Reading the DNA Methylation Signal

A. BIRD AND D. MACLEOD
Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland, United Kingdom

A cell’s properties depend largely on its pattern of gene expression. During the development of complex multicellular organisms, multiple cell lineages emerge among the descendants of the single zygotic cell. Within each lineage, the range of gene expression programs is effectively restricted, culminating in the definitive program of the final differentiated cell. As part of this process, some genes are selected for future activity, whereas others are disqualified by long-term silencing. For a specific gene in one cell type, silence versus activity is not determined solely by the available combinations of transcription factors in that cell, but also by local differentiation of chromatin structure near to or including the gene. A key feature of such local differentiation of the genome is its stability within and between cell generations. The study of these processes has come to be known as epigenetics. Contemporary epigenetics might therefore be defined as, “the study of the structural adaptation of chromosome regions so as to perpetuate local activity states.”

Three sorts of epigenetic system have traditionally been studied: DNA methylation, the polycomb/trithorax system, and histone modification or variant substitution. Other epigenetic categories have emerged, however, and are the subject of intense study as reflected elsewhere in this volume. They include silencing based on noncoding RNAs, transcriptional silencing triggered by double-stranded RNA interference, and silencing or activity that responds to localization of genes within the nucleus. All of these processes appear to be closely interwoven with histone modification, which is itself a diverse and complex system of chromosome marking. For each system there is evidence that the trigger for altered chromatin is a preexisting transcriptional state. In other words, it appears to be the transcription factors that primarily determine whether a gene is active or inactive, after which chromatin structure is adapted accordingly and the state is memorized. This “secondary” status of many epigenetic mechanisms does not diminish their importance. Stable, leakproof gene silencing requires sophisticated mechanisms for sensing the transcriptional status quo without perturbing it and then passing the memory between cell generations without disruption by DNA replication.

Obviously the most direct way of marking genes either for silence or activity would be to apply chemical tags to the DNA itself. In practice, however, the only widespread modification of DNA is methylation, either of the C5 position of the cytosine ring or of the N6 position of the adenine ring. Other modifications are found, but have limited phylogenetic distribution. In multicellular eukaryotes, only m5C is convincingly reported. A plausible explanation for the narrow spectrum of DNA modifications is that chemical moieties, such as acetyl or phosphate groups, or methylation at other positions within the DNA structure would interfere with aspects of DNA function. Therefore the scope for chemically tagging genomic DNA, without compromising its essential information coding and storage properties, may be extremely limited. Even m5C, which does not alter the base-pairing specificity of cytosine and has survived several billion years of scrutiny by natural selection, has drawbacks, as it is a mutagenic base in vertebrates and bacteria (Coulondre et al. 1978; Bird 1980).

This trade-off between benefit and disadvantage may explain the evolutionary volatility of cytosine methylation across evolutionary time. Some organisms have dispensed with DNA methylation whereas their relatives retain it. For example, the fungus Neurospora crassa displays cytosine methylation, but the fission and budding yeasts—also fungi—do not.

The emergence of complex gene expression programs in multicellular organisms with functionally adapted cell types may have depended on the ability to epigenetically mark genes for either activity or silence during development. Without an effective memory system to enforce the activity status established at critical developmental stages, it would be necessary to constantly reiterate the initial conditions in order to reestablish the program. DNA is the obvious substrate for memorable marks of this kind, but, as pointed out above, has a limited capacity for chemical modification. This shortcoming of the genetic material may have provided a driving force behind the evolution of histones. Rather than marking the DNA itself, why not coat it uniformly with proteins that are susceptible to a plethora of chemical modifications? (Jenuwein and Allis 2001) A precondition of such a system is that histones do not move around. If they did, marks would not be loci specific and would lose their value. Histone loss by degradation or stochastic processes would also need to be compensated by ensuring that replacement histones acquired modifications that are appropriate to the region. What we know about histones supports these requirements. The half-life of histones on DNA appears to be long and there is evidence for processes that transfer modification from one nucleosome to its (perhaps new) neighbor. An example is the interaction of H1P1 with both methylated lysine 9 of histone H3 (H3K9) and the histone methyltransferase Suv4.3 (Bannister et al. 2001). Re-
based on work of this laboratory and others.

To understand transcriptional repression, let us consider what happens when m5C becomes spontaneously deaminated. More due to inefficient repair of G:T mismatches that arise during replication, methyl-CpG can become modified, leading to transcriptional silencing. However, the mechanisms by which this silencing is achieved are not entirely clear. For example, it has been suggested that methyl-CpG-binding proteins (MBD) can recruit histone-modifying enzymes to methylated DNA, thereby altering the chromatin structure and repressing transcription.

