Gene Ther Mol Biol Vol 1, 495-508. March, 1998.

Dedicated sites of gene expression in the nuclei of mammalian cells.

Fransisco Iborra, Ana Pombo and Dean A Jackson.

CRC Nuclear Structure and Function Research Group, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, OX1 3RE, UK.

Correspondence to: Dean. A. Jackson, Tel: +44-1865-275527, Fax: +44-1865-275501, E-Mail: Jackson@Path.OX.AC.UK.

Key Words: Nuclear compartments, chromatin domains, transcription factories, nuclear structure.

Summary

Establishing sites of transcription in the nuclei of higher eukaryotic cells is a very complex process. Before transcription can begin, a series of transcription factors must associate with their recognition motifs, within promoters and more remote activating sequences. Once bound, these factors and associated proteins are believed to form a complex that positions the RNA polymerase holoenzyme so that transcription can commence. As a consequence, active genes assume a specialized chromatin state across regions that define functional domains. Global nuclear architecture appears to stabilize these active domains by providing local environments dedicated to gene expression. As the spatial organization of these sites is unaffected by the removal of most chromatin they must be associated with a structural network. This nucleoskeleton, the associated transcription 'factories' and chromatin loops that arise as DNA binds proteins within factories are fundamental features of nuclear structure in higher eukaryotes. We argue that concentrating proteins needed to perform different steps of RNA synthesis within specialized nuclear compartments will be important in orchestrating events required for efficient gene expression.

I. Introduction

Over recent years, remarkable progress has been made in understanding a number of extremely complex functions performed by our genetic material - DNA. For example, we now have a reasonably clear picture of the basic elements required to initiate gene expression and understand the principles - if not the details - that control gene expression in different cell types. It is clear how different sequence motifs in DNA operate as binding sites to position specialized expression activating 'transcription factors' within gene promoters and how different, though often related, motifs might be located within more distant 'enhancer' elements. In some cases, even more remote 'locus control region' (LCR) elements have been show to exert a dominant effect in establishing chromatin domains competent for gene expression.

To state matters rather simply, the fundamental process of gene expression involves the combined action of protein factors bound to these different DNA motifs (Tjian and Maniatis, 1994). Proteins bound at the different sites are then though to associate forming a tertiary complex which, in the presence of secondary transcription factors acting as adaptors, provides a protein surface that first interacts with RNA polymerase and then positions the polymerase on the promoter (Goodrich et al., 1996) so that transcription can begin (Zawel and Reinberg, 1995; Aso et al., 1995). Though the complexity of factors involved in this process appears daunting, it now seems likely that this process will be simplified by the use of preformed sub-assemblies. For example, in addition to the synthetic machinery, RNA polymerase II complexes can contain elongation factors, RNA processing components, enzymes required for DNA repair and chromatin remodelling proteins (Koleske and Young, 1995; Maldonado et al., 1996; McCracken et al., 1997; Kim et al., 1997).

In some genes, promoter-bound transcription factors are sufficient to direct efficient expression. Often, however, remote enhancer sequences are required to establish appropriate levels of gene expression. Like

promoters, enhancers contain complex arrays of factor recognition motifs that bind appropriate factors. Complex that includes factors bound within both promoters and enhancers probably contribute to the activation process. While there are many possible mechanisms for enhancer function the fact that they work when linked to promoters on catenated DNA molecules, suggests that protein complexes assembled on the promoter and enhancer cooperate during the activation process. Enhancers with different efficiencies have been shown to drive uniform rates of transcription from active genes (Boyes and Felsenfeld, 1996; Osheim et al., 1996); differences in transcription rate reflect the number of active genes in a population and not different polymerase densities on individual genes. A similar mode of action has been proposed for locus control regions (Milot et al., 1996). Promoters appear to determine the rate of initiation while these distal sequence motifs control switching between active and inactive states.

Activating expression in individual cells is only part of an even more complex story. Mammalian genomes are estimated to have some 75,000 different genes. Only a minority of these are expressed in different cells. In rough terms, ~1/3rd genes perform house-keeping functions, ~1/3rd specialized functions in ~250 different cell types throughout the body and the remaining ~1/3rd specialized functions in the brain. The activity of highly active genes in expressing tissues and the same gene in non-expressing tissues can vary by up to 10^8 fold. This remarkable difference in levels emphasises the stability of the inactive and active states and confirms the efficiency of mechanisms that define regions of the genome that are competent for gene expression.

II. Gene expression and chromosomal position effects

Factors that influence gene expression are clearly very complex. Even when different genes are introduced together into cells and expressed transiently from plasmids their activity can be influenced drammatically by factors such as their spatial organization (Emerman and Temin, 1984). When genes are integrated into the genome 'position effects' add an additional complexity that commonly results in the eventual extinction of the ectopic gene (Palmiter and Brinster, 1986). Some genomic sites are non-permissive for expression. This phenomenon was first characterized in *Drosophila* when it was observed that genes translocated close to heterochromatin were commonly switched off (Singh, 1994, for early references). The dominant suppressive properties of constitutive heterochromatin (e.g. centromeres) can spread over 1 Mbp or more; cell to cell variations in suppression result in 'position effect variegation'.

In *Drosophila*, heterochromatin protein 1 (HP1) and the related Polycomb (Pc) proteins maintain a repressed chromatin state. The critical chromatin organization modifier (chromo) domain identified in HP1 and Pc is found in many proteins, including human homologues. Pc is found in large, multi-subunit protein complexes (~5 MDa) and controls expression from homeotic genes during *Drosophila* development (Franke et al., 1992). Though chromatin status is clearly the target, Pc-repressed chromatin is not resistant to digestion by restriction endonucleases and so is not subjected to generalized condensation. Modification of chromatin stability (stable chromatin is likely to prevent transcription factors access), sequestration into an inactive nuclear compartment and reduced chromatin flexibility may account for these observations (McCall and Bender, 1996, for discussion). Importantly, this system provides a means of developing chromatin compartments or domains with no absolute requirement for chromatin binding to a nucleoskeleton.

