PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

DNA Methylation and Late Replication Probably Aid Cell Memory, and Type 1 DNA Reeling Could Aid Chromosome Folding and Enhancer Function

A. D. Riggs

Phil. Trans. R. Soc. Lond. B 1990 **326**, 285-297
doi: 10.1098/rstb.1990.0012

Phil. Trans. R. Soc. Lond. B **326**, 285–297 (1990)
Printed in Great Britain

DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function

By A. D. Riggs

Biology Department, Beckman Research Institute of the City of Hope, City of Hope Medical Center, Duarte, California 91010, *U.S.A.*

DNA methylation in mammals is reviewed, and it is concluded that one role of methylation is to aid cell memory, which is defined as the ability of mitotically derived progeny cells to remember and re-establish their proper cellular identity. Methylation of X-linked CpG-rich islands probably stabilizes X-chromosome inactivation, but other mechanisms appear to be involved. Late replication is discussed as a key ancestral mechanism for X inactivation, and it is emphasized that early and late replication domains may each be self perpetuating. Therefore, early–late replication timing becomes another strong candidate mechanism for cell memory. A chromosome-loop folding enigma is discussed, and it is concluded that special mechanisms are needed to explain the formation and maintenance of specific looped domains. DNA reeling, such as done by type I restriction–modification enzymes, is proposed to provide this special mechanism for folding. DNA reeling mechanisms can help to explain the *cis*-spreading of X-chromosome inactivation as well as long-range action by enhancers.

1. Introduction

Modification of DNA by post-replicational methylation probably arose early in evolutionary history because it is widespread throughout all orders of the plant and animal kingdoms (Adams & Burdon 1985). However, it is not part of the primary genetic code and, therefore, is available for special purposes. Among the many roles suggested for DNA methylation in prokaryotes and eukaryotes have been restriction–modification systems, strand-selection in mismatch repair, and control of gene function (Razin *et al.* 1984; Adams & Burdon 1985). Evidence for each of these roles can be found in various organisms, but the focus of this paper will be on the functions of DNA methylation in higher organisms, and especially mammals. The only modified base so far detected in mammals is 5-methylcytosine and much evidence suggests that this modified base is part of a general gene silencing system involving some aspect of chromatin structure (Cedar 1988; Riggs 1989*a*).

There are several known levels of linear and spatial organization of the eukaryotic chromosome (van Holde 1988) and methylation could act at any (or all) of these levels. The first level of higher-order chromatin structure is a 'beaded string' fibre formed by wrapping the DNA duplex around histone octomers to form nucleosomes. The second level is a 30 nm fibre formed by compacting the beaded string, probably by coiling, to form a solenoid. The third level is a 200–300 nm fibre formed by folding of the 30 nm fibre into 50–100 kilobase (kb) looped domains. Each looped domain may be a unit of gene regulation and, possibly, a unit of replication (Goldman 1988). A fourth level of organization, which occurs at least in mitotic

cells, is the further compaction, by coiling, of the 300 nm fibre into the 700 nm diameter chromatid fibre seen in mitotic chromosomes (Rattner & Lin 1988). Without doubt, the foundation of chromosomal order is specific protein–DNA interactions. Enzymatic DNA methylation, the subject of this meeting can certainly affect protein–DNA interactions (Watt & Molloy (1988) and references therein) and thus seems virtually certain to influence chromatin structure.

2. Cell memory and DNA methylation

Mammals and many other higher organisms have highly differentiated cells that continue to proliferate throughout the lifespan of the adult. The proper maintenance of cell commitment and cell phenotype must require a specific pattern of chromatin structure. A complex pattern of active and inactive domains in the interphase chromosome must be remembered and faithfully reproduced at each somatic generation. This is crucially important for most higher eukaryotes because the history of a cell determines its response to differentiating signals, including signals to stop cell replication. A multitude of interacting molecules must be involved in forming the proper pattern of chromatin structure over some 10^9 base pairs (b.p.); therefore, it seems important for the differentiated chromosome structure to be 'self-templating' and to provide a memory function to aid its perpetuation. Progeny cells must not only reproduce their proper phenotype, but they must also remember their commitments. In this context, it is convenient to speak of cell memory and to ask the question: what are the mechanisms of cell memory?

Among the several models that have been proposed for cell memory (reviewed by Riggs (1989*a*)), a model for the self perpetuation of methylation patterns has become increasingly attractive as experimental evidence has accumulated. The essential feature of this proposal is called the maintenance methylase concept (Riggs 1975; Holliday & Pugh 1975), which is illustrated in figure 1*a*. The model was based on the then known properties of a type I restriction/modification enzyme called *Eco* K (reviewed by Yuan & Hamilton 1984). This *Escherichia coli* enzyme is composed of three different subunits, a specificity subunit (s), a methylase subunit (m) and a restriction subunit (r) (figure 2). The s subunit recognizes and binds to the sequence $AACN_6GTGC$. The DNA then acts as an allosteric effector and the resultant action of the enzyme complex is dependent on the modification state of the DNA. If the recognition site is methylated in both strands, the complex dissociates. If the site is hemimethylated, the unmethylated strand is rapidly methylated. If the site is unmethylated in both strands, the methylase subunit is not efficiently activated. Instead, the restriction subunit is activated and cleaves the DNA. More recently it has been learned that DNA translocation occurs before cleavage (Yuan & Hamilton 1984; Endlich & Linn 1985; Studier & Bandypodhyay 1988). This added feature will become important in later sections of this paper.

