the interaction of the repressor proteins Sin3A, Sin3B, and other Sin3 associated proteins (SAPs) in a large multiprotein complex that comprises at least seven subunits (fig 1A). The mammalian Sin3 complex mediates repression for an extensive and ever growing list of transcriptional regulator proteins, which include DNA binding components such as the Mad/
Max heterodimer and nuclear hormone receptors (see below, fig 1B). The members of the Mad/Mxi1 family are able to replace Myc in the Myc/Max heterodimer, and can therefore repress transcription at promoters with Myc consensus DNA binding sites. Mutations in the Sin3 interaction domains (SIDs) of Mad/Mxi1 can abolish binding of the corepressors to the Sin3 proteins and hence HDACs, and this correlates with the abolition of transcriptional repression and anti-oncogenic activity. In addition, transfection studies have shown that HDACs and Mad cooperate to repress cell proliferation. In a similar mechanism, the transcriptional corepressor proteins N-CoR (nuclear hormone receptor corepressor) and SMRT (silencing mediator of retinoid and thyroid hormone receptor) target deacetylase activity to non-liganded thyroid hormone and retinoic acid nuclear receptors and to antagonist bound oestrogen and progesterone receptors. As discussed below, the presence of receptor ligands, for example, retinoic acid, appears to induce an exchange of the corepressor complexes containing HDACs for those with coactivator functions that contain histone acetyltransferase (HAT) activities (compare figs 1B and 2A).

Histone deacetylases are also recruited by the retinoblastoma protein pRb (fig 1A), the product of a tumour suppressor gene, and an inhibitor of cell proliferation. The inhibitory action of pRb is the result, in part, of its ability to bind to the E2F family of DNA binding transcription factors, which results in the sequestration of E2F and repression of E2F target genes during the G1 phase of the cell cycle. The interaction is mediated by the A/B pocket domain in pRb, and it is no coincidence that the great majority of Rb mutations in human tumours are located in this domain. The pocket domain can also interact with a variety of other cellular proteins, including viral transforming oncoproteins (such as E1A from adenovirus and SV40 large T) and histone deacetylases (HDAC1 and HDAC2) that share the common LXCXE motif which allows interactions with pocket proteins. However, the interaction between Rb and either HDACs or viral oncoproteins appears to be competitive, and the Rb-HDAC1 interaction may be one of the intracellular targets for these transforming proteins. The Rb-HDAC interaction has been analysed by transient transfection experiments, which show that Rb and the HDACs cooperate in repressing an E2F1-driven promoter, and the repression exerted by Rb and other pocket proteins during the G1 phase of the cell cycle can be reversed by treatment with chemical inhibitors of HDACs (see below). The HDAC inhibitors can also upregulate some of the E2F target genes. It is probable that aberrant targeting of deacetylase activity and incorrect chromatin remodelling are one step in the process of transformation and implies that these processes have a fundamental role in the suppression of carcinogenesis.

Figure 1  Schematic representation of multiprotein complexes that mediate chromatin modifications. Some of the known components of the complexes are shown on the left, with arrows indicating additional interactions with other proteins on the right (refer to main text for details). Histone deacetylases (HDAC) are shown in purple, components of the DNA methylation system in yellow, and ATPase/helicase that mediates chromatin remodelling in green. Other corepressor accessory proteins in the complexes (panels A-C) are shown in grey. (A) Components of the HDAC/Sin complex and known interacting proteins. (B) Additional interactions of the HDAC/Sin3 complex, mediated by the corepressors NcoR/SMRT, with unliganded nuclear receptors and leukaemogenic fusion proteins. Note that these interactions occur at low concentrations, or in the absence of the receptor ligand. (C) Components and interactions of the NuRD complex.

Figure 2  Schematic representation of interactions mediated by histone acetyltransferases (HATs). Unbroken arrows indicate known interactions of HATs with the proteins listed on the right (refer to main text for details). Histone acetyltransferases are shown in red, the SWI/SNF complex that mediates chromatin remodelling in green, and other accessory proteins in grey. (A) Interactions of three human histone acetyltransferases. The p300/CBP coactivator can interact with nuclear receptors at physiological concentrations of the receptor ligand, with other transcription factors and with the histone acetyltransferase PCAF. PCAF can also interact with the coactivator ACTR and transcription factors. (B) Potential interactions of the human and/or yeast Hat1 protein with both RbAp46 and, through the bromodomain, with histones H3 and H4. Histone H4 can also interact directly with both RbAp46 and RbAp48. The bromodomain may also mediate an interaction with the yeast SWI/SNF complex (broken arrow).
the case of pRb, histone modification is implicated in the induction of cell cycle arrest which may explain why the Rb gene is mutated in almost all cancer cells.