III. Chromatin domains as units of gene expression

The structure of active and inactive chromatin is clearly different (Edmondson and Roth, 1996). Much inactive chromatin is condensed and forms heterochromatic chromatin clumps inside the cell (**Figure 1**). Active chromatin must be accessible to proteins involved in gene expression and is consequently relatively open or dispersed. The ease with which chromatin is cut by nucleases provides the best indicator of activity status (Wolffe, 1995). The most accessible, 'hypersensitive' sites highlight regions of functional importance - promoters, enhancers, locus control regions and sites of nuclear matrix attachment. Hypersensitive sites arise as a consequence of structural changes that appear when the repetitive nature of nucleosomal chromatin is disrupted by proteins such as transcription factors. These interactions are clearly stable during successive rounds of transcription but appear to be displaced during replication.

In addition, transcriptional status correlates with a generalized nuclease 'sensitivity' that results from an open - 10nm - chromatin fibre. However, as sensitivity often extends many kbp outside established expression units, it is clear that other factors must be involved, implying that remote sequences determine the boundaries of functional domains.

Changes in chromatin structure that accompany the transition from an inactive to active state are believed to be controlled by products of the *SWI/SNF* genes, first described in *Saccharomyces cerevisiae* (Kingston et al.,

Figure 1. Morphology of the mammalian cell nucleus.

Growing HeLa cells were processed by standard EM techniques and thin sections stained. Inside the nucleus (n), the most prominent feature is the nucleolus (nu), with many fibrillar centres (fc) surrounded by dense fibrillar components (dfc) and a dispersed granular component (gc). The nucleoplasm contains dense patches of condensed chromatin (hc) that stand out against the relatively amorphous nuclear interior. The nucleus is separated from the organelle-rich cytoplasm (c) by a nuclear membrane (nm). Nuclear pores allow molecules to pass between the nucleus and cytoplasm. Bar, 1 μm.

1996). The *SWI* genes were shown to be required for mating type switching and subsequently, to form a complex capable of disrupting chromatin structure. Interestingly, the SWI/SNF complex has been shown to co-purify with the yeast RNA polymerase II holoenzyme, suggesting that it might operate to disrupt euchromatin, during RNA synthesis (Koleske and Young, 1995)

IV. Nuclear compartmentalization

The influence of gene position on expressional status emphasises the importance of higher-order chromatin structure in establishing patterns of gene expression. Protecting such structures in vivo, through successive rounds of transcription and replication, will be crucial to the maintenance of a cell's expression programme.

The cytoplasm of a typical mammalian cell is highly structured, containing many classes of membranated organelles with specialized roles. The nucleus contains no equivalent structures and appears ill-organized, in comparison. A section of a human cell emphasizes this impression (**Figure 1**). The nuclear membrane, nucleolus and regions rich in condensed heterochromatin stand out in contrast to the relatively amorphous nuclear interior. Within the interior, specialized staining techniques allow different features to be recognized (Monneron and Bernard, 1969). EDTA regressive staining, for example, confirms that roughly half of this region is occupied by dispersed chromatin. The remaining, interchromatin space, is rich in hnRNPs and contains characteristic structures such as perichromatin fibrils, perichromatin granules, interchromatin granules and interchromatin granule clusters. Though the precise functions of these structures remain unresolved, they are believed to play different roles in RNA metabolism (Fakan, 1994).

Recently, the use of immuno-staining as a routine analytical tool has emphasized the structural complexity within eukaryotic nuclei (Spector, 1993; Strouboulis and Wolffe, 1996; Jackson and Cook, 1996). Autoantibodies to protein components of snRNPs were first used to demonstrate that these proteins concentrate in 20-50 nuclear speckles in mammalian cells. These major sites were later shown to be inter-connected by a fibro-granular 'network' of minor sites (**Figure 2**). Like splicing components, many other proteins involved in gene expres-

Figure 2. Nuclear compartments rich in splicing proteins and nascent transcripts.

HeLa cells were permeabilized with saponin in an isotonic buffer and sites of RNA synthesis labelled for 15 min with Br-UTP (A-D) or biotin-CTP (E-H). Sm (B) or SC35 (F) antigens and Br-RNA (C) or biotin-RNA (G) were indirectly immunolabelled and DNA stained with TOTO-3 (A,E). Optical sections (~700 nm) show a classical distribution of Sm and SC35 antigens, with major sites 'speckles' and dispersed minor foci; nucleoli are blank. In the nucleoplasm, most sites of transcription lie adjacent to the minor foci (D and H, merges of B,C and F,G respectively). Note that in (G), anti-biotin antibodies label an extensive mitochondrial network. This indicated the preservation of cellular structure under conditions used. Loss of structure generally correlates with the appearance of mitochondrial transcription, not seen here (C). Bar, 2.5 µm. See Pombo and Cook (1996) for details.

sion give intriguingly punctate staining patterns (van Driel et al., 1996). Understandably, such images are assumed to demonstrate that different proteins accumulate at functionally important sites. In some cases, however, the complexity of staining is greater than anticipated, implying that proteins might accumulate at non-functional, storage, sites.

V. Sites of transcription

In some cases, the inability of antibodies to proteins involved in different nuclear functions to distinguish active and inactive sites can be overcome by labelling active sites directly. For example, sites of transcription in mammalian cells can be labelled using ^{3}[H]-uridine (Fakan and Puvion, 1980). High resolution analyses have identified perichromatin fibrils as the features most closely associated with transcription sites. Under these conditions, however, scatter of emitted irradiation (50% of gains lie more than 200 nm from their source using standard EM autoradiography) prevents the unambiguous identification of transcription sites. Recently, resolution has been improved by labelling transcription sites in permeabilized cells using modified RNA precursors (Br-UTP or biotin-CTP); under these condition it is a simple matter to control the rate of transcription and so ensure that only sites of synthesis are labelled. Cells labelled with Br-UTP and examined by light microscopy, after immuno-labelling sites of incorporation (**Figure 2**), demonstrate that the majority of labelled transcripts are concentrated within a limited number of nuclear sites, and not diffusely spread throughout chromatin (Jackson et al., 1993: Wansink et al., 1993). Labelling under *in vivo* conditions (micro-injected Br-UTP or cells grown in medium supplemented with Br-U) gives the same impression if short labelling periods (~5 min) are used (Fay et al., 1997); after longer labelling intervals staining patterns are complicated by the presence of labelled RNA in transit to the cytoplasm.