For the maintenance methylase concept, the key point is that hemimethylated sites are efficiently methylated, whereas unmethylated sites remain unmethylated. This allows the perpetuation, through DNA replication, of two methylation states. These alternate states of methylation can be thought of as being separated by an energy barrier, even though they are exposed to the same environment and could have the same nucleotide sequence. To convert from one methylation state to another requires special work, either demethylation or *de novo* methylation. In the absence of specific action, the methylation patterns should be

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

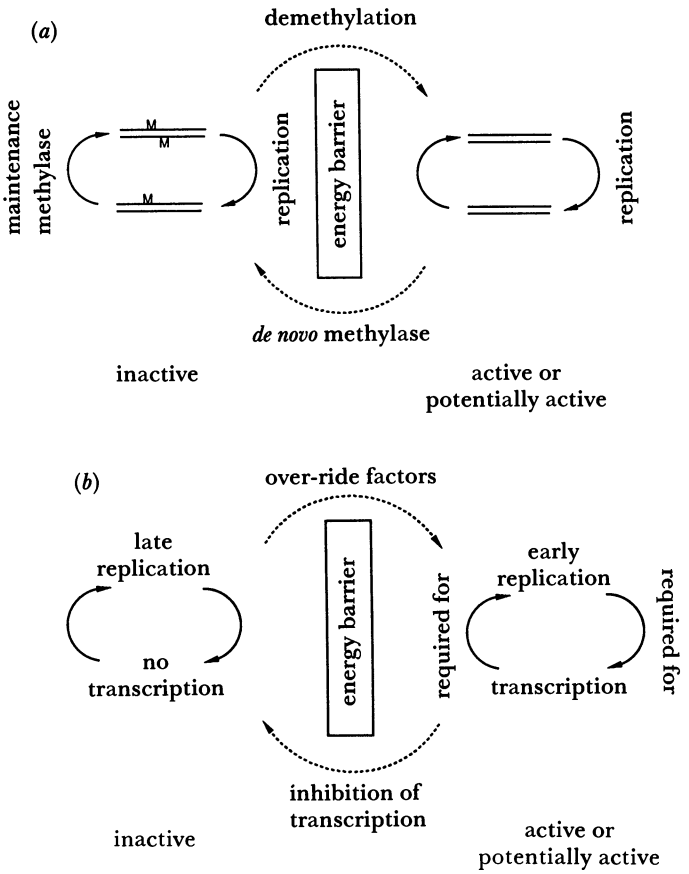

FIGURE 1. (*a*) Alternate stable states maintained by a maintenance methylase system: M represents a DNA cytosine methyl group. The alternate states survive DNA replication because the maintenance methylase acts only on hemimethylated DNA. (*b*) Alternate stable states maintained by early–late replication. Late replicating DNA may assume an inactive chromatin configuration and preclude transcription; early replication may be required for the establishment of a transcriptionally competent domain and transcription may be required for initiation of early-replication origins. Late-replication origins are proposed either not to be dependent on transcription or to be dependent on special late transcripts. This system would allow a pattern of alternate transcriptional states to be established and maintained through DNA replication. Special action by over-ride factors or inhibitory factors would be necessary to cause a change in state of a replication domain.

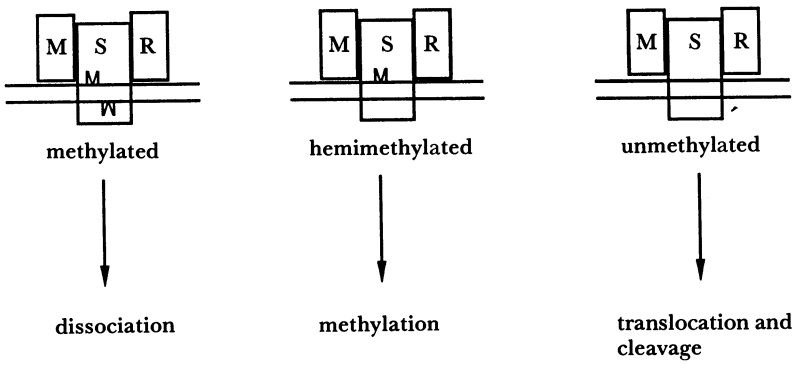

FIGURE 2. *Eco* K type I restriction/modification enzymes. The horizontal lines represent DNA strands; M on the DNA strand represents a DNA cytosine methyl group. The boxed S represents the specificity subunit of the enzyme complex; the boxed R, the restriction subunit; the boxed M, the methylase subunit. Action of the enzyme complex depends on the methylation state of the DNA site.

self-perpetuating. Figure 1*a* illustrates the coexistence of two separately stable states. It should be noted that hypomethylation of critical sites may be necessary for active transcription, but it is clearly not sufficient. Therefore, the state on the right side of figure 1*a* is labelled active or potentially active. A hypomethylated site can be considered as unlocked and poised, ready to respond to inducing signals.

Much experimental evidence now indicates that methylation patterns in mammalian DNA are indeed somatically heritable. Thus DNA methylation probably is part of a secondary, somatically inherited information coding system (Razin *et al.* 1984; Holliday 1987; Riggs 1989*a*). The key question now is: what are the functions of this epigenetic system? Possible answers that will be discussed in this paper are: (i) a primary role in determination and differentiation during development; (ii) a role in cell memory and higher order chromatin structure; (iii) a role in replication timing and (iv) a role in chromosome folding.