In addition to E2F and HDACs, Rb has also been shown to interact with other proteins that regulate chromatin modifications; TAFII 250 is a transcription factor that has intrinsic histone acetyltransferase (HAT) activity; BRG1 is a transcriptional activator and ATPase/helicase that is a component of the mammalian SWI/SNF chromatin remodelling complex; and Rb associated protein 48 (RbAp48) is a component of the chromatin assembly factor CAF-1, interacts with HDAC1, and is a component of the human Hat1 acetyltransferase (fig 2B). RbAp48, and the related protein RbAp46, can bind directly to an alpha helix in histone H4, so it is likely that these proteins mediate core histone binding for the Sin3-HDAC complex, CAF-1 and Hat1.

A similar network of interactions that regulate cell proliferation is also seen with the HATs (see below) and for histone deacetylase complexes other than the Sin3-HDAC complex. For example, a novel multiprotein complex has been isolated recently that contains both nucleosome remodelling and histone deacetylase activities (hence the NuRD complex, fig 1). The NuRD complex contains, in addition to HDAC1 and HDAC2, the dermatomyositis specific autoantigen Mi-2 which contains an ATPase/helicase domain of the SWI/SNF type (see below). In addition, the NuRD complex contains a protein, MTA2, that is homologous to the metastasis associated protein MTA1, which is expressed at high levels in several human cancer cell lines and tissues. The NuRD complex can also be recruited to DNA by specific DNA binding factors, in a similar mechanism to the targeting of the Sin3-HDAC complex. For example, two determinants of the lymphoid lineage in T cells, Ikaros and Aiolos, have been shown to associate with the NuRD complex. However, Ikaros and Aiolos can also function as transcriptional repressors during lymphocyte development by recruiting the Sin3-HDAC complex.

**Table 1** Summary of human diseases in which a defect in chromatin modification and remodelling is believed to contribute to a clinical pathology. The defects can arise from mutations in gene products, or by the aberrant recruitment of other proteins that are components of multiprotein complexes. Refer to the main text for details.

<table>
<thead>
<tr>
<th>Type of chromatin modification</th>
<th>Enzymes and other proteins that mediate the chromatin modification</th>
<th>Mutated genes implicated as a cause of a pathogenesis</th>
<th>Relevant proteins that interact with the enzymes etc</th>
<th>Clinical conditions and pathologies that correlate with a listed mutation or protein-protein interaction</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histone acetylation</td>
<td>Histone deacetylases</td>
<td>Rb</td>
<td>pRb</td>
<td>Neoplasia</td>
<td>24, 25, 26</td>
</tr>
<tr>
<td></td>
<td>HDAC1/2</td>
<td>p53</td>
<td>pRb</td>
<td>Neoplasia</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mi-2</td>
<td>RARα</td>
<td>Dermatoyositis and neoplasia</td>
<td>33, 34, 122</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MTA2</td>
<td>Fusion proteins</td>
<td>Metastasis</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Histone acetyltransferases</td>
<td>ElA</td>
<td>Fusion proteins</td>
<td>Acute promyelocytic leukaemia</td>
<td>39, 40, 47, 48</td>
</tr>
<tr>
<td></td>
<td>p300/CBP</td>
<td>MLL and MOZ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histone phosphorylation</td>
<td>TAF II 250</td>
<td>RARα</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA methylation</td>
<td>DNA methyltransferase</td>
<td>RSK2</td>
<td>Not known</td>
<td>Loss of tumour suppression?</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>DNA methyltransferase</td>
<td>DNMT3B</td>
<td>Not known</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Methyl-CpG binding proteins</td>
<td>MECP2</td>
<td>Not known</td>
<td>ICF syndrome</td>
<td>113, 114</td>
</tr>
<tr>
<td>Chromatin remodelling</td>
<td>ATPase/helicases</td>
<td>MBD4</td>
<td>Not known</td>
<td>Rett syndrome</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>Mi-2</td>
<td>NuRD complex</td>
<td></td>
<td>Micrositate unstable colon cancers</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>ATRX</td>
<td></td>
<td>Neoplasia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BRG-1/BRM</td>
<td></td>
<td>eRb</td>
<td>derivative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hSW1/SN1 complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SW1/SN2 complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HISTONE DEACETYLASES AND LEUKAEMOGENIC FUSION PROTEINS**