The ability to label sites of transcription in this way allows the active sites of RNA synthesis to be compared with the organization of other components involved in gene expression, by immuo-fluorescence (van Driel et al., 1996; Huang and Spector, 1996; Pombo and Cook, 1996; Grande et al., 1997).

Figure 3. Genes in action.

Chromatin templates and associated transcripts can be visualized after 'spreading' nuclear contents. Such 'Miller spreads' are commonly prepared from amphibian oocytes where the nuclear membrane can be removed manually and nuclear contents dispersed in water. As mammalian nuclei cannot be manipulated in this way, HeLa cells (A,B) must be dispersed in a 0.33% solution of a commercial detergent (Joy), during spreading. Ribosomal RNA genes usually appear clustered and are rich in nascent transcripts (A). In contrast, extra-nucleolar transcription unit, in dispersed chromatin, have very few associated transcripts (B). Bar, 1 μm. See Miller and Bakken (1972) for details.

VI. The nucleolus - a dedicated site of ribosomal RNA biosynthesis

The specialized site of transcription that has been studied most intensively in mammalian cells is the nucleolus (Fischer et al., 1991; Shaw and Jordan, 1995). Each diploid human cell has an estimated 300-400 copies of the gene needed to make ribosomal RNA. These are located on chromosomes 13, 14, 15, 21 and 22 where the ~40 kbp repeats are grouped into clusters usually with 3-5 genes in each unit. At any time, ~1/3 of the rRNA genes are active. These are expressed within specialized sub-nuclear organelles called nucleoli (**Figure 1**). Most diploid human cells have 1 or 2 nucleoli but cells in culture can have more; nucleolar morphology is a marker for growth status and can be used as one indicator in the diagnosis of malignancy.

The anatomy of nucleoli is well documented (**Figure 1**). Fibrillar centres (fcs) are rich in the synthetic machinery (RNA polymerase I). Fcs are surround by a zone called the dense fibrillar component (dfc) where the nascent transcripts accumulate. An individual fc is probably coated with the genes and nascent products from a single active rDNA cluster (3-5 genes); in cross section, the products of a single gene appear as densely staining areas measuring ~200 nm across (**Figure 1**). As the nascent transcripts mature they pass into the granular component (gc) where they become associated with the ribosomal proteins.

The high demand for ribosomal RNA molecules means that these genes are the most active in human cells - a proliferating cell must produce ~$5x10^6$ ribosomes during each cell cycle of ~24 hours. Nucleolar structure ensures that the components required for RNA production are organized to optimize efficiency. This is borne out by the appearance of the active genes and their associated nascent transcripts (**Figure 3**). When nucleoli are disrupted and spread, each active gene is seen to have between 100-150 engaged RNAs. The polymerases are loaded with such efficiency (~1/120 bp) that the active transcription units remain devoid of histones (Miller and Bakken, 1972).

VII. Specialized sites of transcription in the nucleoplasm

While early experiments analysed by LM supported the existence of nucleoplasmic transcription centres more detailed studies were required to assess if nucleoli are a paradigm for transcription site organization (Hozák et al.,

Figure 4. Distribution of transcription sites.

HeLa cells were grown for 5 min in medium supplemented with Br-uridine. EM sections were prepared and sites of RNA synthesis immunolabelled with 9 nm gold particles. Within the nucleus (n), most particles lie in clusters (arrowheads). The bulk of the nuclear interior, nuclear membrane (nm) and cytoplasm (c) are unlabelled. Parts of two nucleoli (nu) are seen in this section. Bar, 250 nm. See Iborra et al. (1996) for details.

1994). Electron microscopy of HeLa cells labelled with RNA precursor analogues (Iborra et al., 1996) both in vivo and in vitro has shown a typical cell to have ~2500 distinct transcription compartments (**Figure 4**), with labelled zones measuring 50-150 nm (mean is 80 nm) across. The labelled sites often lie along the borders of nuclear regions rich in condensed chromatin and usually lie adjacent to or are surrounded by 'clouds' of dispersed chromatin (**Figure 5B**). The nascent sites also contain the synthetic machinery (**Figure 5A**) which quite often appears concentrated towards one sub-region of the labelled zone. The sites also contain transcription factors, splicing factors and others proteins involved in RNA processing. Transcription sites often lie close to, but never within, interchromatin granule clusters (**Figure 5C**) that are rich in many splicing factors and appear as 'speckles' by LM when stained with antibodies to splicing components such as Sm and SC35 proteins (**Figure 2**).

The main difference between nucleolar and nucleoplasmic transcription sites concerns the organization of the active genes. Like nucleoli, nucleoplasmic transcription sites must contain many active polymerases. This can be inferred from the number of sites and estimates of the number of active polymerases and genes in a cell at any moment (Cox, 1976). To confirm this crucial point, we have shown recently (Jackson et al., submitted) that HeLa cells with ~2,500 nucleoplasmic sites have ~75,000 active nucleoplasmic RNA polymerases. Unlike nucleoli, however, genes transcribed in the nucleoplasm are known to have rather low densities of associated transcripts. When HeLa nuclei are disrupted by hypotonic treatment and spreading dispersed chromatin fibres rarely have more than a few putative transcripts (**Figure 3**). Even cells infected with adenovirus, at their peak of transcription, have only 1 transcript/7.5 kbp adenovirus DNA (Beyer et al., 1981; Wolgemuth and Hsu, 1981). Using HeLa cells disrupted with sarkosyl - under conditions that retain all engaged RNA polymerases - we have confirmed that most active genes have very few (usually 1-3) transcripts. This suggests, perhaps remarkably, that an average transcription site contains the machinery to simultaneously synthesise and process 30 transcripts associated with ~20 different genes. Because the nascent transcripts only occupy 0.5% of a HeLa cell nucleus (Iborra et al., 1996) these cannot contain all the active chromatin. It seems probable, therefore, that active genes will surround the synthetic sites, as indicated above (**Figure 5B**). Note, however, that despite their small size and evident complexity, the nucleoplasmic sites have less than half the RNA density of equivalent sites in nucleoli.

