Opinion TRENDS in Genetics Vol.23 No.3

Full text provided by www.sciencedirect.com ScienceDirect

What are the molecular ties that maintain genomic loops?

Davide Marenduzzo[1*], Inês Faro-Trindade[2,3*] and Peter R. Cook[2]

[1] School of Physics, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, UK
[2] Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK
[3] Present address: Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Lower Ground Floor, Wernher & Beit Building South, Groote Schuur Campus, Observatory 7925, South Africa

The formation of genomic loops by proteins bound at sites scattered along a chromosome has a central role in many cellular processes, such as transcription, recombination and replication. Until recently, few such loops had been analyzed in any detail, and there was little agreement about the nature of the molecular ties maintaining these loops. Recent evidence suggests that loops are found in both prokaryotes and eukaryotes, and that the transcription machinery is a molecular tie. In addition, results obtained using site-specific recombination in bacteria and chromosome conformation capture in eukaryotes support the idea that active transcription units are in close contact. These data are consistent with a model for genome organization in which active polymerases cluster into transcription 'factories', which, inevitably, loops the intervening DNA. They are also consistent with the ties functioning as barriers, silencers, enhancers or locus control regions, depending on their positions relative to other genes.

Introduction

There is little agreement about how genomes are organized within cells. One enduring idea is that genomes are looped, with the position of a gene within the loop determining its activity. For example, regulatory motifs such as enhancers and silencers can regulate the activity of target genes located many tens of kilobases away on the chromosome, and it is often suggested that they do so by inducing looping to bring the motif and its target into molecular contact. However, conclusive evidence for such ideas has been lacking, mainly because, until recently, so few loops had been characterized in molecular detail. But one thing is clear: genome sequencing projects have failed to uncover any potentially responsible motifs that are conserved among prokaryotes and eukaryotes. We can envisage that distant segments of one DNA molecule might come into contact with each other directly or through an RNA or protein tie. Protein ties have proved the most popular of these possibilities, and a variety of candidates have been proposed (including topoisomerases and matrix-binding proteins). Here, we review new evidence that supports a unifying view of how all genomes are organized: active RNA polymerases and their transcription factors function as the main molecular ties that maintain most loops and determine gene activity. For details about specialized ties that carry out particular functions (e.g. the segregation machinery that orients whole chromosomes), see Ref. [1].

Theoretical considerations suggest that components of the transcription machinery – that is, bound transcription factors and polymerizing complexes – are likely to cluster, resulting in looping of the intervening DNA (Box 1). Recent experimental results also support the idea that these components are the main ties in both prokaryotes and eukaryotes.

Loops and ties in prokaryotes

It has long been known that lysis of bacteria in 1 M NaCl releases naked, supercoiled DNA, which is looped by attachment to clusters of polymerases that are still engaged [2] [Figure 1a(iv)]. The clustering of polymerases depends on transcription, because it is eliminated by pre-treatment with rifampicin or post-treatment with ribonuclease. However, few researchers originally thought that such clusters existed *in vivo*, mainly for two reasons. First, such nucleoids (see Glossary) could be isolation artifacts, occurring when structures similar to those in Figure 1a(i) or Figure 1a(iii) are converted to that in Figure 1a(iv).

Glossary

Chromosome conformation capture (3C): a method used to determine which DNA sequences lie close together in 3D space in fixed cells (Figure 2).
Depletion attraction: a noncovalent force (known to physicists but few biologists). It has an osmotic basis and is seen only in crowded environments, such as those in cells (see Figure Ib in Box 1); it can bring together complexes that do not interact with each other in dilute solutions.
Lampbrush loops: named because of their similarity to the brushes used to clean gas lamps; these brushes are equivalent to test-tube brushes (which are also outmoded). Lampbrush loops can be isolated from the oocytes of nearly all animals, except mammals (presumably because the appropriate conditions have not yet been developed). They are extraordinarily large chromosomes, and their size has made them useful for studies of chromosome structure.
Miller spreads: O.J. Miller invented a simple method for visualizing 'genes in action'. Bacterial spheroplasts (or eukaryotic nuclei) are swollen in what is, in essence, distilled water and a household detergent, and the dispersed transcription units are spun through a fixative onto a grid used for electron microscopy.
Nucleoid: these were originally isolated from bacteria. Spheroplasts were lysed in a detergent and 1 M NaCl to release a rosette of loops of naked, supercoiled DNA attached to a cluster of still-engaged RNA polymerases. Eukaryotic nucleoids are prepared similarly by lysing whole cells.
Transcription factory: a site containing a number of transcript-producing machines (including RNA polymerases and the machinery that processes the transcript) active on several different templates.

Corresponding author: Cook, P.R. (peter.cook@path.ox.ac.uk).
* Authors contributed equally.
Available online 5 February 2007.

www.sciencedirect.com 0168-9525/$ – see front matter © 2006 Elsevier Ltd. All rights reserved. doi:10.1016/j.tig.2007.01.007

Box 1. Two interactions driving looping

Theoretical considerations make the transcription machinery a strong candidate for a molecular tie [64]. Consider two DNA-binding proteins [Figure Ia(i); red diamonds]. If these proteins are present at ~1 nM and interact with a K_d of 10^{-7} M (values typical of nuclear proteins), <1% will be complexed together. But, if they bind to the same DNA molecule at sites 10 kb apart, the resultant local concentration ensures that 66% of these proteins will be in the complex; then, it is inevitable that a loop forms [Figure Ia(ii)]. Chemists call this the chelate effect, and it can do more than just facilitate looping. For example, the *lac* operon repressor and the λ repressor each bind to more than two target sites when they generate one loop, and full occupancy of the repressor-binding site involves large complexes in which individual repressors bind to each other with increased affinities and greater cooperativity. In turn, this should reduce the transcriptional noise within one cell and the variability in transcription rate in different cells [65]. These additional effects are also advantageous to organisms with large genomes; full site occupancy can be achieved using lower affinities, so bound proteins can be displaced easily by the replication and transcription machineries. Importantly, loops generated in this way cannot persist; tagging transcription factors with GFP shows that residence half-lives on DNA are generally <10 s [66].