The interaction of HDACs with chimeric mutant proteins of the retinoic acid receptor is one of the mechanisms that underlies the molecular pathogenesis of acute promyelocytic leukaemia (APL). One of the best characterised forms of acute myeloid leukaemia (AML). In this disease, chromosomal translocations create fusion proteins of retinoic acid receptor-α and either PML (for promyelocytic leukaemia) or, in rare cases, PLZF (for promyelocytic leukaemia zinc finger). In the case of PML/RAR-α forms of APL, the fusion protein retains the ability of the wild type nuclear receptor to recruit the N-CoR/HDAC complex (fig 1B) and to block haematopoietic differentiation. However, physiological concentrations of retinoic acid dissociate the corepressor complex from wild type RAR-α (see above and fig 2A), so that it can therefore function as a transcriptional activator. In contrast, the fusion protein retains the ability to bind to the corepressor complex under these conditions, thereby constitutively repressing RAR-α target genes. Treatment with higher, pharmacological concentrations of the hormone overcomes this interaction and converts the PML/RAR-α fusion protein back into an activator. As a consequence, cell proliferation is inhibited and neutrophilic differentiation of neoplastic cells is induced, which is the basis of differentiation therapy of APL. One of the target genes of PML/RAR-α during RA induced differentiation of APL cells is that encoding the cyclin-CDK inhibitor p21 also known as WAF1 and CIP1, which exerts a G1 cell cycle arrest in response to a variety of stimuli.

In contrast, cells expressing the PLZF/ RAR-α fusion are not sensitive to RA induced differentiation, and patients with this type of
APL respond poorly, if at all, to pharmacological doses of the hormone. The interaction between PLZF/RAR-α and the corepressor complex is resistant to retinoic acid because the wild type PLZF protein can itself interact directly with corepressors such as SMRT and N-CoR. This interaction is mediated by the broad complex/tramtrack/bric a brac/poxviruses and zinc finger (BTB/POZ) repression domain from PLZF, which also allows SMRT to interact with another BTB/POZ oncoprotein, LAZ3/BCL6. In contrast, other transcriptional repressors that also contain the BTB/POZ domain, such as the product of the putative tumour suppressor gene HIC-1 (for hypermethylated in cancer), do not recruit the SMRT/N-CoR-histone deacetylase complex as a general mechanism to repress transcription. However, since HDAC inhibitors restore the retinoic responses of RA resistant APL cell lines, clinical and cytogenetic remission of a PLZF/RAR-α type of APL has been achieved with a combination therapy of retinoic acid and phenylbutyrate, an HDAC inhibitor. Such combination therapies may therefore be applicable to other types of neoplastic diseases that are associated with oncogenic repression of gene transcription by histone deacetylases.