As many genes appear to be transcribed from individual sites, that are also capable of performing a range of downstream processing events, we have called these sites transcription 'factories'. In view of the complexity of these factories, it is interesting to consider whether different genes with common requirements (e.g. activating transcription factors) might accumulate at individual sites. Though the chromosomal arrangement of genes will be dominant in determining the composition of individual transcription sites, circumstances could arise where related genes, on different chromosomes, are transcribed within the same compartment. Preliminary experiments have hinted that this might be so. For example, we (AP and DAJ) have shown that a fraction (~1/10th) of nucleoplasmic

Figure 5. The architecture of transcription sites.

HeLa cells were permeabilized and sites of RNA synthesis labelled with biotin-CTP for 15 minutes. Sites of synthesis were immunolabelled with 9 nm gold (A-C). Most sites also contained RNA polymerase, immunolabelled with 15 nm gold (A); note that the 9 and 15 nm gold particles are generally sub-compartmentalized within transcription sites. Transcription sites commonly appeared to be surrounded a halo of chromatin clouds, visualized after EDTA regressive staining (B). Bismuth binds phosphoproteins (such as RNA polymerase II) and stains transcriptions sites (C), though adjacent regions were often unstained. Interchromatin granule clusters stained with bismuth but never contained labelled transcripts. Bar, 100 nm. See Iborra et al. (1996) for details.

transcription sites contain predominantly (if not exclusively) RNA polymerase III transcription units. In addition, we (AP) have characterized the organization of a number of genes that are activated by the transcription factors PTF and OCT-1 in association with a PTF-rich sub-nuclear compartment.

VIII. The spatial organization of transcription sites

Figure 6. The nucleoskeleton.

Encapsulated HeLa cells were permeabilized, chromatin cut with nucleases, ~90% DNA removed and a 500 nm resinless section prepared. Agarose (A) surrounds the cytoplasmic (C) and nuclear remnants that are separated by the nuclear

Another remarkable feature of the structure of transcription factories is that the spatial organization of labelled nascent RNA persist in nuclei when almost all chromatin is removed (**Figure 6**; Jackson et al., 1993; Wansink et al., 1993: Iborra et al 1996). The nascent RNA, unlike most mRNA en route to the nuclear periphery (Verheijen et al., 1988), is tightly associated with the nuclear matrix (**Figure 7**), confirming that interactions within transcription factories are stable under a range of conditions, independently of chromatin. This emphasises the important point that the transcription factories are structures in their own right and do not arise passively as a consequence of the organization of other nuclear components.

This view of global nuclear organization supports the idea that active RNA polymerases are 'fixed' and is incompatible with transcription complexes that track along the chromatin during RNA synthesis. In addition, transcription factors, polymerases and other proteins involved in gene expression will perform critical structural roles, binding chromatin at functionally important sites that become spatially restricted through their indirect association with the nucleoskeleton. While such an arrangement will have profound organizational consequences in higher eukaryotes, it is worth noting that such an arrangement has been shown to exist within virus particles (Prasad et al., 1996).

The *in vivo* organization of transcription sites has also been analyzed using an approach that determines the spatial distribution of active genes and their products using fluorescent *in situ* hybridization (FISH). Though the processing required for hybridization leads to loss of morphological detail, this approach has provided data on the relative organization of active genes and local processing compartments (Xing et al., 1993) and has suggested interesting mechanisms by which mature mRNA molecules might be transported to the cytoplasm (Rosbach and Singer, 1993).

IX. Sites of pre-mRNA splicing

While it is clear that components involved in major nuclear functions are compartmentalized, the extent to which these compartments correlate with sites of function

lamina (L). The nucleoskeleton, a diffuse network of coated filaments (arrowheads) connects regions of the nuclear interior such as the nucleolus (NU) replication factories (F) and many dense sites (D), some of which arise from transcription sites. Bar, 1 µm. See Jackson and Cook (1988) for details.

is not always easy to assess. This is most apparent when function is difficult to measure directly. Understanding the organization of sites of pre-mRNA splicing serves to emphasise this point (**Figure 2**). A series of sophisticated analyses have shown that many genes lie close to the splicing speckles (Moen et al., 1995). As in situ hybridization shows these sites to be rich in polyA+ nuclear RNA and much, but not all, splicing occurs co-transcriptionally it appears that the speckles represent predominant sites of active splicing. Other observations refute this view. First, cells grown in ^{3}[H]uridine for short times show little or no labelling within IGCs/speckles; even after long incorporations these structures remain poorly labelled (Puvion and Puvion-Dutilleul, 1996). Second, high resolution analyses of nascent transcripts confirm that while many lie close to IGCs no transcription occurs within them (**Figure 5**; Iborra et al., 1996; Puvion and Puvion-Dutilleul, 1996; Pombo and Cook, 1996). Finally, *in situ* hybridization techniques designed to analyze the sites of splicing indicate that speckles are not the predominant active sites (Zhang et al., 1994).

Possible explanations of these controversies range from the obvious limitations, technical capabilities and resolution of different approaches used to cell-specific differences. Perhaps these observations indicate that the major (speckles) and minor sites represent 2 parts of the same population that is able to respond to demand - is dynamic - performing splicing as and when required. The major sites might represent regions of particularly high demand/turnover but could also be sites of storage and/or assembly that are subsequently dispatched elsewhere. The dynamic features of these compartments have been assessed using an essential splicing factor (SF2/ASF) tagged with green fluorescent protein Misteli et al., 1997).

X. RNA transport pathways

In the majority of cases, activities present within transcription factories, adjacent to the site of RNA synthesis, will ensure that mature mRNA molecules leave factories. Though the majority of splicing occurs at the time of transcription, some RNAs containing introns may move into the downstream transport pathways. Details of events that control RNA transport and export are complex and have been reviewed extensively (Gorlich and Mattaj, 1996). It is not appropriate to discuss an extensive literature here, though we would like to consider one interesting aspect of transport that has arisen from our own studies.

It is usually assumed that nuclear RNAs and their associated proteins move inside the nucleus as independent entities (Dreyfuss et al., 1993). Movement is probably diffusion driven in inter-chromatin channels. A detailed analysis of large (200S) RNP particles has demonstrated that RNA molecules are associated with protein according to size: RNA from 1.5 to 35 kb is found in particles of similar mass (Spann et al., 1989; Sperling et al., 1997). This observation is believed to reflect the high protein

Figure 7. Transcription at the nuclear matrix.