2. Cell memory and *Drosophila*

Mammalian cells convert about 4% of cytosine to 5-methylcytosine, but *Drosophila* has very little 5-methylcytosine, perhaps none (Achwal *et al.* 1984; Patel & Gopinathan 1987). It still remains possible that *Drosophila* has one or two modified bases per chromosome loop or replicon, but, at the very least, we know that the most fundamental aspects of gene control can proceed adequately without the relatively high 5-methylcytosine levels seen in mammals. What more can be deduced about the function of methylation in mammals if we assume that *Drosophila* has no cytosine modification at all? *Drosophila* manages to carry out a fairly complicated developmental programme so one can conclude that it is unlikely that methylation changes are always the key, initiating, step in cellular determination. This was *a priori* unlikely, however, given knowledge gained from mammalian studies. Since the work of Singer *et al.* (1978), a favoured model has been that *de novo* methylation takes place in the early mammalian embryo, with DNA reaching, shortly after implantation, a high-methylation ground state, from which the major changes in methylation patterns take place by specific demethylation events (see Monk *et al.* 1987; Sanford *et al.* 1987). Much of mammalian development is therefore from a high-methylation ground state, suggesting that critical control elements may initially be methylated. To activate such an element, an already-bound repressor must be removed, or an activator, insensitive to methylation, must bind to the methylated control region. In either case, the primary event is a change in the binding of a methylation insensitive protein, not a methylation change. One can predict that true determinator proteins (Riggs & Jones 1983) should not be sensitive to methylation. I have discussed this point previously and suggested that methylation patterns may just be a footprint of regulatory protein binding (Riggs 1989*a*). However, this is not to belittle the importance of these footprints, which have altered potential for protein–DNA interactions, and probably are self-templating and thus stably heritable in somatic cells.

The greatest difference between the developmental systems of *Drosophila* and man is in the number of mitoses required for forming and maintaining the organism. Only about 20 mitoses are needed to provide the approximately 10^6 somatic cells in an adult *Drosophila* and there is no DNA replication in the adult, except in germ cells (Raman & Lakhotia 1989). In man, in the order of 10^6 mitoses per second occur, leading to an estimate of 10^{16} mitoses in the course of a lifetime (Alberts *et al.* 1983). Not only are the numbers astoundingly different, but

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

out-of-control dividing cells are dangerous for long-lived organisms and must be guarded against. The fidelity of information transfer to progeny cells is thus much more important for mammals and other organisms having continued cell division in the adult. This reasoning leads to the suggestion that the primary function of methylation is to aid the mistake-free transfer of chromatin structural information to progeny cells; that is, methylation probably aids cell memory. Information about *Drosophila* is therefore consistent with methylation serving a locking function, ensuring the stable memory of a differentiated state, as suggested by Razin & Riggs (1980).

3. X-chromosome inactivation and cell memory

A prime example of cell memory is provided by the phenomenon of X-chromosome inactivation. Female mammalian embryos begin development with two active X chromosomes, but at about the time of implantation, almost all genes on one of the two X chromosomes become transcriptionally silent and the inactive chromosome becomes relatively condensed and late replicating (for recent reviews see Lyon (1988); Grant & Chapman (1988); Riggs (1989*b*)). The initial choice between inactivation of the maternal or the paternal X is random, when averaged over many cells in the embryo proper. However, once an inactive X chromosome forms in a given cell, it becomes a somatically heritable entity, with the same X remaining inactive in all progeny cells. Female mammals are cellular mosaics, composed of cells functionally hemizygous for X-linked genes such as *hprt*. The reactivation frequency of the inactive X chromosome (usually measured by the reactivation of *hprt*) is less than 10^{-8} (not detectable) in normal, diploid human cells. The reactivation frequency in mouse–human hybrid cells is detectable, but still low (10^{-6}).

For three reasons, it is extremely likely that methylation is part of the maintenance mechanism for X-inactivation (reviewed by Monk (1986); Grant & Chapman (1988); Riggs (1989*b*)). First, DNA from the inactive chromosome transfects less efficiently than DNA from an active chromosome, indicating that the DNA is modified in some way. Second, inhibitors of DNA cytosine methylation, such as 5-azacytidine, cause efficient reactivation of inactive X-linked genes. Third, DNA reactivated by 5-azacytidine treatment now functions normally in transfection experiments. The overall level of methylation is the same on DNA from the active and inactive X chromosome, but, as will be discussed below, all data are consistent with methylation differences at critical sites. One type of critical site may be CpG-rich islands.

Methylation of CpG-rich islands

An important feature of the mammalian genome is the presence of CpG-rich islands spaced at approximately 100 kilobases (kb) intervals. These islands are characteristically unmethylated and are often near transcriptionally active genes (Bird 1986). In contrast to most CpG islands, several found at the 5′ regions of X-linked genes are highly methylated on the inactive X chromosome of placental mammals (Toniolo *et al.* 1988, and references therein). As an example, we have found that seven *Hpa* II sites (CCGG) over an 800 b.p. region at the 5′ end of X-linked human *pgk* are completely unmethylated on the active X chromosome, but completely methylated on the inactive X chromosome (Keith *et al.* 1986). Hansen *et al.* (1988) have studied methylation of the human *pgk* 5′ region in mouse–human hybrids selected only for the reactivation of human *hprt*. Some reactivants express both human *hprt* and human

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

X-linked *pgk*, but most do not. In their study a perfect correlation was found between inactivity of *pgk* and methylation of certain sites just downstream of the transcription start site, at the 3′ end of the CpG-rich island. We have recently studied the corresponding promoter and CpG island of *mouse* X-linked *pgk* and have found that whereas the upstream region of the *pgk* island is only partially methylated on the inactive X chromosome, the region just downstream from the transcription start site seems fully methylated (J. Singer-Sam *et al.* 1989). These studies, and similar correlations between methylation of CpG islands near other X-linked genes, indicate that the CpG islands may be critical sites for the maintenance of X-chromosome inactivation.