In this context, it is interesting to note that the H3K9 histone modification is known to be a key player in transcriptional repression. When H3K9 is methylated, it recruits histone deacetylases (HDACs) to the promoter, leading to the transcriptional silencing of genes. Similarly, the H3K27me3 modification is also involved in transcriptional repression, with the recruitment of HDACs and methyl-CpG-binding proteins to methylated DNA.

One of the most important mechanisms of transcriptional repression is the recruitment of histone-modifying enzymes. These enzymes include HDACs and histone methyltransferases, which can alter the chromatin structure and repress transcription. For example, HDACs can deacetylate histones, leading to a more compact and transcriptionally repressed chromatin structure. Similarly, histone methyltransferases can add methyl groups to histones, leading to the recruitment of repressor complexes and transcriptional silencing.

In conclusion, the mechanisms of transcriptional repression are complex and involve a variety of cellular proteins and enzymes. Understanding these mechanisms is crucial for identifying new therapeutic targets for various diseases, including cancer and autoimmune disorders. Further research is needed to fully understand the mechanisms of transcriptional repression and develop new strategies for treating these diseases.
MBD of MeCP2 and several of these (Mbd1, Mbd2, and Mbd4) were shown to bind methylated DNA (Cross et al. 1997; Hendrich and Bird 1998). Methylated DNA binding in vivo was demonstrated by assaying localization of exogenous MBD proteins to heterochromatin foci of mouse nuclei. These foci contain the mouse major satellite, which happens to possess ~40% of all methyl-CpG in the mouse nucleus. Mutations in the DNA methyltransferase gene Dnmt1 that greatly reduce genomic CpG methylation (Li et al. 1992) prevent efficient localization of MeCP2 (Nan et al. 1996), Mbd1, Mbd2, and Mbd4 (Hendrich and Bird 1998), confirming the need for methyl-CpG as their in vivo chromosomal target.

What about the function of these proteins in the cell? Attempts to answer this question are still ongoing, but much is already known. As the first MBD protein to be purified, MeCP2 led the way in functional studies. DNA methylation-dependent transcriptional repression was established using in vitro extracts and transient transfections (Nan et al. 1997). In this way, a transcriptional repression domain was identified. A significant advance in understanding occurred when our group and the group of Alan Wolfe showed that the TRD of MeCP2 recruited the Sin3A-histone deacetylase corepressor complex (Jones et al. 1998; Nan et al. 1998). Repression by MeCP2 could be relieved by the histone deacetylase inhibitor trichostatin A. These studies provided a molecular explanation for the long suspected link between DNA methylation and a repressive chromatin structure. Since then there has been progress in characterizing all MBD protein complexes. Interestingly, each MBD protein appears to specialize by associating with a different corepressor complex. At the time of writing, the MBD1 complex is not fully described, but the involvement of MCAF suggests that a histone deacetyltransferase may be involved (Fujita et al. 2003). Unexpectedly, an isoform of Mbd1 turns out to have a second DNA-binding domain specific for nonmethylated CpG, suggesting that this protein can potentially interpret CpG as a repressive signal whether it is methylated or not (Jørgensen et al. 2004). The biological rationale for this dual DNA-binding specificity is currently unknown. MBD2 associates with the NuRD or Mi-2 complex (Fig. 3) which contains histone deacetylases and the chromatin remodeling protein Mi-2. NuRD is recruited to DNA by a number of alternative DNA binding proteins (TR), of which Mbd2 is just one.

**TARGET GENES OF METHYL-CpG-BINDING PROTEINS**

Early studies of the function of methyl-CpG-binding proteins relied on model reporter gene systems. Next, it became important to assess their significance in cells of the organism by identifying target genes whose regulation is disrupted in the absence of a particular protein. Mouse gene knockouts are the system of choice for these studies. The first MBD protein gene to be disrupted was the X-linked Mecep2 gene (Chen et al. 2001; Guy et al. 2001). Two laboratories found that Mecep2-null mice are...
born and develop normally for several weeks, but they ac-
quire a variety of neurological symptoms at about 6
weeks of age leading to death at ~10 weeks. This delayed-
onset phenotype, which is fully penetrant, recalls human
Rett Syndrome, which is caused by MECP2 mutations
(Amir et al. 1999; see below). Interestingly, deletion of the
Mecp2 gene only in mouse brain cells using cre ex-
pression driven by the brain-specific nestin promoter
caused the same symptoms as deletion in the whole
mouse (Chen et al. 2001; Guy et al. 2001). Therefore, al-
though MeCP2 is ubiquitously expressed in cells of the
mouse, the Mecp2-null phenotype appears to be entirely
due to its absence in the brain. Biochemical and immuno-
cytochemical studies established that MeCP2 expression
levels are highest in the brain; specifically in neurons
(Shahbazian et al. 2002).