HeLa cells with nascent transcript containing Br-UMP (**Figure 4**) were extracted with 2 M NaCl and Br-RNA immunolabelled with 5 nm gold particles. Samples were then embedded and sections prepared (A,B). A typical nuclear region (A) shows numerous dense areas separated by the fibro-granular 'nuclear matrix'. Many of the dense areas are rich in Br-RNA (B; higher magnification of area indicated in A); these are clearly remnant transcription sites. Note that immunolabelling before embedding gives much higher particle densities because label is not restricted to the detection of antigens on a section surface. This increases sensitivity but does not alter the number of transcription sites seen. Bars, 100 nm.

content of the particles, that contain a modular structure thought to be a remnant of earlier splicing events. These particles are stable even though ~95% RNA present is fully spliced mRNA.

In our analysis of the movement of nuclear RNA labelled with bromouridine in vivo, we have shown that the majority of RNA transport occurs in association with structures equivalent to these 200S particles. The particles contain Br-RNA and are rich in SR-proteins, a family of proteins involved in splicing (Manley and Tacke, 1996). Interestingly, the size distribution and numbers of labelled particles in transit suggest that each contains many mRNAs.

XI. Global nuclear structure

Many lines of evidence indicate the functional importance of global nuclear structure. We know, for example, that chromosomes occupy discrete nuclear domains during interphase and that these often assume preferred though never precise nuclear locations (Cremer et al., 1993). Preferred orientations are not uncommon (probably a vestige of mitosis) and may differ at different stages of the cell cycle; even at this level nuclear structure must be dynamic (Ferguson and Ward, 1992). Though chromosomes in interphase are decondensed relative to their familiar mitotic counterparts, it is clear from *in situ* hybridization analyses that individual domains remain locally restricted and that the linear arrangement of genes, together with local structural features, closely reflect those seen in mitosis (Yokota et al., 1995). Apparently dramatic differences in the organization of chromosomes during mitosis and corresponding structures during interphase probably reflect the dispersal of chromatin domains through elongation of the chromosome axis, with relatively minor changes in chromatin condensation. As the nucleus reforms after mitosis, chromosome decondensation will generate inter-chromatin channels, both between and within chromosomes, allowing access of different nuclear components to all functionally important parts of the genome.

XII. The nuclear matrix and nucleoskeleton

By their very nature, nuclear functions such as transcription and replication must be dynamic processes. It was surprising, therefore, when the products of these processes were shown to retain their spatial organization in extracted nuclei from which almost all chromatin had been removed. To explain such observations, it was argued that the active enzymes were associated with and organized by a structural 'nuclear matrix'. Subsequently, the nuclear matrix of extracted cells has been shown to participate in almost all aspects of nuclear function, in higher eukaryotes (reviewed in Berezney et al., 1996; Nickerson et al., 1996; Stein et al., 1996).

A detailed analysis of functional sites in cells extracted under 'physiological' conditions (Jackson and Cook, 1996) has ruled out the possibility that this organization arises during matrix preparation and confirmed that active transcription sites maintain their spatial disposition when almost all chromatin, representing about 40% of nuclear mass, is removed. This crucial observation shows that the active sites are intimately associated with an organizational solid phase, probably the major structural feature within eukaryotic nuclei (**Figures 6** and **7**). At the heart of this structure lies a network of intermediate filament-like core filaments that provides important structural continuity by connecting sites of functional importance to the nuclear periphery. *In situ*, this structure is coated and stabilized by hnRNP complexes, many of which will be *en route* to the cytoplasm. Though structural, this nucleoskeleton must be dynamic - dramatic changes accompany mitosis and more subtle ones result from changes in growth.

XIII. Chromatin domains and loops

The nucleoskeleton and functional compartments that it binds are fundamental features of internal nuclear architecture. As these structures are reasonably stable they are likely to be a major source of protein-DNA interactions that form DNA loops and chromatin domains

inside the nucleus. Once again a complex literature addresses this issue (Laemmli et al., 1992; Bode et al., 1996).

In brief, scaffold and matrix attached sequences (SARs and MARs) are commonly AT-rich sequences with many topoisomerase II binding motifs. Various other proteins have been shown to bind strong S/MAR DNA sites. It is not clear, however, how the organization of these extracted structures reflects that existing inside the cell (Jacks and Eggert, 1992). In contrast, in domains analyzed under isotonic conditions, most sequences responsible for binding chromatin to the nucleoskeleton were of functional importance (Jackson and Cook, 1993; Jackson et al., 1996). It appears that different extraction protocols must accentuate different classes of interactions existing *in vivo*.

Probable chromatin domain boundaries have been characterized in the fruit fly, *Drosophila*. 'Specialized chromatin structures' (scs) flanking two heat shock genes (HSP70) were shown to correlate with the boundaries of an ~15 kbp active domain - at 87A7 on polytene chromosomes - following heat treatment (Schedl and Grosveld, 1995). These elements contain pairs of very strong DNAse hypersensitive sites flanking a nuclease resistant sequence of ~300 bp. Boundary element attachment factors and their recognition motifs within the scs may be important determinants of chromosome structure (Zhao et al., 1995; Strick and Laemmli, 1995).

XIV. Global organization and dedicated sites of gene expression

A priori, we might assume that genes expressed from their natural chromosomal sites operate with the required efficiency. Wide variations in the expression of identical ectopic sequences, at different chromosomal sites, implies that local factors can influence gene expression (Allen et al., 1988; Bonnerot et al., 1990). Different factors must contribute to the range of expression seen. It is clear how chromosomal position, through the suppressive influence of local heterochromatin, might extinguish expression. Partial activities, in contrast, must reflect transcription from sites with different combination of factors that together determine the frequency with which transcription occurs.

Other lines of evidence imply that nuclear organization can influence gene function. For example, RNA molecules transcribed from chimeric genes - with RNA polymerase III promoters and polymerase II coding sequences - introduced into mammalian tissue culture cells by transfection, produced normal pre-mRNA but were unable to splice or form mature polyA+ tails and hence mature mRNA (Sisodia et al., 1987). This suggests that splicing and polyadenylation pathways are coupled to transcription by RNA polymerase II and that different pathways are spatially independent.