Marsupials and X-linked CpG islands

Recent studies on promoter-associated CpG islands near marsupial X-linked genes (Kaslow & Migeon 1987) are consistent with methylation helping to preserve the inactive state of X-linked genes. Marsupials, which appear earlier in the fossil record, are thought to represent the evolutionary precursors of placental mammals. Marsupials preferentially inactivate the paternal X chromosome, similar to the X-inactivation seen in the extra-embryonic tissues of placental mammals, so it is of importance that Kaslow & Migeon (1987) have found no sex-related differences in the methylation of CpG islands associated with three X-linked genes. However, the inactivation of paternal genes is only partial in most marsupial tissues, and readily detectable reactivation of the paternal *hprt* gene occurs in tissue culture (Migeon *et al.* 1988). Therefore, these results suggest that methylation of CpG islands, as seen in placental mammals, increases the stability of the inactive state; that is, methylation probably enhances the fidelity of cell memory. Studies of marsupials, however, suggest that methylation of CpG islands is not the whole story.

4. Early–late replication and cell memory

The above discussion indicates that other control systems in addition to methylation of CpG-rich islands are likely to be part of X-chromosome inactivation. Marsupials and monotremes (egg-laying mammals) share with placental mammals the late-replication feature of the inactive X chromosome (McKay *et al.* 1987) and the onset of asynchronous replication is closely correlated with the onset of condensed sex chromatin and transcriptional inactivation (VandeBerg *et al.* 1987). This correlation, and the recent results showing lack of CpG island methylation in a marsupial, led McKay *et al.* (1987), Grant & Chapman (1988) and Riggs (1989*b*) to suggest that late replication may be the key ancestral mechanism used for the establishment of X-chromosome inactivation.

Recent results in two areas of research make late replication an especially attractive candidate for gene silencing. First, evidence is accumulating that late replication (replication in the second half of S phase) precludes transcription and that early replication is a necessary but not sufficient condition for transcriptional activity (Holmquist 1987; Goldman 1988). Second, considerable evidence now indicates that transcription (or at least active transcriptional elements) is required for the initiation of some replicons (DePamphilis 1988). Grant & Chapman (1988) and Riggs (1989*b*) have proposed that combination of these two concepts leads to self-perpetuating states (figure 1*b*) analogous to the maintenance methylase scheme shown in figure 1*a*. If early replication is required for transcription, and transcription

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

is required for early replication, then a stable, active state is maintained through DNA replication and cell division. If late replication precludes active transcription, and location in an early replicating domain is required for transcription, then a stable, inactive state is maintained. Conversion between the two states requires special intervention, either by override or inhibitory factors, as indicated in figure 1*b*.

A key prediction of this scheme is that late-replication origins are different from early replication origins. Either late origins do not require transcription, or special late transcripts are used. Also, late-replication origins could be either different sequences, or they could be differently modified. Are late origins methylated? This now seems to be an important question, especially as: (i) methylation inhibitors are known to advance the replication timing of rodent X chromosome and satellite DNA (Jablonka *et al.* 1985; Selig *et al.* 1988) and (ii) cytosine methylation may affect choice of replication origins in slime moulds (Cooney *et al.* 1988). It also is of interest that adenine methylation may control replication timing in *E. coli* (Ogden *et al.* 1988). N-6-methyladenine has not been detected in eukaryotic DNA; however, the presence of this modified base in replication origins spaced every 50–100 kb would have gone undetected.

Because replication origins may be important for X-chromosome inactivation, a model consistent with all present data is that the methylation of replication origins controls replication timing, which in turn influences higher-order chromatin structure and transcription. Methylation of cell type-specific enhancers, CpG islands and critical sites near promoters could be additional locking features, further ensuring high-fidelity cell memory.

5. Type I DNA reeling and chromosome folding

The maintenance methylase concept was based on the properties known in 1973 for *E. coli* type I restriction–modification enzymes. It is therefore intriguing that DNA translocation, a more recently discovered property of type I enzymes, may help explain the specificity of chromosome folding and some other phenomena that have been especially difficult to explain at the molecular level. One such molecular enigma is the *cis*-spreading of inactivation to include most genes on one entire X chromosome. It is generally believed that the X-inactivation process begins at a master locus, *Xce*, the inactivation centre, and then spreads along the chromosome. Only one of the two X chromosomes is affected, so X-inactivation provides an example of *cis*-action over 10^8 b.p.!

A folding enigma

Looped domains are probably an essential part of chromosome organization. How are they formed and maintained? The DNA loop is thought to be attached at the base to a protein scaffold or matrix (Cockerill & Garrard 1986; Gasser & Laemmli 1987). The DNA elements that bind to form loops have been called SARs (scaffold associated regions) for metaphase chromosomes or MARs (matrix associated regions) for interphase chromosomes. SARs and MARs may be the same, but this is not an esssential point. For convenience, SAR will often be used as a general term for both of these elements. Most important for this discussion are questions of specificity. Do the SAR–protein complexes at the bases of loops interact only with their *cis*-located neighbours (about 100 kb distant), or is there limited specificity and thus promiscuous *intrastrand* intermixing of folding complexes (see figure 3)? Also, what prevents *interstrand* mixing of homologous SARs or folding complexes?