HISTONE DEACETYLASE INHIBITORS AND CANCER CHEMOTHERAPY

Inhibitors of HDACs have also received considerable attention as possible therapeutic agents to induce growth arrest and terminal differentiation in malignant cells and therefore prevent the progression of cancers. HDAC inhibitors can be classified on the basis of structure and mode of inhibition. Reversible inhibitors include n-butyric acid and other related short chain fatty acids, the microbial antibiotic trichostatin A (TSA), and hybrid polar compounds such as suberoylanilide hydroxamic acid (SAHA). The carboxylic and hydroxamic acid groups in this set of compounds are likely to be specific ligands of a catalytic zinc ion at the active site of HDAC because a similar zinc binding site exists in a prokaryotic homologue of the enzyme. Irreversible inhibitors include trapoxin and trapoxin related natural products, such as chlamydacin and HC toxin. These compounds are tetrapartite cyclic peptides with 2-amino-9,10-epoxy-8-oxodecanoic acid (Aeo) as one residue. Aeo contains an epoxyketone group that is isosteric with N-acetyl lysine and is presumed to inhibit HDACs by binding covalently and irreversibly to nuleicophile groups in the active site of the enzymes. TSA, trapoxin, and depudicin (a fungal metabolite that resembles Aeo) have all been shown to revert the morphologies of oncogene transformed cells and cells derived from tumours to those with a normal cytoskeletal architecture. TSA and SAHA are also potent inducers of transformed cell differentiation and apoptosis. A novel antibiotic and HDAC inhibitor, FR901228, can strongly inhibit proliferation of tumour cells in vitro by arresting cell cycle transition at G1 and G2/M phases, a property that is common between the unrelated types of HDAC inhibitors. For this reason, HDAC inhibitors are being assessed as therapeutic agents for cancer chemotherapy and as adjuncts to established agents such as retinoic acid (see above). In view of this, it is interesting to note that n-butyric acid causes growth inhibition, differentiation, and apoptosis in colon cancer derived cell lines. Butyrate is a natural fermentation product of certain dietary fibres by anaerobic bacteria in the lumen of the colon. The production of this short chain fatty acid may be the key factor that allows dietary fibre, for example fibre from wheat, to protect against colon carcinogenesis. The molecular mechanism that underlies this protection is presumed to be inhibition of histone deacetylase activity and induction of histone hyperacetylation by butyrate. Archer et al showed that butyrate induces expression of the G1 cell cycle inhibitor p21 gene (see above), presumably in response to hyperacetylation of the p21 promoter, which induces growth arrest in colon cancer cells. Trapoxin has also been shown to increase p21 expression in human tumour cells, with an increase of histone H3 acetylation at the p21 promoter and the induction of apoptosis in a cell line that contained wild type p53.

HISTONE ACETYLTRANSFERASES AND COACTIVATOR COMPLEXES

In parallel with the developments in our understanding of HDACs, there have also been the recent identification and functional analysis of several human histone acetyltransferases (HATs) and the mechanisms by which histone acetylation can activate transcription. Histone acetylation, at lysines within the N-terminal tails of H3 and H4, appears to be a prerequisite for the process of transcriptional activation in vivo. The acetylation appears to mediate chromatin remodelling (which makes the chromatin more accessible to transcription factors) by the specific, targeted interaction of coactivator proteins that contain a bromodomain with acetyl lysines. In particular, the bromodomain of a yeast nuclear HAT is able to bind in vitro to acetylated peptides of the H3 and H4 N-terminal tails and the bromodomain is required for the subsequent recruitment and coordination of in vivo remodelling activity by the yeast SWI/SNF complex (fig 2B).

Recent work has shown the intimate relationship between histone modifications and the processes of cell proliferation, cell differentiation, and oncogenesis. As discussed in previous sections, nuclear receptors have a key role in determining the balance of cell proliferation and cell differentiation in response to extracellular signals, such as the hormone retinoic acid. In the absence of ligand, the receptors remain constitutively bound to target promoters and recruit a transcriptional repressor complex (see above), which maintains the cell in a proliferating state. However, once the nuclear receptors bind the ligand, the repressor complex is replaced by an activator complex that induces cell differentiation (fig 2A). Components of the activator complexes include CREB binding
Chromatin modification and disease