Knockout of the Mbd2 gene in mice is compatible with
viability and fertility, but some phenotypic effects were
noticed (Hendrich et al. 2001). Mbd2–/– mothers failed to
fully nurture their young, which consequently were under-
weight regardless of genotype. This weak maternal re-
sponse to pups was also evident from the delayed re-
trieval of pups after they were removed from the nests of
Mbd2−/− mothers. Other phenotypes were initially less
obvious, but led to identification of the first target genes
for a methyl-CpG binding protein. First, cells derived
from Mbd2−/− mice were found to be unable to repress
methylated reporter genes effectively. This finding
matched early experiments using in vitro extracts, which
showed that the MeCP1 complex—whose DNA-binding
component is Mbd2—is largely responsible for the re-
pression of methylated reporter genes. This evidence that
Mbd2-deficiency might compromise repression of methylated genes raised the possibility
that a careful search might uncover misregulation of en-
dogenous genes in these mice. Success came with an
analysis of T cell differentiation in the mouse (Hutchins
et al. 2002). Sorting of Mbd2−/− cells by fluorescence ac-
nabled detection of a small fraction of naive and differen-
tiated T cells in which expression of both interleukin 4
(Il4) and interferon gamma (Ifg) genes was derepressed.
Expression of Il4 and Ifg is normally mutually exclusive and is initiated by triggering differentiation into either the Th2 or the Th1 pathway, respectively. In the absence of
Mbd2, differentiated Mbd2−/− thymocytes often ex-
pressed both genes and this misexpression effect could be enhanced by addition of the histone deacetylase inhibitor
trichostatin A. Chromatin immunoprecipitation showed
that Mbd2 was normally present at the Il4 promoter and
was lost when the gene was transcriptionally activated. It
is noteworthy that absence of Mbd2 does not lead to
100% reactivation of the inappropriate gene; the effect is
instead stochastic: in that only a fraction of all Th1 cells
derepress Ifg, but derepression in those cells appears to
be complete. Thus the removal of Mbd2 appears to increase
the normally very low probability that these silent genes
will become fully reactivated. Similarly, reactivation of the aberrantly silenced pi-class glutathione S-transferase
gene can be induced in cancer cell lines by Mbd2 defi-
ciency (Lin and Nelson 2003). It seems likely that care-
ful investigation of specific cell types will uncover fur-
nier examples of the loosening of gene repression caused
by absence of Mbd2.

The status of the Mecp2-null mouse as a model for Rett
Syndrome has stimulated the search for MeCP2 target
genes. Initial attempts to detect misregulated genes in the
brains of the mutant animals offered partial success (Tu-
dor et al. 2002). Deregression of gene expression was ap-
parent, but the degree of misexpression was subtle and
only acquired statistical significance when groups of af-
fected genes were considered together. The “candidate
gene” approach proved more successful. Two groups
(Chen et al. 2003; Martinowich et al. 2003) hypothesized
that MeCP2 might play a role in activity-dependent ex-
pression of the gene for brain-derived neurotrophic factor
(Bdnf). Only one of several Bdnf promoters responds to
calcium-dependent activation and this promoter was shown
to associate with MeCP2 in cultured neurons. Upon stimulation of the neurons, MeCP2 became phos-
phorylated and was displaced from this promoter (Chen
et al. 2003). Although the effect of MeCP2 deficiency on
expression proved to be small in this in vitro system, it is
possible that deregulation has important consequences
in the animal itself. That MeCP2 deficiency can cause ma-
jor gene expression changes was shown in a nonmam-
alian system, the embryo of the amphibian, Xenopus
laevis. Here, MeCP2 was found associated with the pro-
moter of the hairy-2e gene, whose product is important
for limiting the number of embryonic cells that become
neurons (Stancheva et al. 2003). Activation of the gene
during development was accompanied by displacement of
MeCP2, suggesting that, as in the case of the mouse
Bdnf gene, MeCP2 is involved in the dynamic control of
gene expression. These findings suggest that the view
of methyl-CpG-binding proteins as long-term silenters of
gene expression may need to be modified. It remains pos-
sible, however, that they may be involved in both stable
and dynamic repression, depending on the locus con-
cerned.