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

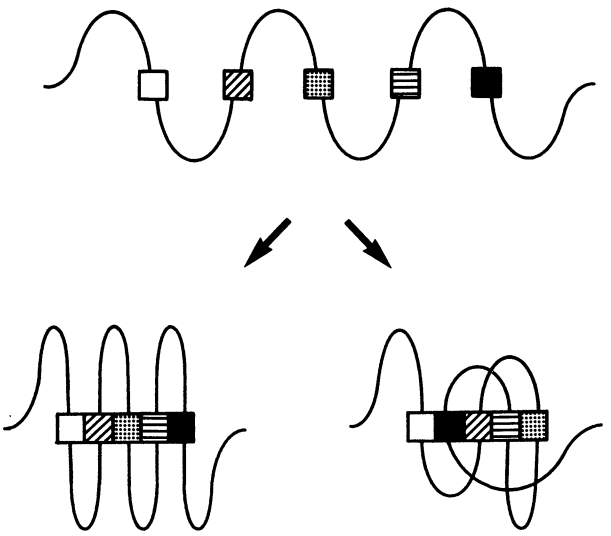

FIGURE 3. The SAR or MAR intermixing enigma. At the top is shown a DNA fibre with various SAR (or MAR)–protein complexes indicated (boxes). If each SAR complex binds only to its immediately neighbouring SAR, then the precisely folded structure depicted on the bottom left is formed. If each SAR can indiscriminately bind to all SARS, a tangled structure would result, as illustrated, bottom right.

A chaotic tangle seems to be inevitable unless: (i) a special mechanism exists, (ii) each SAR is different and specificity is high, or (iii) the folding pattern is templated by a prior folding pattern. This latter possibility implies that folding structure is never lost. Folding is probably retained through interphase, but some reformation of lost structure seems necessary after DNA replication or accidents. Over time and replication cycles, order would be lost unless a specific maintenance process repairs perturbations in folded structure. Can a complicated folding pattern survive 10^{16} mitoses without a folding-repair mechanism? Random diffusion as a repair mechanism for relatively large structures separated by 100 kb intuitively seems inadequate. More likely a special mechanism exists to aid specific folding. It will be seen that the properties of type I enzymes provide such a special mechanism.

Type I DNA reeling

For several years, type I restriction–modification enzymes, such as *Eco* K, have been known to translocate along DNA and cut at sites far removed from the original binding site; moreover, it seems that the enzyme remains attached to the original binding site (Yuan *et al.* 1980; Yuan & Hamilton 1984; Endlich & Linn 1985). Studier & Bandyopadhyay (1988) have suggested the mechanism illustrated in figure 4. After binding to an unmethylated recognition site, the type I enzyme remains attached to the recognition site and begins translocating DNA towards itself from each side; that is, it begins a bidirectional reeling-in of DNA. The DNA reeling requires ATP and can proceed for at least 20 kb at an average speed of about 200 b.p. per second. When the reeling process brings two neighbouring enzymes into contact a new activity is triggered and the DNA between them is cut. For chromosome structure, and possibly for enhancer function, two points deserve emphasis. First, the type I DNA reeling process forms specific loops, bringing only *cis*-located sites together. Second, when neighbouring DNA–protein complexes are reeled into close contact, a conformational change is triggered.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

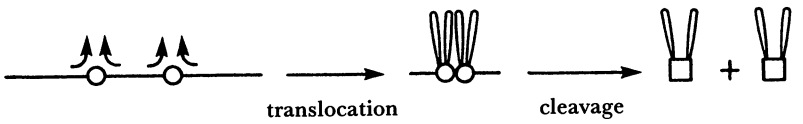

FIGURE 4. DNA reeling and cleavage by *Eco* K type I restriction–modification enzymes. The enzyme binds to specific sites and then begins reeling in DNA from both sides. When two enzyme molecules are brought together by the reeling process, a conformational change activates nuclease activity and the DNA is cut between the two enzymes. (Adapted from Studier & Bandypodhyay (1988).)

A chromosome folding model

Figure 5 illustrates the type I DNA-reeling concept applied to a eukaryotic chromosome with folding elements spaced at 50–100 kb intervals. Proteins evolutionary related to bacterial type I enzymes, but having no restriction activity, are proposed to bind to, and remain positioned at, specific DNA sites (folding elements). DNA-reeling then proceeds. The reeling process causes the chromosome to fold specifically and a 'scaffold' of contiguous proteins forms at the base of the loops. If proteins at the base of the 1000 loops in a chromosome were brought into linear contact, a protein backbone of about 5 μm would result. This is approximately the size of mammalian mitotic chromosomes.

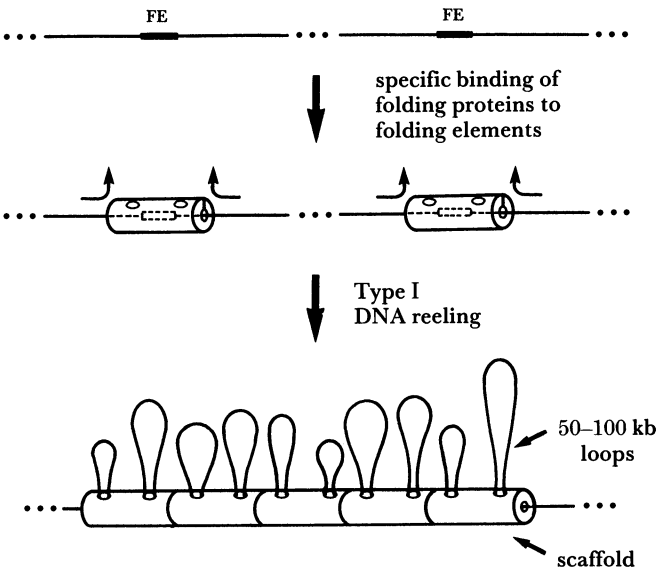

FIGURE 5. Chromosome folding by type I DNA reeling. Folding elements (FE) are proposed to be located every 50–100 kb along the DNA fibre. Folding proteins bind to these FE sites, forming a folding complex. The folding complex reels in DNA bidirectionally until a scaffold of contiguous folding complexes is formed with protruding 50–100 kb loops.