A poorly appreciated force known as 'depletion attraction' (see Glossary) probably also contributes to looping [67]. Consider Figure Ib(i), in which several large spheres (red) are surrounded by many small spheres (purple), which represent the many small macromolecules crowded into a cell [68]. In physicists' terminology, both types of sphere are hard and noninteracting, so none of the forces familiar to biologists (i.e. ionic bonds, hydrogen bonds, van der Waals forces and hydrophobic interactions) acts between these spheres. The small spheres bombard the large ones from all sides (brown arrows). When two large spheres come into contact, the small ones are excluded from the volume between the two. Therefore, the small spheres exert an unopposed force equivalent to their osmotic pressure on opposite sides of the two large spheres, keeping the large spheres together. This osmotic effect depends on the volume that is inaccessible to the small spheres. If the small spheres could access this (depleted) volume, they would force the two large ones apart. The scale of the depletion attraction depends on size: the larger the inaccessible volume, the larger the attraction. This force will also be larger if the two large objects fit snugly together and smaller if surface irregularities or charge limit close contact. This depletion attraction force acts on all large complexes in a cell.

Consider the case in which the large spheres represent active transcription complexes (whether they be pre-assembled or not [69]). In this case, the force could act between polymerizing complexes (probably together with other specific attractions) for as long as those polymerases remained engaged: that is, seconds in bacteria, minutes in humans and, perhaps, even hours (because ~21% of active RNA polymerase II is paused in *Drosophila melanogaster* [70]). Each complex might contain a multisubunit polymerase, the transcript and its neutralizing proteins, plus associated ribosomes (in bacteria) or a spliceosome (in eukaryotes); therefore, these complexes are large enough to generate an attraction approximately equal to the entropic costs of looping [Figure Ib(ii)]. Figure Ic illustrates the results of a computer simulation in which 0.4 Mb of human DNA [associated with active RNA polymerase II complexes (modeled as beads) spaced every 20 kb] was allowed to diffuse. The energy associated with the depletion attraction between two beads was conservatively estimated to be 4 k_BT (where k_B is the Boltzmann constant and T the absolute temperature); this is approximately equivalent to the energy contained in 3 hydrogen bonds. Eighteen of the 21 beads are in clusters, and the larger clusters are stabilized, because one bead can interact with more than one other bead.

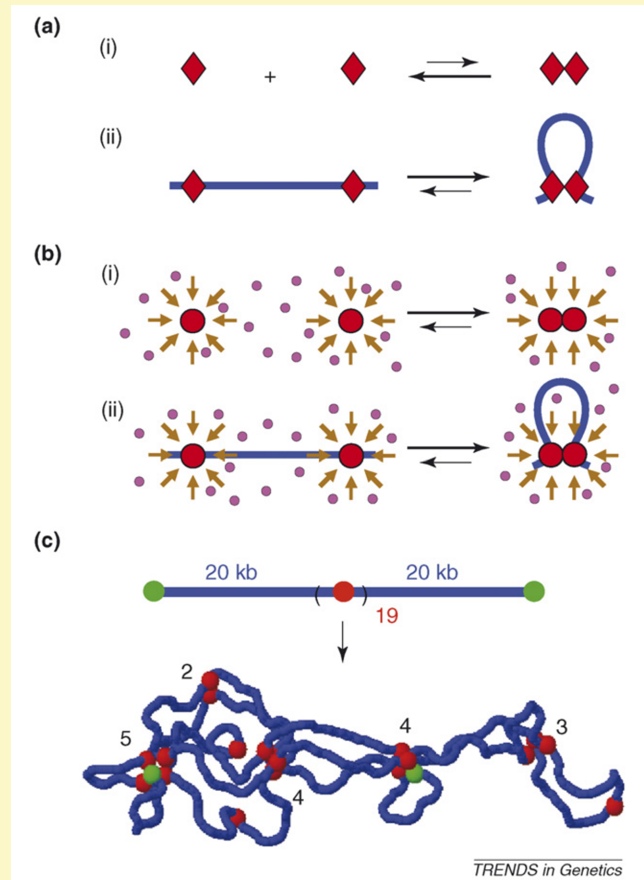

Figure I. Forces driving looping. **(a)** Specific interactions between DNA-binding proteins. (i) If two DNA-binding proteins are present at ~1 nM and interact with a K_d of 10^{-7} M, <1% dimerize. (ii) On adding a DNA molecule with two binding sites that are 10 kb apart, protein binding creates a local concentration that drives 66% of DNA-binding proteins into the complex (forming a loop). These values are typical of many transcription factors. **(b)** Nonspecific (entropic) depletion attraction between polymerizing complexes. (i) In cells, which are crowded, many small soluble macromolecules bombard large complexes from all sides (brown arrows). When two complexes come into contact, the small molecules exert a force equivalent to their osmotic pressure on opposite sides of the two large molecules, keeping the large molecules together. The scale of the attraction can be calculated reasonably accurately, and the equilibrium is to the right if the complexes are much larger than the small molecules. (ii) When large spheres (representing polymerases) are threaded on a string (representing DNA or chromatin), the depletion attraction is partially countered by the entropic cost of looping the string. **(c)** Monte Carlo simulation of looping induced by the depletion attraction. Starting with a linear string (chromatin fiber containing 0.4 Mb DNA) containing equally spaced 21 beads (polymerase-containing complexes; red and green circles), modeling diffusion in the presence of the depletion attraction yields the structure shown. The number of beads in each cluster is indicated, with the position of terminal beads (green) and internal beads (red) shown. Clustering will be augmented by specific forces [e.g. those shown in (a)]. Modified, with permission, from Ref. [67] © (2006) the Biophysical Society.

And, second, polymerases are immobilized and, therefore, cannot track along their templates, and it was assumed that this would prevent them from working efficiently. Nevertheless, the presence of the supercoiling that is associated with such loops can be detected. A single-strand break releases supercoiling only in one loop (or topological domain), because the broken end of one DNA strand can rotate about the other intact strand until all of the energy of supercoiling is lost in that loop. In addition, the nicking of DNA in live cells, by γ-irradiation, progressively reduces the binding of [^{3}H]trimethylpsoralen – a probe that binds preferentially to supercoiled DNA – and these results are consistent with the existence of many loops [2]. But what is the identity of the molecular ties that maintain such loops?