Transcription during apoptosis by recruiting acetyltransferase.82 Associated factor (PCAF) that is itself a histone associated with other DNA binding transcription factors, such as c-Jun, c-Myb, c-Fos, and MyoD83 (fig 2A), as well as p300/CBP associated factor (PCAF) that is itself a histone acetyltransferase.82 The studies by Yang et al82 also show that the E1A oncoprotein stimulates proliferation by disrupting the interaction of p300/CBP and PCAF, which would normally suppress cell growth. PCAF is also implicated in nuclear receptor coactivation.81 83 Two additional nuclear hormone receptor coactivators, SRC-1 and ACTR, which are members of the p160 family, are also HATs84 85 (see below). The human transcription factor TAF II 250, a component of the TFIID complex that recognizes the TATA element at promoters, also contains HAT activity.86 A further level of complexity in the regulation of gene expression by HATs is that the enzymes can acetylate lysine groups of non-histone proteins. For example, p300/CBP and PCAF are able to acetylate transcription factors such as p53 and components of the general transcription machinery such as the TFIIEβ subunit.87 88 Acetylation of both p53 and the haematopoietic transcription factor GATA-1 increased their DNA binding activities.87 88 In the latter study, the acetylation of GATA-1 in vivo was implicated in the activation of target genes. A similar mechanism of regulation has been shown for the nuclear hormone receptor coactivator and acetyltransferase ACTR, which can itself be acetylated by p300/CBP.89 As expected, hormone treatment of cells caused an increase in histone acetylation at receptor target genes. However, this effect was transient because the subsequent acetylation of ACTR led to dissociation of the receptor-coactivator complex and down regulation of transcription, and suggests that non-histone protein acetylation is the key element that autoregulates hormone induction. In a separate study, the acetylation of E2F1 by both PCAF and p300/CBP appears to enhance the function of this transcription factor,90 which would presumably stimulate the transcription of target genes during S phase of the cell cycle. As discussed above, the E2F transcription factors repress transcription of target genes during G1 phase by interacting with the Rb tumour suppressor gene product pRb and histone deacetylases (HDACs). In addition, the p53 tumour suppressor gene product has also been shown to repress transcription during apoptosis by recruiting the Sin3-HDAC complex,91 and itself to become acetylated at defined lysine residues in vivo.92 These observations suggest that transcriptional regulation involves the addition and removal of acetyl groups not only from histones but also from other nuclear proteins, which has important implications for the understanding of cell growth and oncogenesis.

Other histone modifications

The N-terminal tails of the core histones are subject to other post-translational modifications, which include phosphorylation, methylation, ADP ribosylation, and ubiquitination.93 Phosphorylation of histone H3 at serine 10 has been the most extensively studied modification to date. It occurs during mitosis in many cells94 as a transient and rapid response after mitogen stimulation, which correlates with the expression of c-fos and c-jun.95 H3 phosphorylation in response to epidermal growth factor (EGF) appears to be mediated by the Rsk-2 kinase,96 which is a member of the p90rsk (ribosomal S6 kinase) family that is implicated in cell proliferation and differentiation.97 It is interesting to note that mutations in Rsk-2 (but not in any other kinase in this family) are associated with Coffin-Lowry syndrome,98 which is characterized by severe psychomotor retardation, facial and digital dysmorphism, and skeletal deformations. However, it remains unclear what role H3 phosphorylation has during cellular response to mitogens and if it has a direct involvement in gene regulation.

Methylation of histone H3 has been linked to gene activation by the p160 family of coactivators, which include the histone acetyltransferases SRC-1 and ACTR (see above). Both of these proteins, and a third member of this family called GRIP1, have been shown to interact with the coactivator associated arginine (R) methyltransferase (CARM1), which has extensive homology to other arginine methyltransferases.99 CARM1 can also methylate histone H3 in vitro, and enhances the expression of a reporter gene in a transient transfection experiment. Methylation of histones, or other proteins in the transcription initiation complex, may therefore be a gene regulatory mechanism that cooperates with histone and protein acetylation.
DNA methylation and gene silencing

At the level of DNA modifications, methylation of the C5 atom of cytosine residues is a powerful and prevalent mechanism for the repression and inactivation of genes. Methylation of promoter regions, on opposite DNA strands of the dinucleotide sequence CpG, correlates, in general, with transcriptional inhibition. Hydrolytic deamination of 5'-methylcytosine to thymine residues generates T-G mispairs, which contribute to many germline point mutations associated with human genetic disease and other somatic mutations that lead to cancer. For this reason, CpG dinucleotides tend not to be found in coding regions, but are clustered near the promoters of widely expressed housekeeping genes, but remain unmethylated at all levels of expression of the gene. CpG islands can become de novo methylated during normal development (to silence imprinted genes and the genes on the inactive X chromosome of female mammals), but the CpG islands of autosomal genes can also become methylated as a result of in vitro cell culture or neoplasia. It is possible that the silencing of certain tumour suppressor genes may, in part, be the consequence of de novo methylation of an adjacent CpG island. For example, the frequency with which promoter methylation contributes to the gene inactivation is 33% for VHL and 84% for MLH1, in von Hippel-Lindau (VHL) disease and microsatellite unstable colorectal tumours, respectively. A causal link between promoter methylation and carcinogenesis is implied from studies of transgenic mice with targeted deletions of the DNA methyltransferase 1 (Dnmt1) gene, in which the formation of intestinal polyps was suppressed. Dnmt1 can also recruit histone deacetylase activity and can interact with HDAC1 in vitro. Dnmt1 methylates DNA containing hemimethylated CpG dinucleotides more efficiently than unmethylated DNA, and is therefore presumed to be the major maintenance methyltransferase in vivo. Since Dnmt1 protein colocalises with replication foci, it will be interesting to see if HDACs have a particular role at the replication fork. Other mammalian DNA methyltransferases (Dnmt3a and b) are likely to de novo methylate promoters, with Dnmt3b specifically required to methylate centromeric minor satellite (in mouse embryonic stem cells). Mutations in one of the conserved catalytic domains of DNMT3B, which presumably cause a partial loss of function of the enzyme, are associated with ICF syndrome (for immunodeficiency, centromeric instability, facial anomalies). ICF syndrome is characterised by immunological defects, hypomethylation, and instability of centromeric heterochromatin, and facial anomalies such as hypertelorism, epicanthic folds, and macroGLOSSIA.

The mechanism by which methylated promoters are inactivated has also become clearer and appears to involve a long term remodelling of the chromatin at the promoter. DNA methylation is able to reduce the binding affinity of sequence specific transcription factors or recruit sequence specific DNA binding proteins, such as methylated DNA binding protein (MDBP), which may act as transcriptional repressors. However, a more general process recruits sequence non-specific methyl-CpG binding proteins (MeCPs) that exclude transcription factors from the methylated promoter. It is now clear that the key molecular mechanism that underlies this repression involves the recruitment of histone deacetylases (HDACs, see above and fig 1A). The MeCP2 protein contains a methyl binding domain (MBD), which allows it to bind to a single, symmetrically methylated CpG site and a transcriptional repression domain (TRD) which recruits the Sin3-HDAC corepressor complex. This mechanism is now known to be implicated in repression by members of the MBD protein family, which were identified on the basis of homology with the MBD domain of MeCP2. MBD2 forms the so called MeCP1 complex, together with HDAC1, HDAC2, and RbAp46 and 48, and MBD3 is a component of the Mi2-NuRD deacetylase complex (see above). It is interesting to note that MBD4 contains the canonical methyl CpG binding domain in addition to a thymine DNA glycosylase catalytic domain, which binds preferentially to methyl CpG-TpG mismatches in DNA. It is therefore improbable that the function of the MBD4 enzyme is to suppress mutation at methyl CpG, rather than to act as a transcriptional repressor. Frameshift mutations in MBD4 that would cause truncation of the protein between the MBD and glycosylase domains, and hence cause a defect in mismatch repair, correlate with over 40% of microsatellite unstable sporadic colon cancers.

Mutations in the MECP2 gene, which encodes the X linked methyl CpG binding protein 2 (MeCP2, see above), are of particular clinical significance. Genetic linkage analysis has established that de novo missense mutations in the methyl CpG binding domain (MBD), and other mutations that disrupt the transcriptional repression domain (TRD), are a cause of Rett syndrome in about a quarter of the sporadic patients studied. Rett syndrome is a progressive neurodevelopmental disorder that occurs almost exclusively in females, and is typified by the onset of autism, dementia, ataxia, and loss of purposeful hand movements from the ages of 6 to 18 months. An additional study has made the suggestion that the effect of MECP2 mutations may not be limited to Rett syndrome, since heterozygote females with skewed X inactivation patterns may have a mild disease phenotype. However, it is clear that the mechanism that underlies the pathogenesis of Rett implicates a dysregulation in chromatin remodelling, although it is not yet known if a similar pathogenesis is seen for mutations in the MBD family, or their partners in multiprotein complexes.

Chromatin remodelling and carcinogenesis

A further causal link between epigenetic dysregulation and carcinogenesis is provided.
by defects in ATP utilising chromatin remodelling complexes. The role of these complexes is to physically reposition or disrupt nucleosomes by altering histone-DNA contacts. The ATPase/helicase subunits of the complexes are members of three related families: homologues of yeast SWI2/SNF2 (for mating type switch/sucrose non-fermenting 2), the Mi-2 family (also known as the CHD family), and the ISWI (imitation SWI) family.