Rather surprising observations suggest that different steps in the expression pathway cooperate to produce mature mRNA. For example, the _-globin promoter was shown to drive synthesis of transcription from an intron-containing gene but not the same sequence from which the non-coding DNA had been removed (Neuberger and Williams, 1988). Remarkably, the same intron-less transcription unit was expressed from CMV or heat-shock promoters.

Such observations suggest that nuclei are assembled so that the local organization of components required for gene expression will reproduce a level of activity that reflects that organization. In addition to establishing the desired level of synthesis, this nuclear 'set-up' appears to direct the transcription product onto an appropriate pathway that couples synthesis to the desired combination of post-synthetic events - RNA processing, export and perhaps even cytoplasmic location and function.

XV. Conclusion

It is now clear that different layers of organization contribute to the complex processes required for gene expression. Simple recognition motifs in chromatin first bind transcription factors and set the activation process in motion. Multi-component pre-initiation complexes then facilitate the binding of the RNA polymerase complex to the promoter (Goodrich, et al., 1996) so that elongation can proceed (Aso et al., 1995; Zawel and Reinberg, 1995). Intellectually, these steps are easy to describe in molecular detail.

These processes must occur in the context of systems established to maintain active and inactive chromatin states. Though we know surprisingly little about the signals that define chromatin domains it is clear the certain LCR elements can 'open' chromatin and establish domains that are permissive for transcription (Schedl and Grosveld, 1995). The partial activation of expression from an incomplete LCR together with equivalent experiments on enhancer elements are consistent with the view that these remote elements make vital contributions to the active promoter complex (Wijgerde et al., 1995). The behaviour of genes with incomplete LCRs, situated close to heterochromatin, indicates that the complete LCR maintains transcription activity by ensuring that transcription occurs in all cells, at all times, but does not directly control the rate of transcription (Milot et al., 1996).

In vivo, additional organizational features appear to influence gene expression. Over recent years a significant body of literature has described the compartmentalization of different nuclear functions. Rather than being dispersed uniformly throughout the nuclear interior, different

functions required for expression occur at specialized, dedicated sites. In mammalian cells the density of transcription sites is far lower than expected from estimates of active transcripts and transcription unit complexity, implying than multiple transcription units are active in each site. Morphological analyses together with immuno-staining indicate that these sites have a zonal organization, with different regions performing specialized roles. Furthermore, as the transcript containing regions occupy only 0.5% of the volume of a HeLa nucleus this arrangement is inconsistent with the view that transcripts will be uniformly dispersed throughout euchromatin (occupying ~10% of the nucleus) bound to tracking RNA polymerases. Images like **Figure 5B** support the idea that many chromatin clouds are served by a single, active transcription compartment.

This ordered view of transcript synthesis and maturation has interesting functional consequences. Before any genes can be expressed, transcription factor complexes must assemble on its promoter (Goodrich et al., 1996). It is commonly assumed that these factors scan chromatin for potential binding sites. The observation that significant fractions of many factors are nuclear matrix associated complicates this view, implying that factors are part of some nuclear 'solid phase'. Of course such impressions present a time averaged image of nuclear order and give no indication of either spatial or temporal dynamics. But even if factors can scan for binding sites, it may be that they then serve to deliver and subsequently confine genes to appropriate nuclear compartments. Once established, such complexes could have a major influence on higher-order chromatin structure, throughout interphase and mitosis.

Another advantage of dedicated nuclear compartments is that by concentrating components required to perform coupled functions in a limited number of sites, it will be possible to execute these functions with optimal efficiency. If components were released into a 'soluble' nucleoplasm between each activity, a dramatic fall in efficiency might be expected. It is clear that active chromatin domains must maintain their active configuration under circumstances where the majority of chromatin is inert. Though different mechanisms will tend to stabilise these forms, it is important that they are both sufficiently stable to maintain expression and sufficiently unstable to allow reprogramming. Interestingly, proteins capable of remodelling chromatin have been shown to form an integral part of the yeast RNA polymerase II holoenzyme (Koleske and Young, 1995). Restricting the spatial distribution of the complex in this way will tend to preserve existing chromatin states. In addition, compartments rich in transcription factors would also help to re-establish active sites following replication. Elongating DNA polymerases are known to displace certain transcription factor complexes and these must be re-established if expression is to be maintained. This must occur, however, in the presence of large histone pools that could alter promoter activity by competing for the same sites. Replicating active genes early in S-phase within compartments that contain the components required to re-establish an active chromatin configuration provides the best opportunity of protecting existing patterns of gene expression within eukaryotic cells.

Finally, it is worth considering how nuclear organization could influence the behaviour of particular genes in differentiated cells. With the exception of nucleoli, we known very little about the constellations of genes that contribute to individual transcription sites or whether individual sites can be totally or partially dedicated to the synthesis of particular classes of transcripts. While groups of genes on particular chromosomes will dominate local organization of transcription sites, some contribution from active genes on adjacent chromosomes might be expected. This could be especially prevalent in non-proliferative cells where long periods without the disruptive forces of the cell cycle might allow genes with similar activation requirements to occupy particular sites. Such an arrangement could influence the frequency of common translocations during the development of malignancy. Specialized transcription sites could also influence the performance of non-chromosomal genes. If cells have a limited number of sites that are competent for the synthesis of specific gene products, understanding how to access these sites would be important in gene therapy.

Acknowledgements

We thank the Cancer Research Campaign (DAJ), The Wellcome Trust (FI) and the Junta Nacional de Investigaç_o Científica e Tecnológica (Portugal, Program PRAXIS XXI; AP) for support.

References

Allen ND Cran DG Barton SC Hettle S Reik W Surani MA (**1988**) Transgenes as probes for active chromosomal domains in mouse development. **Nature** 333, 852-855.

Aso T Conaway JW and Conaway RC (**1995**) The RNA polymerase II elongation complex. **FASEB** 9, 1419-1428.

Berezney R Mortillaro MJ Ma H Wei X and Samarabandu J (**1996**) The nuclear matrix: a structural milieu for genomic function. **Int. Rev. Cytol.** 162A, 1-65.