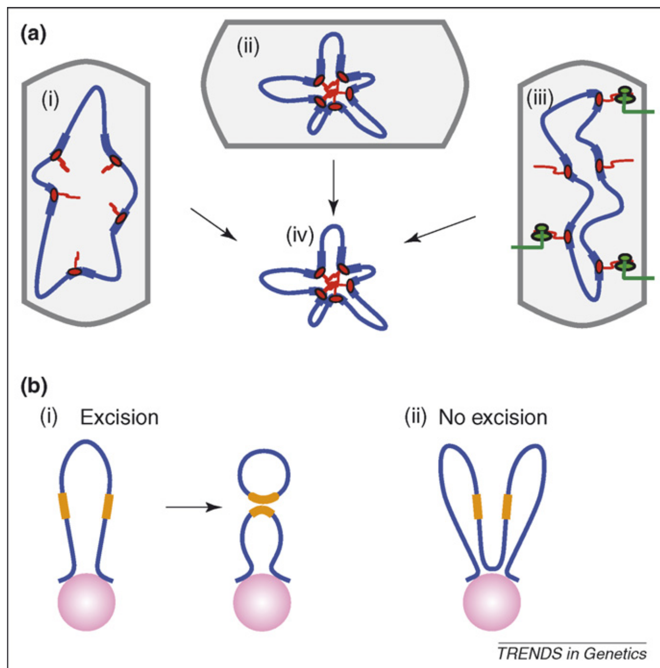

Figure 1. Genome organization in bacteria. **(a)** Models of genome organization. (i) The genome might be randomly organized. (ii) Active polymerases (red ovals) might cluster. (iii) Loops might be attached to the nuclear membrane. Some transcription units encode membrane proteins, and – owing to coupled transcription and translation – nascent peptides insert (cotranscriptionally) into membranes to form loops (between the membrane-attachment points). (iv) Isolated nucleoids could be derived directly from genomes organized as shown in (ii) or artifactually from those shown in (i) and (iii). **(b)** Mapping attachments (barriers) using resolvases. (i) Resolvases are dimers. One half binds to one *res* site (orange), and the other tracks along the DNA until it finds a second; this brings the two *res* sites together to form a loop. The enzyme now excises the intervening DNA, and this excision can be detected if the excised fragment contains the appropriate genetic marker. (ii) An attachment (barrier) between these two *res* sites that creates two loops prevents the resolvase from tracking from one site to the other and, therefore, excision does not occur. Barriers are formed by active transcription units.

Many candidates have been suggested [3]. These include the following: (i) knots or tangles in DNA; (ii) topoisomerases that can form a bridge between two DNA segments (e.g. DNA gyrase); (iii) nucleoid-associated proteins, such as the histone-like protein HU, integration host factor (Ihf), histone-like nucleoid-structuring protein (Hns) and factor for inversion stimulation (Fis), the last two of which make loops *in vitro*; (iv) repressors such as Gal, AraC and λ, which also make loops *in vitro*; (v) the actin homolog MreB; (vi) several unexpected proteins (e.g. phosphoglucomutase and a transketolase); and (vii) DNA translocases such as FtsK and the MukBEF complex. The MukBEF complex contains conserved SMC (structural maintenance of chromosome) proteins that include condensins and cohesins, but there are too few copies of MukB to maintain many loops. Because no single candidate has all of the right credentials, is it possible that they work together? In this case, loops would appear and disappear depending on the vagaries of Brownian motion, being tied transiently by some (or all) of the proteins noted.

Membrane-bound proteins have also been popular candidates; the coupled transcription, translation and insertion of nascent proteins into membranes would form loops [Figure 1a(iii)]. Indirect support for membrane-bound proteins as ties came from analyses of supercoiling in plasmids carrying *tetA* – which encodes a membrane pump – in cells with mutant topoisomerases, but these candidates have been discounted recently [3]. The *Tn*3 and γδ resolvases can detect two of their target *res* sites in one domain in living *Salmonella enterica* serovar Typhimurium (*S. typhimurium*); half of the enzyme binds to one site, and the other scans until it finds a second (bringing the two together), and then the intervening DNA is excised [Figure 1b(i)]. No deletion occurs if the sites are in different chromosomal loops, because the resolvase cannot track through the attached ties that create the loops [Figure 1b(ii)]. Moreover, when pairs of sites were inserted at increasing separations, followed by induction of the resolvase and monitoring of the excision frequency, the frequency was found to decline with increasing separation, consistent with loops of 11–20 kb. And although *tetA* was a barrier that prevented tracking (and, therefore, excision), genes encoding non-membrane proteins (e.g. *lacZ* and *kan*) were equally effective.

What property of these other genes makes them effective barriers? The answer is transcriptional activity; any gene can function as a barrier but only if it is transcribed [3]. This can be demonstrated by the insertion between two *res* sites of a promoter driving *lacZ*. When *lacZ* is not induced, DNA is excised (and is present in one long loop; Figure 2); when induced, DNA is not excised (because a tie now blocks tracking between *res* sites) [Figure 1b(ii)]. Excision efficiency inversely correlates with promoter strength, and this is consistent with strong promoters being better barriers.

If active transcription units lie at the points of attachment, then loop length should equal the spacing between engaged polymerases. This is the case, and the length is ~11 kb [4], consistent with the results from the resolvase assay discussed here. Microarrays also yield similar results: the activity of ~300 promoters in *Escherichia coli* changes when expression of a restriction enzyme releases supercoiling, and analysis of the distances between restriction sites and affected promoters (which lie in the same

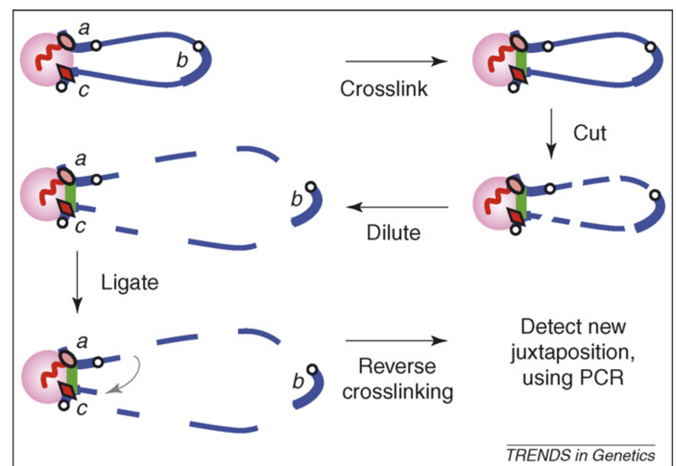

Figure 2. The 3C method. A loop containing genes *a*, *b* and *c* is shown. 3C involves fixation to crosslink DNA sequences that lie next to each other (usually through DNA–protein–DNA links; green), before cutting with a restriction enzyme, dilution and ligation. Dilution favors intramolecular ligation: that is, the end of one DNA molecule in a DNA–protein–DNA complex is joined to the end of the second DNA molecule in the same complex more frequently than to the end of a different molecule or complex. Then, two DNA sequences that were initially in close proximity are ligated (i.e. *a* with *c*, but not *a* with *b*), and (after reversing the crosslinks) the novel juxtaposition is detected by PCR.

domain) indicates that loops are ~9 kb [4]. Inspection of Miller spreads also shows that polymerases on open reading frames are spaced every 10–20 kb [5]. [The 'Christmas trees' in the same spreads are packed with polymerases about every 80 bases, but these are rRNA (*rrn*) operons, which are exceptions.] Moreover, polymerases must be spaced every 9–24 kb to maintain mRNA levels at the levels that are seen using microarrays [6]. We conclude that the experimental evidence suggests that active polymerases are the molecular ties that maintain genomic loops.