As discussed above, the dermatomyositis specific autoantigen and ATPase/helicase Mi-2β (also called CHD4) is a component of the NuRD complex which contains the histone deacetylases HDAC1 and HDAC2 (fig 1C). This observation suggests that chromatin remodelling by the ATPase activity makes the histone tails more accessible to the NuRD deacetylases, with the subsequent formation of a repressive chromatin structure. However, neither the effect of NuRD mediated histone deacetylation on gene expression nor the targets of this type of chromatin remodelling are known, although the MBD3 component of the complex may recruit it to extensive, methylated regions of the genome such as heterochromatin. Since 15% to 30% of patients with dermatomyositis develop cancer, it is probable that changes in chromatin modification and remodelling can affect cell proliferation.

It is interesting to note that the Mi-2β protein contains two cysteine rich PHD (plant homeodomain)/zinc finger regions and two chromodomains, in addition to the helicase/ATPase domain. Homeodomains and chromodomains are common features of other transcriptional regulators that bind to chromatin, which include the DNMT3 family of DNA methyltransferases and ALL-1 (also known as MLL, HRX, or HTRX, see below). Another such protein is ATRX (for α thalassaemia retardation on the X chromosome) which, like Mi-2β, also contains PHD-like fingers and other domains that classify it as a member of the ATPase/helicase superfamily. The functional importance of the PHD-like fingers is shown by the observation that two thirds of all mutations that cause ATRX syndrome lie in this region. ATRX syndrome comprises a severe form of mental retardation, characterised by the presence of α thalassaemia, urogenital abnormalities, and facial dysmorphism.

In view of the interaction of Mi-2β protein with HDAC1, it is probable that PHD-like domains in other transcriptional regulators also mediate the same interaction. In addition, mutations in ATRX are associated with changes in DNA methylation patterns at highly repeated sequences, such that rDNA repeat arrays are hypomethylated whereas others (the Y specific repeat DYZ2) are hypermethylated. This observation implies a link between chromatin remodelling, mediated by the ATRX protein, and the DNA methylation system. However, it is unclear if ATRX functions as a transcriptional coactivator (similar to the SWI2/SNF2 proteins) to increase the expression of a component of the DNA methylation system, or if it increases the accessibility of chromatin to a DNA methyltransferase.

The human SWI/SNF complex has been found to act as a coactivator for several nuclear receptors and to remodel promoter regions to facilitate the binding of other factors. The two human homologues of SWI/SNF2, BRG-1 and hBRM, are also implicated in transcriptional repression, since both can interact with the tumour suppressor gene product Rb. As discussed above, Rb regulates cell cycle progression by inhibiting the activity of the transcriptional activator E2F and recruiting histone deacetylases. Transient transfection of BRG-1 or hBRM into cell lines that do not express these proteins but that do express Rb leads to growth arrest and, furthermore, fibroblasts transformed with ras and lacking endogenous expression of BRM revert to a flattened, growth arrested phenotype after reintroduction of hBRM. These observations establish the link between chromatin remodelling by SWI/SNF and control of the cell cycle. A further connection is provided by another subunit of the SWI/SNF complex, called hSNF5/INI1. Genetic linkage has identified biallelic deletions or mutations in the hSNF5/INI1 gene to be responsible for malignant rhabdoid tumours, which are very aggressive cancers of early childhood that tend to occur in the kidney, brain, and soft tissues. Another study has shown that deletions of hSNF5/INI1 can be acquired during leukaemogenesis in patients with chronic myeloid leukaemia.