Beyer AL Bouton AH Hodge LD and Miller OL (**1981**) Visualization of the major late R strand transcription unit of adenovirus serotype 2. **J. Mol. Biol.** 147, 269-295.

Bode J Schlake T Rios-Ramirez M Mielke C Stengert M Kay V and Klehr-Wirth D (**1996**) Scaffold/matrix-attached regions: structural properties creating transcriptionally active loci. **Int. Rev. Cytol.** 162A, 389-454.

Bonnerot C Grimber G Briand P and Nicolas J-F (**1990**) Patterns of expression of position-dependent transgenes in mouse embryo. **Proc. Natl. Acad. Sci. USA.** 87, 6331-6335.

Boyes J and Felsenfeld G (**1996**) Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. **The EMBO J.** 15, 2496-2507.

Cox RF (**1976**) Quantitation of elongating form A and B RNA polymerases in chick oviduct nuclei and effects of estradiol. **Cell** 7, 455-465.

Cremer T Kurz A Zirbel R Dietzel S Rinke B Schrock E Speicher MR Mathieu U Jauch A Emmerich P Scherthan H Ried T Cremer C and Lichter P (**1993**) Role of chromosome territories in the functional compartmentalization of the cell nucleus. **Cold Spring Harb. Symp. Quant. Biol.** 58, 777-792.

Dreyfuss G Matunis MJ Pinol-Roma S and Burd C (**1993**) hnRNP proteins and the biogenesis of mRNA. **Annu. Rev. Biochem.** 62, 289-321.

Edmondson DG and Roth SY (**1996**) Chromatin and transcription. **FASEB J.** 10, 1173-1182.

Emerman M and Temin H (**1984**) Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. **Cell** 39, 459-467.

Fakan S (**1994**) Perichromatin fibrils are *in situ* forms of nascent transcripts. **Trends Cell Biol.** 4, 86-90.

Fakan S and Puvion E (**1980**) The ultrastructural visualization of nucleolar and extranucleolar RNA synthesis and distribution. **Int. Rev. Cytol.** 65, 225-299.

Fay FS Taneja KL Shenoy S Lifshitz L and Singer RH (**1997**) Quantitative digital analysis of diffuse and concentrated nuclear distributions of nascent transcripts, SC35 and poly(A). **Exp. Cell Res.** 231, 27-37.

Ferguson M and Ward DC (**1992**) Cell cycle dependent chromosomal movement in pre-mitotic human T-lymphocyte nuclei. **Chromosoma** 101, 557-565.

Fischer, D Weisenberger D and Scheer U (**1991**) Assigning functions to nucleolar structures. **Chromosoma** 101, 133-140.

Franke A DeCamillis M Zink D Cheng N Brock HW and Paro R (**1992**) *Polycomb* and *polyhomeotic* are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. **The EMBO J**. 11, 2941-2950.

Goodrich JA Cutler G and Tjian R (**1996**) Contacts in context: promoter specificity and macromolecular interactions in transcription. **Cell** 84, 825-830.

Gorlich D and Mattaj IW (**1996**) Nucleocytoplasmic transport. **Science** 271, 1513-1518.

Grande MA van der Kraan I de Jong L and van Driel R (**1997**) Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. **J. Cell Sci.** 110, 1781-1791.

Hozák P Cook PR Schöfer C Mosgöller W and Wachtler F (**1994**) Site of transcription of ribosomal RNA and intranucleolar structure in HeLa cells. **J. Cell Sci.** 107, 639-648.

Huang S and Spector DL (**1996**) Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. **J. Cell Biol.** 133, 719-732.

Iborra FJ Pombo A Jackson DA and Cook PR (**1996**) Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. **J. Cell Sci.** 109, 1427-1436.

Jacks RS and Eggert H (**1992**) The elusive nuclear matrix. **Eur. J. Biochem.** 209, 503-509.

Jackson DA and Cook PR (**1988**) Visualization of a filamentous nucleoskeleton with a 23 nm axial repeat. **The EMBO J.** 7, 3667-3677.

Jackson DA and Cook PR (**1993**) Transcriptionally-active minichromosomes are attached transiently in nuclei through transcription units. **J. Cell Sci**. 105, 1143-1150.

Jackson DA and Cook PR (**1996**) The structural basis of nuclear function. **Int. Rev. Cytol.** 162A, 125-149.

Jackson DA Hassan AB Errington RJ and Cook PR (**1993**) Visualization of focal sites of transcription within human nuclei. **The EMBO J.** 12, 1059-1065.

Jackson DA Bartlett J and Cook PR (**1996**) Sequences attaching loops of nuclear and mitochondrial DNA to underlying structures in human cells: the role of transcription units. **Nucl. Acids Res.** 24, 1212-1219.

Kim E Du L Bregman DB and Warren SL (**1997**) Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. **J. Cell Biol.** 136, 19-28.

Kingston RE Bunker CA and Imbalzano AN (**1996**) Repression and activation by multiprotein complexes that alter chromatin structure. **Genes Dev.** 10, 905-920.

Koleske AJ and Young RA (**1995**) The RNA polymerase II holoenzyme and its implications for gene regulation. **Trends Biochem. Sci.** 20, 113-116.

Laemmli UK Kas E Poljak L and Adachi Y (**1992**) Scaffold-associated regions: *cis*-acting determinants of chromatin structural loops and functional domains. **Curr. Opin. Genet. Dev.** 2, 275-285.

Maldonado E Shiekhattar R Sheldon M Cho H Drapkin R Rickert P Lees E Anderson CW Linn S and Reinberg D (**1996**) A human RNA polymerase II complex associated with SRB and DNA-repair proteins. **Nature** 381, 86-89.

Manley JL and Tacke R (**1996**) SR proteins and splicing control. **Genes Dev.** 10, 1569-1579.

McCall K and Bender W (**1996**) Probes for chromatin accessibility in the *Drosophila* bithorax complex respond differently to *Polycomb*-mediated repression. **The EMBO J.** 15, 569-580.

McCracken S Fong N Yankulov K Ballantyne S Pan G Greenblatt J Patterson SD Wickens M Bentley DL (**1997**) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. **Nature** 385, 357-361.

Miller OLJr and Bakken AH (**1972**) Morphological studies of transcription. **Acta Endocrinol.** Suppl. 167, 155-177.

Milot E Strouboulis J Trimborn T Wijgerde M de Boer E Langeveld A Tan-Un K Vergeer W Yannoutsos N Grosveld F and Fraser P (**1996**) Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. **Cell** 87, 105-114.

Misteli T Cáceres JF and Spector DL (**1997**) The dynamics of a pre-mRNA splicing factor. **Nature** 387, 523-527.

Moen PT Smith KP and Lawrence JB (**1995**) Compartmentalization of specific pre-mRNA metabolism: an emerging view. **Hum. Mol. Genet.** 4, 1779-1789.

Monneron A and Bernhard W (**1969**) Fine structural organization of the interphase nucleus in some mammalian cells. **J. Ultrastruct. Res.** 27, 266-288.

Neuberger MS and Williams GT (**1988**) The intron requirement for immunoglobulin gene expression is dependent upon the promoter. **Nucl. Acids Res.** 16, 6713-6724.

Nickerson JA Blencowe BJ and Penman S (**1996**) The architectural organization of nuclear metabolism. **Int. Rev. Cytol.** 162A, 67-123.

Osheim YN Mougey EB Windle J Anderson M O'Reilly M Miller OLJr Beyer A and Sollner-Webb B (**1996**) Metazoan rDNA enhancer acts by making more genes transcriptionally active. **J. Cell Biol.** 133, 943-954.

Palmiter RD and Brinster RL (**1986**) Germ-line transformation of mice. **Annu. Rev. Genet.** 20, 465-499.

Pombo A and Cook PR (**1996**) The localization of sites containing nascent RNA and splicing factors. **Exp. Cell Res.** 229, 201-203.

Prasad, BVV Rothnagel R Zeng CQ-Y Jakana J Lawton JA Chiu W and Estes MK (**1996**) Visualization of ordered genomic RNA and localization of transcriptional complexes in rotavirus. **Nature** 382, 471-473.

Puvion E and Puvion-Dutilleul F (**1996**) Ultrastructure of the nucleus in relation to transport and splicing: roles of perichromatin fibrils and interchromatin granules. **Exp. Cell Res.** 229, 217-225.

Rosbash M and Singer RH (**1993**) RNA travel: tracks from DNA to cytoplasm. **Cell** 75, 399-401.

Schedl P and Grosveld F (**1995**) Domains and boundaries. In: Elgin SCR (Ed) Chromatin structure and gene expression (pp 172-196). IRL Press, Oxford.

Shaw PJ and Jordan EG (**1995**) The nucleolus. **Ann. Rev.Cell Dev. Biol.** 11, 93-121.

Singh PB (**1994**) Molecular mechanisms of cellular determination: their relation to chromatin structure and parental imprinting. **J. Cell Sci.** 107, 2653-2668.

Sisodia SS Sollner-Webb B and Cleveland DW (**1987**) Specificity of RNA maturation payhways: RNAs transcribed by RNA polymerase III are not substrates for splicing or polyadenylation. **Mol. Cell. Biol.** 7, 3602-3612.

Spann P Feinerman M Sperling J and Sperling R (**1989**) Isolation and visualization of large compact ribonucleoprotein particles of specific nuclear RNAs. **Proc Natl. Acad. Sci. USA** 86, 466-470.

Spector DL (**1993**) Macromolecular domains within the cell nucleus. **Annu. Rev. Cell Biol.** 9, 265-315.

Sperling R Koster AJ Melamed-Bessudo AR Angenitzki M Berkovitch-Yellin Z and Sperling J (**1997**) Three-dimensional image reconstruction of large nuclear RNP (lnRNP) particles by automated electron tomography. **J. Mol. Biol.** 267, 570-583.

Stein GS van Wijnen AJ Stein J Lian JB and Montecino M (**1996**) Contributions of nuclear architecture to transcriptional control. **Int. Rev. Cytol.** 162A, 251-278.

Strick R and Laemmli UK (**1995**) SARs are *cis* DNA elements of chromosome dynamics: synthesis of a SAR repressor protein. **Cell** 83, 1137-48.

Strouboulis J and Wolffe AP (**1996**) Functional compartmentalization of the nucleus. **J. Cell Sci.** 109, 1991-2000.

Tjian R and Maniatis T (**1994**) Transcriptional activation: A complex puzzle with few easy pieces. **Cell** 77, 5-8.

van Driel R Wansink DG Van Steensel B Grande MA Schul W and de Jong L (**1996**) Nuclear domains and the nuclear matrix. **Int. Rev. Cytol.** 162A, 151-189.

Verheijen R van Venrooij W and Ramaekers F (**1988**) The nuclear matrix: structure and composition. **J. Cell Sci.** 90, 11-36.

Wansink DG Schul W van der Kraan I van Steensel B van Driel R and de Jong L (**1993**) Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. **J. Cell Biol.** 122, 283-293.

Wijgerde M Grosveld F and Fraser P (**1995**) Transcription complex stability and chromatin dynamics *in vivo*. **Nature** 377, 209-213.

Wolffe A (**1995**) Chromatin structure and function. Academic Press, 2nd Edition.

Wolgemuth DJ and Hsu M-T (**1981**) Visualization of nascent RNA transcripts and simultaneous transcription and replication in viral nucleoprotein complexes from adenovirus 2-infected HeLa cells. **J. Mol. Biol.** 147, 247-268.

Xing Y Johnson CV Dobner PR and Lawrence JB (**1993**) Higher level organization of individual gene transcription and RNA splicing. **Science** 259, 1326-1330.

Yokota H van den Engh G Hearst JE Sachs RK and Trask BJ (**1995**) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. **J. Cell Biol.** 130, 1239-1249.

Zawel L and Reinberg D (**1995**) Common themes in assembly and function of eukaryotic transcription complexes. **Ann. Rev. Biochem.** 64, 533-561.

Zhang G Taneja KL Singer RH and Green MR (**1994**) Localization of pre-mRNA splicing in mammalian nuclei. **Nature** 372, 809-812.

Zhao K Hart CM and Laemmli UK (**1995**) Visualization of chromosomal domains with boundary element-associated factor BEAF-32. **Cell** 81, 879-889.