A target for Kaiso in cancer cell lines has clearly been
down shown to depend on recruitment of this protein plus the N-
CoR corepressor (Yoon et al. 2003). The expression of such
genes are, at the time of writing, yet to be reported. Im-
munoprecipitation experiments have identified multiple
DNA sequences associated with various MBD proteins
(Ballestar et al. 2003), but the regulatory consequences of
MBD protein withdrawal at these loci have not yet been
determined. It is possible that all genomic regions that asso-
ciate with one of these proteins are to some extent regu-
lated by them. Alternatively, MBD proteins may spend
time unproductively at methyl-CpG sites where their reg-
ulatory input is superfluous. In depth study of specific
genes is required to distinguish these possibilities. A re-
lated issue concerns the redundancy or otherwise of these
proteins. Do they compete for access to all methylated
sites and therefore functionally back each other up? Or
does each protein act at a subset of genes that are adapted
to respond specifically to its particular regulatory influ-
ence? The answer to these questions is not yet clear, al-
though we know that combining Mecp2 and Mbd2 muta-
tions does not lead to an obvious “synthetic” enhancement of
phenotype (Guy et al. 2001). Chromatin immunopre-
METHYL-CpG-BINDING PROTEINS AND DISEASE

Dramatic demonstration of the medical relevance of methyl-CpG-binding proteins came with the discovery that at least 80% of patients with Rett Syndrome have new mutations in the MECP2 gene (Amir et al. 1999; Shahbazian and Zoghbi 2002; Kriaucionis and Bird 2003). Rett Syndrome affects females that are heterozygous for the mutations. Because of X chromosome inactivation, the patients are mosaic for expression of either the mutant or the wild-type gene. They develop apparently normally for 6–18 months, at which time they show regression of motor skills, repetitive hand movements, abnormal breathing, microcephaly, and other symptoms. Given the role of MeCP2 as a transcriptional repressor, an obvious hypothesis to explain the disease is that genes in the brain that should be silenced by MeCP2 escape repression in its absence, leading to aberrant neuronal function. As mentioned above, the discovery that MeCP2 is involved in silencing the Bdnf gene, which encodes a neuronal growth factor, is compatible with this theory, although much remains to be done to connect the medical condition with misexpression of this particular gene. Meanwhile, attempts to define the molecular pathology of Rett Syndrome continue with the search for additional MeCP2 target genes.

The relationship between DNA methylation and cancer is the subject of much research, as well as animated debate (Haylin and Bestor 2002). At the experimental level, the connection is well illustrated by the demonstration that the Min mouse model of intestinal tumorigenesis depends on the DNA methyltransferase Dnmt1 (Laird et al. 1995; Eads et al. 2002). Reduced levels of this enzyme are accompanied by a significant drop in tumorigenesis, virtually abolishing the tumor-susceptible phenotype at presentation indirectly via a methyl-CpG binding protein. As a result, this raises the possibility that misregulation of downstream genes is somehow involved. A significant difference between Dnmt1 and Mbd2 mutant phenotypes is that the former is an embryonic lethal whereas the latter results in viable and fertile mice. Thus both the tumor and its host depend on Dnmt1 function, but only tumors depend on Mbd2. This distinction has therapeutic implications that deserve to be pursued.

CONCLUSIONS

DNA methylation can be considered together with histone modifications as a mechanism for adapting chromatin structure to local functional needs. Like histone modifications, DNA methylation can be read by binding proteins. Methyl-CpG-binding proteins were initially characterized in model reporter gene systems, but gene knockouts in mice have begun to functionally relate each protein to specific target genes. In doing so, the mouse studies have illuminated the relationships between methyl-CpG-binding proteins and human disease.

ACKNOWLEDGMENTS

Our research is supported by the Wellcome Trust, United Kingdom.

REFERENCES


Daniel J.M., Spring C.M., Crawford H.C., Reynolds A.B., and Baig A. 2002. The p120(crt)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. Nuclease Acids Res. 30: 2931.

Lachner M., O'Carroll D., Rea S., Mechtler K., and Jenuwein T. 2001. Closely related proteins Mbd2 and Mbd3 play distinct but interacting roles in mouse development. Genes Dev. 15: 23:

Reading the DNA Methylation Signal

A. BIRD and D. MACLEOD

Access the most recent version at doi:10.1101/sqb.2004.69.113

References
This article cites 54 articles, 18 of which can be accessed free at:
http://symposium.cshlp.org/content/69/113.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to Cold Spring Harbor Symposia on Quantitative Biology go to:
http://symposium.cshlp.org/subscriptions