Maintenance of chromatin states and carcinogenesis

Recent work has shown that the maintenance of active and inactive chromatin states is an important determinant of gene expression. In Drosophila, the coordinated expression of the homeotic genes determines segmentation and body plan along the anterior-posterior axis. The maintenance of this expression throughout development is mediated by the Polycomb-group (Pc-G) repressor proteins and the trithorax-group (trx-G) activator proteins. Many of the polycomb- and trithorax-group proteins contain the SET domain (for Suvar3-9, Enhancer-of-zeste, Trithorax), that is conserved in a number of mammalian homologues. ALL-1, for example, is the human homologue of Drosophila Trithorax protein, which is a positive regulator of homeotic gene expression. A similar role is implied for the mammalian MLL/ALL-1 because heterozygous mll+/- mice have defects in axial skeletal development and haematopoiesis, which arise from alterations in the pattern of Hox gene expression. The ALL-1 gene, at the 11q23 locus, is one of the most common targets of chromosomal translocations in acute lymphocytic leukaemia and other acute leukemias, and there are at least 30 partner genes that produce in frame leukogenic fusion proteins with ALL-1/MLL, including the histone acetyltransferase CBP (see above). In general, the C-terminal SET domain of
ALL-1/MLL is lost during chromosomal translocation. It is interesting to note that the hSNF5/INI1 component of the SWI/SNF complex (see above) has been shown to interact with the SET domain of ALL-1. One of the consequences of ALL-1/MLL translocations would be the inability to recruit SWI/SNF, and it is therefore possible that the transformation of haematopoietic cells involves the dysregulation of chromatin remodelling, as well as the aberrant expression of MLL/ALL-1 target genes.

Mammalian polycomb-group homologues are thought to be negative regulators of homeotic gene expression. EED (for embryonic ectoderm development) is required at a very early stage in embryonic development and is the only known homologue of the *Drosophila* extra sex combs (esc) Pc-G protein. Disruption of the *eed* gene in mice causes defects in anterior mesoderm production, followed by death at day 8.5 of gestation. In addition, the histone deacetylases HDAC1 and HDAC2 have been shown to interact with EED, but not with other vertebrate Pc-G proteins. This result again underlines the complex and subtle interplay between the separate mechanisms of gene regulation.

**Conclusion**

In the past few years, the exciting progress in the field of gene regulation has made it clear that chromatin is not just a static structure, but that it has a pivotal role in regulating transcription. Chromatin has a dynamic structure that can be modulated during cell differentiation and transformation, and the nucleosome is a substrate that can both receive and transmit intracellular signals. In addition, chromatin has the capacity to encode epigenetic information on levels of gene expression, which is independent of the genetic information encoded by the sequence of genes. A parallel development is the realisation that many transcription factors can function by recruiting large, multi-protein complexes that mediate several types of chromatin modification and remodelling events. It is therefore not surprising that mutations that cause loss or perturbation of chromatin modification and remodelling activities will cause changes in gene expression levels. This points to abnormal epigenetic regulation as a general mechanism that underlies human carcinogenesis and the pathogenesis of other genetic diseases (see table 1 for a summary). For example, the types of mutations that are associated with most classes of acute myeloid leukaemia are typified by balanced chromosomal translocations that result in the expression of chimaeric proteins. Most of these chimaeric proteins involve the in frame fusion of transcriptional regulator that results in tran-chimaeric proteins. Most of these involve the in frame fusion of transcriptional regulator that results in transcriptional dysregulation. Some of the proteins that mediate chromatin modification and remodelling are therefore novel molecular targets for future treatments of cancer and other genetic diseases.

One of the most intriguing aspects of this recent work is the probable interplay between modification and remodelling events. Future work will undoubtedly uncover further mechanistic links between distinct gene regulation systems. But a hint of this interplay is seen with the possible deacetylations of histone deacetylase. As discussed above, HDAC1 and HDAC2 can interact with the DNA methyltransferase Dnmt1, the polycomb-like protein EED, and the ATPase/helicase Mi-2. To add to this complexity, it is not known if other histone and protein post-translational modifications (notably histone phosphorylation and methylation) can modulate the effects of chromatin modification and remodelling, or if they participate in other gene regulation systems of their own. Another important aspect is a description and understanding of the signal transduction pathways that use chromatin as a target, and how these signals can mediate changes in gene expression. For example, a very recent development has shown that recombinant yeast and mouse Sir2 proteins, and other yeast Sir2 homologues, are novel, NAD dependent histone deacetylases. In yeast, Sir2 is a mediator of transcriptional silencing at the heterochromatic regions associated with silent mating loci, telomeres, and ribosomal DNA. However, the absolute requirement of NAD for the deacetylation reaction suggests that this form of chromatin modification can be regulated by the metabolic levels of NAD and NADH, and hence by calorific intake.
Chromatin modification and disease


83 Okano M, Bell DW, Haber DA and Li E. DNA methyltransferase Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99:249-59.
84 Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qin X, Russo J, Vegas-Pequignot E. Chromosome instability and immunodeficiency syn-