There are several appealing aspects of type I DNA-reeling mechanisms for chromosome organization.

1. The SAR intermixing problem discussed earlier is eliminated. A given folding element and its associated proteins is pulled into contact only with its flanking neighbours. Only a few different kinds of folding elements and associated folding proteins could bring precise order

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

from tangled chaos. Consider an extreme example: perhaps there are only two families of folding elements and folding proteins, one for G band DNA and one for R band DNA. This could result in the banding seen in mitotic chromosomes. The main point here is that only a relatively small family of different proteins is needed, because the final structure is dictated by the linear placement of DNA elements, rather than by non-DNA supramolecular structures or careful control of protein–protein specificity or protein levels. Within each family, the location of folding elements relative to other elements (e.g. promoters, enhancers, replication origins) would influence gene function.

2. Type I reeling provides an ATP-driven repair mechanism for faithfully reforming the proper folded structure after any perturbation (see below and figure 6).

3. The mechanism is evolutionarily straightforward, with precedent for the basic mechanisms.

4. Folding is an exclusively *cis* process. Random diffusion is not part of the mechanism except for the initial formation of the folding complex.

5. A new model for the spreading of X-chromosome inactivation seems feasible at the molecular level. It has been proposed (Riggs 1989*b*) that a process of SAR modification begins at the *Xce* and spreads along tightly juxtaposed loop-fastening proteins of the scaffold. With the DNA reeling concept, spreading is only *cis* because *trans* complexes cannot be in tight juxtaposition.

During replication, DNA may be reeled through the replication complex (see Zehnbauer & Vogelstein 1985). These experiments provide precedence in mammalian cells for a reeling process, but type I DNA reeling seems much more appealing as a mechanism to restore order after replication, and to maintain order in the interphase nucleus, especially in non-dividing cells where appropriately organized chromatin must be maintained for many years.

Figure 6 illustrates repair in a chromatid that has had a 'loop-fastening' complex broken by stretching forces, or, alternatively, figure 6 could represent chromatid reorganization after homologous recombination in the loop and the beginning of daughter chromatid separation during metaphase. If folding complexes become active for type I DNA reeling whenever they are not in close contact with another complex, a loop-repair process automatically ensues. An analogy with catching a large fish on a small fishing line is illustrative. The best strategy is to

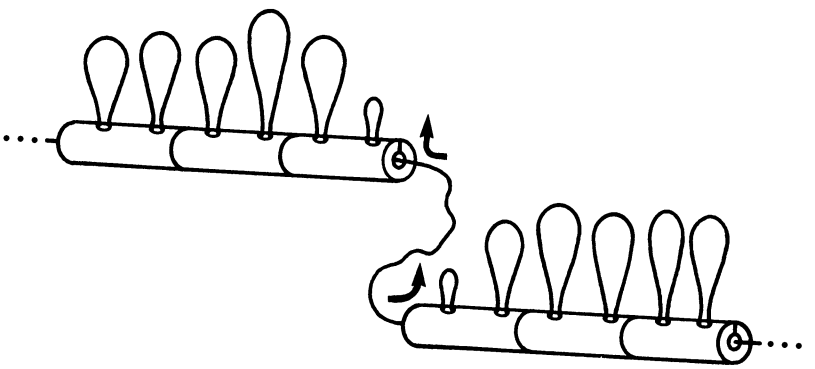

FIGURE 6. Maintenance of loop and scaffold by type I DNA reeling. Shown is a chromatid fibre after some event has pulled apart two adjacent folding complexes by breaking protein–protein interactions at the base of a loop. It is proposed that DNA reeling would begin after separation, reforming the loop and repairing the continuity of the scaffold.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

continuously apply a weak unidirectional force by reeling. If the drag on the fishing reel is loose, the fish eventually is brought in without breaking the line.

The model shown in figure 5 predicts that at least some components of the folding complex become part of a scaffold structure at the base of the loops. The metaphase scaffold contains two major proteins, a 170 kDa protein and a 135 kDa protein. The 170 kDa protein has been identified as topoisomerase II (Gasser & Laemmli 1987). The nuclear matrix also has topoisomerase II as a major component. It is conceivable that topoisomerase II is a 'loop-fastener' (Gasser & Laemmli 1987) and will be found to have reeling activity, but it seems more likely that reeling would be the function of another, as yet less well characterized, scaffold- or matrix-associated protein. Mammalian DNA methylase should be checked for ATP hydrolysis and type I reeling, because some DNA methylase activity is associated with the nuclear matrix (Burdon *et al.* 1985) and DNA methylase has been proposed to 'walk' along DNA (Drahovsky & Morris 1971). Also, at least one report indicates that the methylated state of a MAR is more strongly associated with the nuclear matrix than is the unmethylated state (Ellis *et al.* 1988).