Do active units cluster? The *E. coli* genome encodes seven *rrn* operons, but a rapidly growing cell contains ~22 operons, owing to re-replication before division [6]. Because these operons are transcribed by ~70% of all polymerases active in the cell, they can be visualized by tagging with green fluorescent protein (GFP). Most cells contain only one to three GFP foci (and none have more than six), consistent with the clustering of active operons [7].

Loops and ties in eukaryotes

The evidence for looping in eukaryotes is similar to that for bacteria. Imaging reveals loops in lampbrush chromosomes attached to axial chromomeres [8], and supercoiling is found in living cells and isolated nucleoids [9,10]. In addition, after permeabilizing cells or isolating nuclei, cutting an unlooped fiber should progressively release shorter fragments; however, no long fragments are seen, and the kinetics fit two cuts releasing short fragments from loops [11,12]. In addition, enhancers on plasmids influence promoters on other plasmids if the two make contact, implying that they might do so in their natural chromosomal location [13,14].

Results from chromosome conformation capture (3C; Figure 2) now provide conclusive evidence for looping: if two linked genes lie together in nuclei, the intervening DNA must be looped. Many different loops have now been detected using 3C [15–18], and we describe one that has been analyzed thoroughly [19,20]. This loop brings the mouse *Hbb-b1* gene (which encodes β-globin) into contact

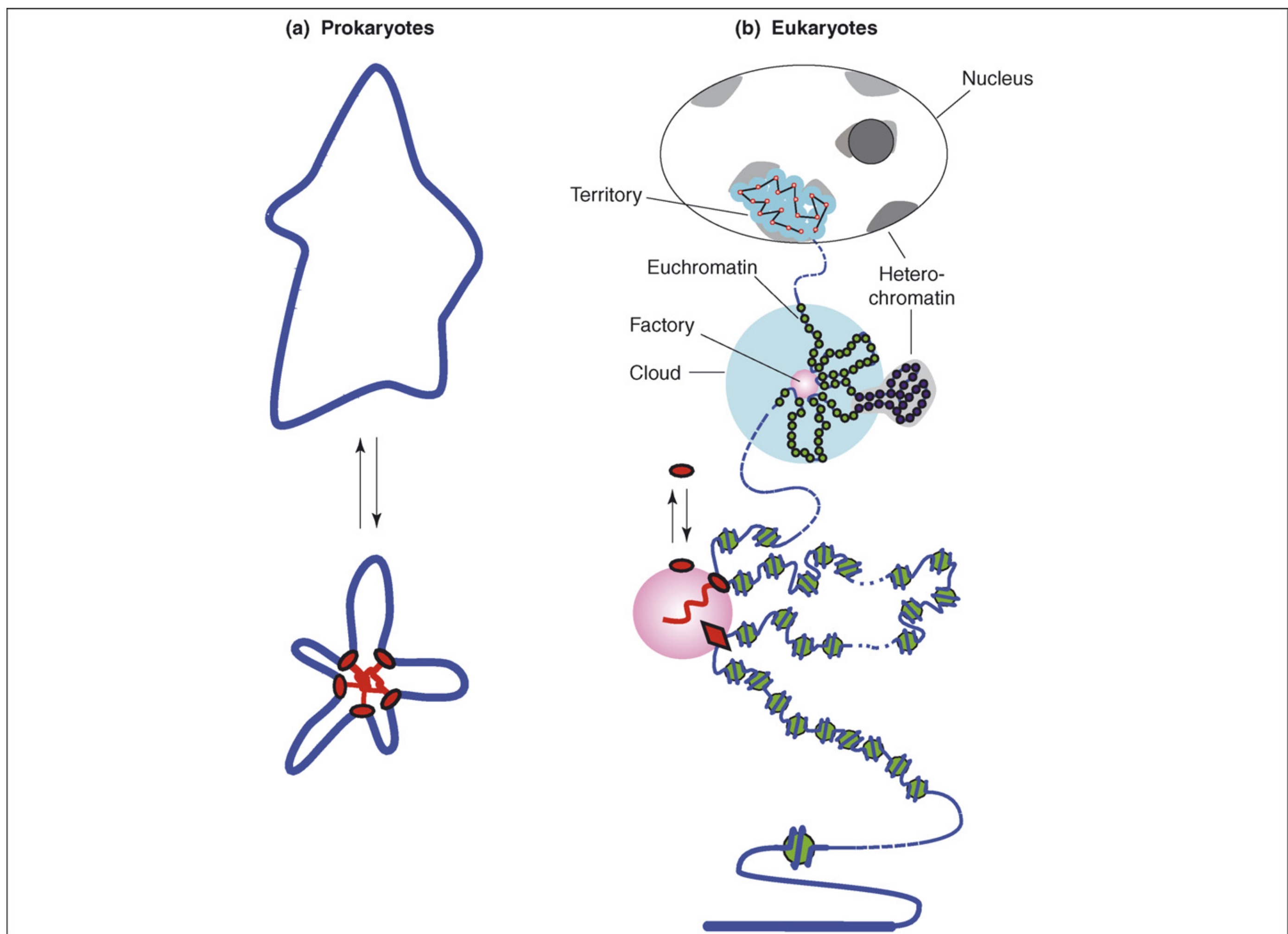

Figure 3. Genome structure in bacteria and eukaryotes. **(a)** In bacteria, transcription of the circular chromosome, followed by aggregation of active transcription complexes, generates loops. **(b)** In eukaryotes, a zigzagging string of nucleosomes (green circles) is attached in loops to a factory (pink circle) through an active polymerase (red oval) and transcription factors (red diamonds). A cloud of loops envelopes the factory, resulting in a structure equivalent to the looped bacterial genome shown in (a), and successive clouds form a territory. Components in the factory exchange with the soluble pool, and attachments are made and broken as polymerases initiate and terminate (or transcription factors bind and dissociate). Nucleosomes in long loops are less mobile and acquire a heterochromatic histone code that spreads down the fiber; they also aggregate on the lamina, nucleoli and chromocenters. In a HeLa cell, there are ~16 loops (average length ~86 kb; range 5–200 kb) and ~8 active RNA polymerase II molecules per factory, with ~100 clouds per chromosome. Part (a) modified, with permission, from Ref. [40] © (2002) Macmillan Publishers. Part (b) modified, with permission, from Ref. [61] © (2001) John Wiley & Sons.

with its locus control region (LCR) and with a gene (*Eraf*) encoding an α-globin-stabilizing protein, which lie ~0.02 and ~25 Mb away, respectively. This LCR is thought to nucleate an 'active chromatin hub' (also known as a factory) that facilitates expression of globin-related genes. Fixed cells are used for 3C, so the observed loops could – in principle – be artifacts; however, the Dam identification (DamID) method provides evidence for looping in living cells in *Drosophila melanogaster* [21].

Studies using nucleases and both isolated nucleoids and cells permeabilized in isotonic buffers originally pointed to active transcription units being important ties: the nucleases detach little nascent RNA and few active units [19,22–24]. These results have now been confirmed using 3C and DamID. For example, the contacts between *Hbb-b1*, the LCR and *Eraf* are seen only in erythroid nuclei (in which all three are transcribed) but not in brain cell nuclei (in which *Hbb-b1* is inactive [19,20]). Moreover, the contacts that *Hbb-b1* makes with other genomic regions depend on its activity; in erythroid nuclei, 80% of contacts are with other active genes, but, in the brain, this value falls to only 13% [25]. It also turns out that many other LCRs are transcribed [26], with transcription of the human growth hormone LCR even being required for LCR function [27]. Thus, 3C also shows that sequences lie close together only when they are being transcribed. Moreover, a range of transcription factors have been implicated in mediating looping [26,28], and, in the case of the *Hbb-b1* locus, these include EKLF (erythroid Kruppel-like factor; also known as KLF1), GATA1 (GATA-binding protein 1) and the zinc-finger protein FOG1 (also known as ZFPM1) [29,30].

If active polymerases are important ties, then inhibiting them should eliminate looping and decondense genomes; this is the case. Actinomycin D prevents lampbrush loops from forming when sperm heads (which contain unlooped DNA) are injected into frog oocytes [31], whereas DRB (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole) and α-amanitin disperse DNA and increase genomic mobility [32,33].

Do active transcription units segregate domains, as is the case in bacteria? Some prototypical examples showing that they do include the following [26].

(i) In budding yeast, the best-studied barriers (those that flank silent *HMR*) contain promoters, and mutating the promoter in $tRNA^{Thr}$ reduces barrier activity [34].
(ii) In fission yeast, barriers that block heterochromatic spread from *cen1* contain tRNA genes, and mutating the $tRNA^{Ala}$ promoter reduces barrier activity [35].
(iii) The first 'barriers' to be defined (*D. melanogaster scs* and *scs′*) encode promoters [26].
(iv) The best-characterized vertebrate boundary (HS4 of the chicken β-globin gene) is a CpG island with all of the hallmarks of a promoter (i.e. DNase-I hypersensitivity, histone H3 hyperacetylation and Lys4 methylation), although an initial study failed to detect any promoter activity [36].
(v) Any active unit can insulate GFP reporters from human heterochromatin [37].

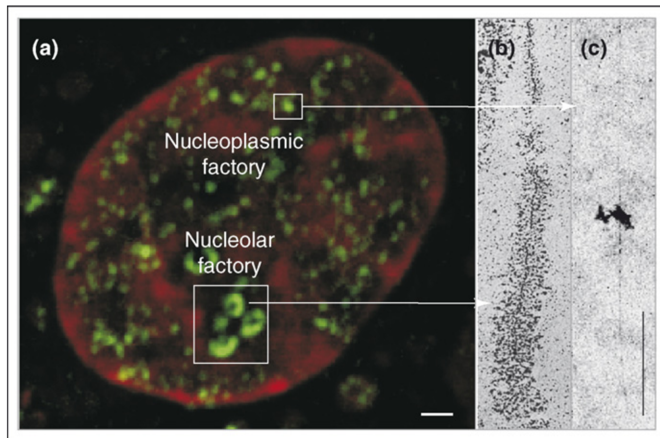

Figure 4. Different views of nascent transcripts. **(a)** Fluorescent micrograph of nascent RNA. HeLa cells were permeabilized, and nascent RNA was extended in the presence of 5-bromouridine 5′-triphosphate. Cells were then cryosectioned (100 nm), and the resultant bromo-RNA was immunolabeled with fluorescein isothiocyanate (FITC; green) and nucleic acids counterstained with TOTO-3 (red). Fluorescence images were collected by confocal microscopy. Newly made bromo-RNA is concentrated in factories in the cytoplasm (where it is made by mitochondrial polymerases), nucleoplasm and nucleoli. **(b,c)** Electron micrographs of spread transcription units. In (b), a crescent similar to one of the four in the nucleolar factory in (a) has been stripped from the underlying structure; ~125 transcripts can be seen engaged on the rRNA unit. In (c), one active transcription unit is shown; it is formed by breaking up a nucleoplasmic factory such as the one shown in (a) to release ~8 active units. Scale bars, 1 μm. Part (a) reproduced, with permission, from Ref. [26] © (2003) The Company of Biologists. Part (b) reproduced, with permission, from Ref. [62] © (1972) the Society of the European Journal of Endocrinology. Part (c) reproduced, with permission, from Ref. [63] © (1998) The American Society for Cell Biology.

A model for all genomes

These recent results support a model for genomic organization (Figure 3) in which the central feature is the clustering of active units into 'transcription factories', with engaged polymerases and transcription factors functioning as the ties [38–40]. We do not wish to suggest that these are the only ties, just that they are the main ones. Further support for this model comes from imaging (which reveals nascent RNA in foci; Figure 4a), and quantitative analysis (which indicates that one focus contains many active units, enzymes and transcripts [41,42]). These foci also remain in place after removal of >90% of chromatin and, therefore, are attached to the underlying structure [43]. Clustering ensures high local concentrations of polymerase, enabling efficient interaction: HeLa cell nuclei contain a pool of RNA polymerase II of ~1 μM, but the local concentration in a factory (diameter ~75 nm) is ~1 mM. So, few transcripts are made outside factories [40]. Moreover, repeated transcription of related genes will lead to the incorporation of certain transcription factors into some factories that now specialize in transcribing only those genes [44]. For example, active RNA polymerase I, II and III are each concentrated in dedicated factories [45], with some RNA polymerase II factories transcribing globin-related genes (as discussed earlier).

At first glance, this model has disadvantages. For example, in this model, polymerases do not move along their templates, as is commonly thought; instead, they are immobilized, and they reel in their templates. Evidence for tracking by polymerases seems to be of two types [39]. First, there is a perception that a polymerase is smaller than its template and, therefore, that the polymerase moves; however, it is now known that polymerizing

complexes dwarf their templates and that immobilized enzymes work efficiently [46–48]. A polymerase that tracks along a helix also generates a transcript entwined about its template, but no satisfactory untwining mechanism has been proposed so far. No such problem arises with a fixed enzyme. [Topoisomerases solve a distinct problem – removing the 'twin domains' of supercoiling created by a polymerase (whether mobile or not).] Second, tracking polymerases seem to be caught in the act in the 'Christmas tree' shown in Figure 4b and in the lone transcription unit in Figure 4c. But these images are prepared by tearing active units away from the underlying structures. Spreading under different conditions yields clusters of transcription units (not 'trees' [9]), and GFP tagging confirms that active bacterial enzymes cluster (as discussed earlier). Images of lampbrush chromosomes are similarly interpreted as showing that polymerases track around loops. But nascent RNA and active RNA polymerase II are even more concentrated in axial chromomeres to which the loops are attached, and no loops are seen in whole-cell sections in which chromatin appears as a granular aggregate [31,49]. Once again, units must be torn away during spreading. Therefore, we conclude that there is little evidence for tracking.

Another apparent disadvantage of this model stems from the considerable amounts of heterochromatin found in higher eukaryotes; because this is gene free and transcriptionally inert (so it is assumed), there would be few engaged polymerases to function as ties. But we now know that some heterochromatin is transcribed [50], that centromeric heterochromatin even depends on continuing transcription for compaction [51], and that euchromatin and deep heterochromatin contain similar densities of transcription sites [52]. Clearly, this assumption that heterochromatin is transcriptionally inert is incorrect, so euchromatin and heterochromatin could be organized as shown in Figure 3b.

Finally, it has been stated that active genes are found only on the surfaces of chromosome territories, but this conclusion was based on evidence from only a few genes. It is now known that active genes are more widely spread [53,54]. Moreover, it can be envisaged that, although adjacent genes would tend to be transcribed in the same factory (as shown in Figure 3b), they would sometimes be transcribed in different ones (see Figure Ic in Box 1, in which neighboring beads can be found in different clusters, and a bead near one end of the string can be found in the same cluster as a bead near the other end). Furthermore, the intermingling of chromosome territories will enable genes on different chromosomes to attach to the same factory [54,55].

Changing ties during development

How does proximity to a factory affect the initiation of transcription? Consider a typical loop in a HeLa cell (Figure 5a). Simulations show [56] that segment a in the loop has a tether that is too short to reach a binding zone in the factory and, therefore, will never be active while the loop persists. This might underlie 'transcriptional interference', in which the activity of one gene prevents transcriptional initiation by a neighbor [26]. (Such interference is rarely detected, because most cells contain two gene copies, and the activity of one masks the inactivity of the other.) Segment b often accesses the binding zone; segment c, less so. In this case, euchromatin and heterochromatin would be in the 'hot' segment (b), and the 'cold' segment (c), respectively. It turns out that c constitutes ~75% of the loop – the known heterochromatic fraction in a HeLa cell. Moreover, increasing loop length, thickness and rigidity (all typical of heterochromatin) reduce access to the binding zone, reinforcing the idea that distant segments are heterochromatic [56].

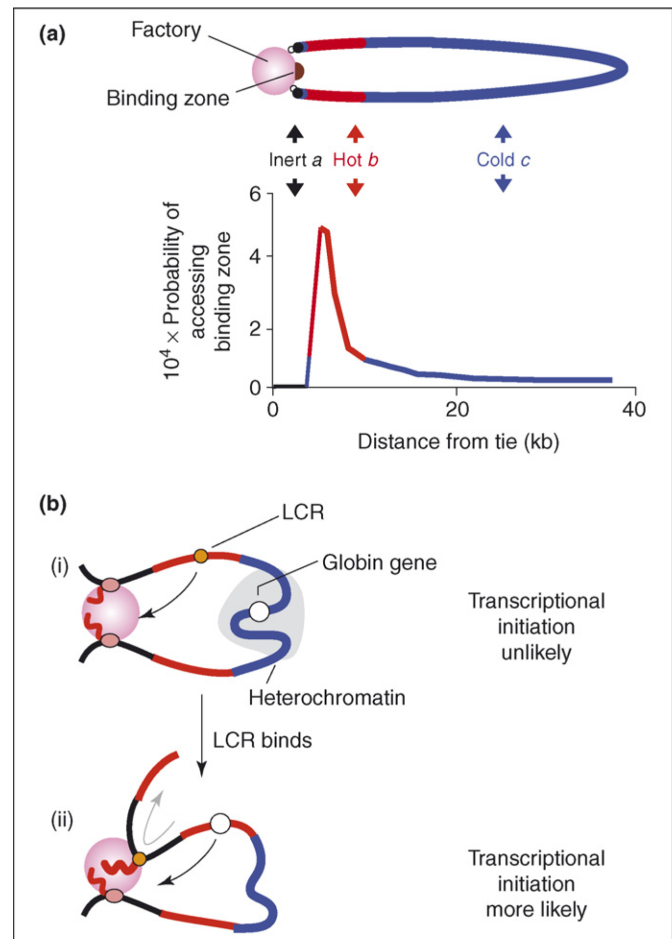

Figure 5. Regulation of eukaryotic gene expression. **(a)** A 77 kb chromatin fiber attached to a 75 nm factory (typical of a HeLa cell) is indicated. The graph (which was determined by computer simulation of this loop as it diffused around the factory) shows the probabilities that promoters at different distances from a tie are found in the binding zone (brown) on the factory (pink circle) surface. 'Inert' (black, *a*), 'hot' (red, *b*) and 'cold' (blue, *c*) segments of DNA are indicated. For simplification, ties are assumed to be permanent. **(b)** Activating expression of a gene encoding globin in erythrocyte development. (i) A loop is attached to a factory through two engaged polymerases (pink ovals). (For simplification, the bottom one is paused.) The gene is in the heterochromatic (cold) segment (blue) and rarely accesses the factory. The LCR is in the euchromatic (hot) segment (red) and frequently comes into contact with the factory; however, transcription of the LCR is not initiated, because the appropriate transcription factors are not present. (ii) During erythroblast development, the concentration of the relevant transcription factors increases, allowing a promoter in (or around) the LCR to initiate transcription. Because the upper polymerase in (i) has terminated, the loop is now attached through one that is transcribing the LCR. This reels in the template (gray arrow) to 'open' the heterochromatin. The promoter of the globin-encoding gene is now in a hot segment, and transcription is much more likely to be initiated. Transcription factors also function as additional ties, creating additional loops (not shown). In practice, activation of the globin-encoding gene probably involves a complicated cascade of attachments and de-attachments. Part (a) modified, with permission, from Ref. [56] © (2006) Elsevier. Part (b) modified, with permission, from Ref. [20] © (2005) Macmillan Publishers.

How might a globin-encoding gene become active during development? A plausible scheme involves the creation of subloops to move the globin-encoding gene into a 'hot' segment, where there is a high probability that transcription will be initiated (Figure 5b). What then happens later, when transcription declines and the nucleus becomes small and pycnotic? Recent studies of the differentiation of pluripotent embryonic stem cells into parietal endoderm show that nuclear volume and the number of active polymerases both halve, whereas factory density and diameter remain constant [41]. In this case, fewer (longer) loops would be expected. This is consistent with results from the 1970s that show that differentiation of transcriptionally active chicken erythroblasts into erythrocytes is accompanied by a progressive increase in loop length until no loops (or transcription) remain [9]. It can also be envisaged that the activation of a promoter might silence a neighboring one (by shifting it into zone *a* or *c*) or create a barrier that prevents heterochromatic spread [26].

Conclusions

It is usually thought that different molecular ties organize genomes in prokaryotes and eukaryotes. Theory and recent evidence provide strong evidence that the transcription machinery is a major component that is common to the organization of all genomes. Such molecular ties will, of necessity, change from moment to moment. Each bacterial cell in a culture (or each eukaryotic cell at the same stage in a developmental lineage) contains approximately the same range of active transcription units strung along the chromosome and approximately the same number of polymerase clusters (or factories). But a specified gene might be present in a loop at one moment and attached in the next moment, and, in another cell, the precise attachments around that gene would rarely be the same. When a bacterium adapts to changing culture conditions (or when a eukaryotic cell differentiates), a different 'constellation' of loops forms. In principle, this model can be tested. Nascent transcripts copied from many different genes simultaneously [57] can be localized to determine whether they cluster. Moreover, as microscopes with increased resolution become available, GFP-tagged polymerases can be observed in living cells to determine whether they cluster when they are active [58]. However, many questions remain. Do clusters associate with an actin and/or lamin nucleoskeleton [59,60]? What happens in mitosis when transcription stops but loops persist [12]? And how are the ties discussed here integrated with others involved in replication and segregation?

Acknowledgements

We thank our colleagues for helpful discussions. We also thank the following organizations for support: the Biotechnology and Biological Sciences Research Council, Cancer Research UK, the Engineering and Physical Sciences Research Council, the Portuguese Foundation for Science and Technology, the Medical Research Council, and the Wellcome Trust.

References

1 Spector, D.L. (2003) The dynamics of chromosome organization and gene regulation. *Annu. Rev. Biochem.* 72, 573–608
2 Pettijohn, D.E. (1996) The nucleoid. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F.C. *et al.*, eds), pp. 158–166, American Society for Microbiology Press
3 Deng, S. *et al.* (2005) Organization of supercoil domains and their reorganization by transcription. *Mol. Microbiol.* 57, 1511–1521
4 Postow, L. *et al.* (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev.* 18, 1766–1779
5 French, S.L. and Miller, O.L. (1989) Transcription mapping of the *Escherichia coli* chromosome by electron microscopy. *J. Bacteriol.* 171, 4207–4216
6 Bon, M. *et al.* (2006) Many expressed genes in bacteria and yeast are transcribed only once per cell cycle. *FASEB J.* 20, 1721–1723
7 Cabrera, J.E. and Jin, D.J. (2003) The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues. *Mol. Microbiol.* 50, 1493–1505
8 Morgan, G.T. (2002) Lampbrush chromosomes and associated bodies: new insights into principles of nuclear structure and function. *Chromosome Res.* 10, 177–200
9 Jackson, D.A. *et al.* (1984) Replication and transcription depend on attachment of DNA to the nuclear cage. *J. Cell Sci. Suppl.* 1, 59–79
10 Matsumoto, K. and Hirose, S. (2004) Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of *Drosophila*. *J. Cell Sci.* 117, 3797–3805
11 Igo-Kemenes, T. and Zachau, H.G. (1977) Domains in chromatin structure. *Cold Spring Harbor Symp. Quant. Biol.* 42, 109–118
12 Jackson, D.A. *et al.* (1990) The size of chromatin loops in HeLa cells. *EMBO J.* 9, 567–571
13 Mueller-Storm, H.P. *et al.* (1989) An enhancer stimulates transcription *in trans* when attached to the promoter via a protein bridge. *Cell* 58, 767–777
14 Mahmoudi, T. *et al.* (2002) GAGA can mediate enhancer function *in trans* by linking two separate DNA molecules. *EMBO J.* 21, 1775–1781
15 Spilianakis, C.G. and Flavell, R.A. (2004) Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* 5, 1017–1027
16 Murrell, A. *et al.* (2004) Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat. Genet.* 36, 889–893
17 Blanton, J. *et al.* (2003) Protein:protein interactions and the pairing of boundary elements *in vivo*. *Genes Dev.* 17, 664–675
18 Lomvardas, S. *et al.* (2006) Interchromosomal interactions and olfactory receptor choice. *Cell* 126, 403–413
19 de Laat, W. and Grosveld, F. (2003) Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* 11, 447–459
20 Chakalova, L. *et al.* (2005) Replication and transcription: shaping the landscape of the genome. *Nat. Rev. Genet.* 6, 669–677
21 Cleard, F. *et al.* (2006) Probing long-distance regulatory interactions in the *Drosophila melanogaster* bithorax complex using Dam identification. *Nat. Genet.* 38, 931–935
22 Jackson, D.A. and Cook, P.R. (1985) Transcription occurs at a nucleoskeleton. *EMBO J.* 4, 919–925
23 Dickinson, P. *et al.* (1990) Active RNA polymerase I is fixed within the nucleus of HeLa cells. *EMBO J.* 9, 2207–2214
24 Jackson, D.A. and Cook, P.R. (1993) Transcriptionally active minichromosomes are attached transiently in nuclei through transcription units. *J. Cell Sci.* 105, 1143–1150
25 Simonis, M. *et al.* (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38, 1348–1354
26 Cook, P.R. (2003) Nongenic transcription, gene regulation and action at a distance. *J. Cell Sci.* 116, 4483–4491
27 Ho, Y. *et al.* (2006) Locus control region transcription plays an active role in long-range gene activation. *Mol. Cell* 23, 365–375
28 West, A.G. and Fraser, P. (2005) Remote control of gene transcription. *Hum. Mol. Genet.* 14, R101–R111
29 Drissen, R. *et al.* (2004) The active spatial organization of the β-globin locus requires the transcription factor EKLF. *Genes Dev.* 18, 2485–2490
30 Vakoc, C.R. *et al.* (2005) Proximity among distant regulatory elements at the β-globin locus requires GATA-1 and FOG-1. *Mol. Cell* 17, 453–462
31 Gall, J.G. and Murphy, C. (1998) Assembly of lampbrush chromosomes from sperm chromatin. *Mol. Biol. Cell* 9, 733–747

32 Haaf, T. and Ward, D.C. (1996) Inhibition of RNA polymerase II transcription causes chromatin decondensation, loss of nucleolar structure, and dispersion of chromosomal domains. *Exp. Cell Res.* 224, 163–173

33 Chubb, J.R. *et al.* (2002) Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr. Biol.* 12, 439–445

34 Donze, D. and Kamakaka, R.T. (2002) Braking the silence: how heterochromatic gene repression is stopped in its tracks. *Bioessays* 24, 344–349

35 Scott, K.C. *et al.* (2006) A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. *Curr. Biol.* 16, 119–129

36 Burgess-Beusse, B. *et al.* (2002) The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl. Acad. Sci. U. S. A.* 99 (Suppl. 4), 16433–16437

37 Sutter, N.B. *et al.* (2003) Chromatin insulation by a transcriptional activator. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1105–1110

38 Cook, P.R. (1995) A chromomeric model for nuclear and chromosome structure. *J. Cell Sci.* 108, 2927–2935

39 Cook, P.R. (1999) The organization of replication and transcription. *Science* 284, 1790–1795

40 Cook, P.R. (2002) Predicting three-dimensional genome structure from transcriptional activity. *Nat. Genet.* 32, 347–352

41 Faro-Trindade, I. and Cook, P.R. (2006) A conserved organization of transcription during embryonic stem cell differentiation and in cells with high C value. *Mol. Biol. Cell* 17, 2910–2920

42 Faro-Trindade, I. and Cook, P.R. (2006) Transcription factories: structures conserved during differentiation and evolution. *Biochem. Soc. Trans.* 34, 1133–1137

43 Jackson, D.A. *et al.* (1993) Visualization of focal sites of transcription within human nuclei. *EMBO J.* 12, 1059–1065

44 Bartlett, J. *et al.* (2006) Specialized transcription factories. *Biochem. Soc. Symp.* 73, 67–75

45 Pombo, A. *et al.* (1999) Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J.* 18, 2241–2253

46 Schafer, D.A. *et al.* (1991) Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature* 352, 444–448

47 Cook, P.R. and Gove, F. (1992) Transcription by an immobilized RNA polymerase from bacteriophage T7 and the topology of transcription. *Nucleic Acids Res.* 20, 3591–3598

48 Harada, Y. *et al.* (2001) Direct observation of DNA rotation during transcription by *Escherichia coli* RNA polymerase. *Nature* 409, 113–115

49 Snow, M.H.L. and Callan, H.G. (1969) Evidence for a polarized movement of the lateral loops of newt lampbrush chromosomes during oogenesis. *J. Cell Sci.* 5, 1–25

50 Cheng, J. *et al.* (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 308, 1149–1154

51 Martienssen, R.A. *et al.* (2005) RNA interference and heterochromatin in the fission yeast *Schizosaccharomyces pombe*. *Trends Genet.* 21, 450–456

52 Abranches, R. *et al.* (1998) Transcription sites are not correlated with chromosome territories in wheat nuclei. *J. Cell Biol.* 143, 5–12

53 Branco, M.R. and Pombo, A. (2006) Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.*4, e138 DOI: 10.1371/journal.pbio.0040138 (biology.plosjournals.org)

54 Wang, J. *et al.* (2004) Promyelocytic leukemia nuclear bodies associate with transcriptionally active genomic regions. *J. Cell Biol.* 164, 515–526

55 Brown, J.M. *et al.* (2006) Coregulated human globin genes are frequently in spatial proximity when active. *J. Cell Biol.* 172, 177–187

56 Bon, M. *et al.* (2006) Modeling a self-avoiding chromatin loop: relation to the packing problem, action-at-a-distance, and nuclear context. *Structure* 14, 197–204

57 Capodieci, P. *et al.* (2005) Gene expression profiling in single cells within tissue. *Nat. Methods* 2, 663–665

58 Egner, A. and Hell, S.W. (2005) Fluorescence microscopy with super-resolved optical sections. *Trends Cell Biol.* 15, 207–215

59 Goldman, R.D. *et al.* (2005) Nuclear lamins: building blocks of nuclear structure and function. *Novartis Found. Symp.* 264, 3–16

60 Visa, N. (2005) Actin in transcription: actin is required for transcription by all three RNA polymerases in the eukaryotic cell nucleus. *EMBO Rep.* 6, 218–219

61 Cook, P.R. (2001) *Principles of Nuclear Structure and Function,* John Wiley & Sons

62 Miller, O.L. and Bakken, A.H. (1972) Morphological studies of transcription. *Acta Endocrinol. Suppl. (Copenh.)* 168, 155–177

63 Jackson, D.A. *et al.* (1998) Numbers and organization of RNA polymerases, nascent transcripts and transcription units in HeLa nuclei. *Mol. Biol. Cell* 9, 1523–1536

64 Rippe, K. (2001) Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.* 26, 733–740

65 Saiz, L. and Vilar, J.M. (2006) DNA looping: the consequences and its control. *Curr. Opin. Struct. Biol.* 16, 344–350

66 Phair, R.D. *et al.* (2004) Global nature of dynamic protein–chromatin interactions *in vivo*: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol. Cell. Biol.* 24, 6393–6402

67 Marenduzzo, D. *et al.* (2006) Entropy-driven genome organization. *Biophys. J.* 90, 3712–3721

68 Minton, A.P. (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* 256, 10577–10580

69 Schneider, D.A. and Nomura, M. (2004) RNA polymerase I remains intact without subunit exchange through multiple rounds of transcription in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 15112–15117

70 Law, A. *et al.* (1998) Direct cloning of DNA that interacts *in vivo* with a specific protein: application to RNA polymerase II and sites of pausing in *Drosophila*. *Nucleic Acids Res.* 26, 919–924

The Genetics Society, BSDB and BSCB joint Spring meeting

29 March to 1 April 2007

Sessions include

Protein modification; Cell growth; Ubiquitin; Biological clocks; Nuclear dynamics; Genetics of behaviour; Genomes, chromosomes and disease; Cell polarity and migration; and Systems biology

Featuring The Genetics Society Medal Lecture 2007, by Caroline Dean
Poster Prize sponsored by *Trends in Genetics*

For further details see http://www.genetics.org.uk