Enhancers

Enhancers are control elements that can act at a distance of many kilobases. Folding, to bring distant sites into contact, has emerged as a favoured mechanism to help explain enhancer action (Ptashne 1988). Obviously, type I DNA reeling could increase enhancer efficiency by causing folding and bringing the various other elements of the enhancer into close proximity with the target promoter. Some SARs involved in general chromosome folding could provide this function, and it is intriguing that enhancers have been found associated with SARs and MARs (Gasser & Laemmli 1987; Cockerill & Garrard 1986). But it is also possible that a different set of folding proteins is used to aid enhancer function. These proteins would retain the bidirectional reeling function, but would stop upon contact with a promoter complex.

6. Concluding comments

The models discussed here for early–late replication, chromosome folding and enhancer function could each involve methylation and each suggests experiments. It is propitious that experimental techniques for the identification, cloning, and study of mammalian replication origins are now being perfected (reviewed by Umek *et al.* (1989)). It should soon be possible to identify replication origins and determine, for example, whether early and late origins are different sequences or just the same sequence differently methylated. Techniques for identifying mammalian SARs and MARs are available. The methods used to study *E. coli* type I restriction enzymes can now be applied to the study of SARs and SAR-associated proteins. Constructs with two SARs or MARs on the same plasmid should be informative. Clearly the methylation of SARs and MARs needs to be studied. Are X-linked SARs differentially methylated between the active and inactive X chromosome?

It seems very likely that 5-methylation plays a role in cell memory by stabilizing the inactive chromatin state. Although the details of this role remain to be learned, it can be hoped that studies stimulated by the models discussed here may provide new information.

I thank Gerald Holmquist for encouragement, helpful criticism, and many stimulating discussions.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References

Achwal, C. W., Ganguly, P. & Chandra, H. S. 1984 Estimation of the amount of 5-methylcytosine in *Drosophila melanogaster* DNA by amplified ELISA and photoacoustic spectroscopy. *EMBO J.* **3**, 263–266.

Adams, R. L. P. & Burdon, R. H. 1985 *Molecular biology of DNA methylation.* New York: Springer–Verlag.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. 1983 *Molecular biology of the cell*, p. 940. New York: Garland Publishing.

Bird, A. P. 1986 CpG-rich islands and the function of DNA methylation. *Nature, Lond.* **321**, 209–213.

Burdon, R. H., Qureshi, M. & Adams, R. L. P. 1985 *Biochem. biophys. Acta* **825**, 70–79.

Cedar, H. 1988 DNA methylation and gene activity. *Cell* **53**, 3–4.

Cockerill, P. N. & Garrard, W. T. 1986 Chromosomal loop anchorage of the κ-immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* **44**, 273–282.

Cooney, C. A., Eykholt, R. L. & Bradbury, E. M. 1980 Methylation is coordinated on the putative replication origins of Physarum rDNA. *J. molec. Biol.* **204**, 889–901.

DePamphilis, M. L. 1988 Transcriptional elements as components of eukaryotic origins of DNA replication. *Cell* **52**, 635–638.

Drahovsky D. & Morris, N. R. 1971 Mechanism of action of rat liver DNA methylase: I. Interaction with double-stranded methyl-acceptor DNA. *J. molec. Biol.* **57**, 475–89.

Ellis, G. C., Grobler-Rabie, A. F., Hough, F. S. & Bester, A. J. 1988 Location and methylation pattern of a nuclear matrix-associated region in the human pro-alpha2(I) collagen gene. *Biochem. biophys. Res. Commun.* **157**, 500–506.

Endlich, B. & Linn, S. 1985 The DNA restriction endonuclease of *E. coli* B: Studies of the DNA translocation and the ATPase activities. *J. biol. Chem.* **260**, 5720–5728.

Gasser, S. M. & Laemmli, U. K. 1987 A glimpse at chromosomal order. *Trends Genet.* **3**, 16–22.

Goldman, M. A. 1988 The chromatin domain as a unit of gene regulation. *BioEssays* **9**, 50–55.

Grant, S. G. & Chapman, V. M. 1988 Mechanisms of X-chromosome inactivation. *A. Rev. Genet.* **22**, 199–233.

Hansen R. S., Ellis, N. A. & Gartler, S. M. 1988 Demethylation of specific sites in the 5′ region of the inactive X-linked human phosphoglycerate kinase gene correlates with the appearance of nuclease sensitivity and gene expression. *Molec. Cell Biol.* **8**, 4692–4699.

Holliday, R. 1987 The inheritance of epigenetic defects. *Science, Wash.* **238**, 163–170.

Holliday, R. & Pugh, J. E. 1975 DNA modification mechanisms and gene activity during development. *Science, Wash.* **187**, 226–232.

Holmquist, G. P. 1987 The role of replication time in the control of tissue-specific gene expression. *Am. J. hum. Genet.* **40**, 151–173.

Jablonka, E., Goitein, R., Marcus, M. & Cedar, H. 1985 DNA hypomethylation causes an increase in DNase-I sensitivity and an advance in the time of replication of the entire inactive X chromosome. *Chromosoma* **93**, 152–156.

Kaslow, D. C. & Migeon, B. R. 1987 DNA methylation stabilizes X-chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6210–6214.

Keith, D. K., Singer-Sam, J. & Riggs, A. D. 1986 Active X-chromosome DNA is unmethylated at eight CCGG sites clustered in a guanine-plus-cytosine-rich island at the 5′ end of the gene for phosphoglycerate kinase. *Molec. Cell Biol.* **6**, 4122–4125.

Lyon, M. F. 1988 The William Allan Memorial Award address: X-chromosome inactivation and the location and expression of X-linked genes. *Am. J. hum. Genet.* **42**, 8–16.

McKay, L. M., Wrigley, J. M. & Graves, J. A. M. 1987 Evolution of mammalian X-chromosome inactivation: sex chromatin in monotremes and marsupials. *Aust. J. biol. Sci.* **40**, 397–404.

Migeon, B. R., Axelman, J. & Beggs, A. H. 1988 Effect of ageing on reactivation of the human X-linked HPRT locus. *Nature, Lond.* **335**, 93–96.

Monk, M. 1986 Methylation and the X chromosome. *BioEssays* **4**, 204–208.

Monk, M., Boubelik, M. & Lehnert, S. 1987 Temporal and regional changes in DNA methylation in the embryonic, extraembryonic, and germ-cell lineages during mouse embryo development. *Development* **99**, 371–382.

Ogden, G. B., Pratt, M. J. & Schaechter, M. 1988 The replication origin of the *E. coli* chromosome binds to cell membranes only when hemimethylated. *Cell* **54**, 127–135.

Patel, C. V. & Gopinathan, K. P. 1987 *Analyt. Biochem.* **164**, 164–169.

Ptashne, M. 1988 How eukaryotic transcriptional activators work. *Nature, Lond.* **335**, 683–689.

Raman, R. & Lakhotia, S. C. 1989 Comparative aspects of chromosome replication in *Drosophila* and mammals. In *Trends in chromosome research* (ed. T. Sharma) New Delhi: Narosa–Springer–Verlag. (In the press.)

Rattner, J. B. & Lin, C. C. 1988 Radial loops and helical coils coexist in metaphase chromosomes. *Cell* **42**, 291–296.

Razin, A. & Riggs, A. D. 1980 DNA methylation and gene function. *Science, Wash.* **210**, 604–610.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Razin, A., Cedar, H. & Riggs, A. D. (ed.) 1984 *DNA methylation: biochemistry and biological significance*. New York: Springer–Verlag.

Riggs, A. D. 1975 X chromosome inactivation, differentiation and DNA methylation. *Cytogenet. Cell Genet.* **14**, 9–25.

Riggs, A. D. 1989*a* DNA methylation and cell memory. *Cell Biophys.* **15**, 1–13.

Riggs, A. D. 1989*b* Marsupials and mechanisms of X chromosome inactivation. *Aust. J. Zool.* **37**. (In the press.)

Riggs, A. D. & Jones, P. A. 1983 5-Methylcytosine, gene regulation, and cancer. *Adv. Cancer Res.* **40**, 1–30.

Sanford, J. P., Clark, H. J., Chapman, V. M. & Rossant, J. 1987 Differences in DNA methylation during oogenesis and spermatogenesis, and their persistence during early embryogenesis in the mouse. *Genes Devel.* **1**, 1039–1046.

Selig, S., Ariel, M., Goitein, R., Marcus, M. & Cedar, H. 1988 Regulation of mouse satellite DNA replication time. *EMBO J.* **7**, 419–426.

Singer, J., Roberts-Ems, J., Luthardt, R. W. & Riggs, A. D. 1978 DNA methylation levels in the early mouse embryo. *Nucl. Acids. Res.* **7**, 2369–2385.

Singer-Sam, J., Yang, T. P., Mori, N. Tanguay, R. L., Le Bon, J. M., Flores, J. C. & Riggs, A. D. 1989 DNA methylation in the 5′ region of the mouse *PGK*-1 gene and a quantitative PCR assay for methylation. In *UCLA symposia on molecular and cellular biology, new series*, vol. 128, Nucleic acid methylation. (ed. G. Clawson, D. Willis, A. Weissbach & P. Jones). New York: Alan R. Liss.

Studier, F. W. & Bandyopadhyay, P. K. 1988 Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4677–4681.

Toniolo, D., Martini, G., Migeon, B. R. & Dono, R. 1988 Expression of the G6PD locus on the human X chromosome is associated with demethylation of three CpG islands within 100 kb of DNA. *EMBO J.* **7**, 401–406.

Umek, R. M., Linskens, M. H. K., Kowalski, D. & Huberman, J. A. 1989 New beginnings in studies of eukaryotic DNA replication origins. *Biochim. biophys. Acta* **1007**, 1–14.

van Holde, K. E. 1988 *Chromatin*. New York: Springer–Verlag.

VandeBerg, J. L., Robinson, E. S., Samollow, P. B. & Johnston, P. G. 1987 X-linked gene expression and X-chromosome inactivation: marsupials, mouse, and man compared. *Isozymes: Curr. Top. Biol. med. Res.* **16**, 225–253.

Watt, F. & Molloy, P. L. 1988 Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Develop.* **2**, 1136–1143.

Yuan, R. & Hamilton, D. L. 1984 Type I and type II restriction-modification enzymes. In *DNA methylation: biochemistry and biological significance* (ed. A. Razin, H. Cedar & A. D. Riggs), pp. 11–37. New York: Springer–Verlag.

Yuan, R., Hamilton, D. L. & Burckhardt, J. 1980 DNA translocation by the restriction enzyme from *E. coli* K. *Cell* **20**, 237–244.

Zehnbauer, B. A. & Vogelstein, B. 1985 Supercoiled loops and the organization of replication and transcription in eukaryotes. *BioEssays* **2**, 52–54.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES