Cell, Vol. 91, 291–294, October 31, 1997, Copyright ©1997 by Cell Press

Compartmentalization of Eukaryotic Gene Expression: Causes and Effects

Minireview

Robert H. Singer* and Michael R. Green†
*Department of Anatomy and Structural Biology and Cell Biology
Albert Einstein College of Medicine
Bronx, New York 10461
†Howard Hughes Medical Institute
Program in Molecular Medicine
University of Massachusetts Medical Center
Worcester, Massacusetts 01605

Expression of eukaryotic structural genes is a multistep process that includes transcription of the gene, splicing, and polyadenylation of the primary transcript, and transport of the fully processed mRNA to the cytoplasm. The intricate and complex nature of these biochemical events raises the legitimate question of whether there are underlying organizing principles of the eukaryotic nucleus required for proper gene expression. Contemporary manuscripts dealing with this subject often begin with a phrase, such as: "It is becoming increasingly clear that the nucleus is highly organized." This minireview presents our thoughts on various aspects and implications of this statement. Specifically, we address several questions: What is the molecular definition of "highly organized"? How does this putative organization manifest itself structurally and functionally? Does such organization involve the components involved in gene expression (e.g., transcription and splicing factors), the activities required to express genes (e.g., transcription and splicing), or both? If gene expression is compartmentalized, do all of these activities reside in a single subnuclear compartment, or is there a separate nuclear subcompartment for each of the activities?

Lessons from the Nuclear Pore Complex and Nucleolus

We begin with a brief overview of two well-established but substantially different nuclear subcompartments: the nuclear pore complex (NPC) and nucleolus. The NPC has been known for over twenty years to specifically transport RNAs out of and proteins into the nucleus, and there is no evidence that nucleocytoplasmic transport can occur outside of the NPC. Over the past several years, extraordinary progress has been made in developing methods to purify the NPC, identify its protein components, and study how it functions. In addition, genetic experiments have confirmed and extended the relationship between components of the NPC with one another and with other factors involved in nuclear-cytoplasmic trafficking. The two conclusions that emerge from these considerations is that all nuclear-cytoplasmic trafficking occurs via the NPC (i.e., it is an obligatory compartment) and the NPC is a distinct entity that can be identified microscopically, genetically, and biochemically.

A different type of nuclear compartment is exemplified by the nucleolus. This nuclear organelle is easily visualized microscopically and is functionally correlated with the synthesis and processing of ribosomal RNA and the assembly of ribosomes. Genes on several chromosomes assemble together in order to organize the nucleolar structure. Within this nucleolar "factory," there are subcompartments that may be associated with cleavage and processing of the rRNA: a dense fibrillar component appears to contain the nascent chains, whereas a granular component appears to contain the mature RNA and partially assembled ribosomes. The nucleolus has been referred to as the "paradigm for nuclear compartmentalization" (Strouboulis and Wolffe, 1996), and it is therefore worthwhile to review the genesis of this structure.

It is now widely accepted that the initiating event for formation of the nucleolus is transcription of the tandemly repeated rDNA genes. When rDNA is inserted into a euchromatic region in Drosophila, or is coupled to a pol II promoter in a yeast pol I deletion strain, transcription of these genes initiates formation of a nucleolar-like structure, and processing of the ribosomal RNAs is normal when expressed from a plasmid containing a single copy of the rRNA gene (Nierras et al., 1997 and references therein). Therefore, the integrity of the bona fide nucleolus is not obligatory for ribosome biogenesis. Taken together, these results indicate that the "higher order" structure of the nucleolus is a coalescence of individual transcription units into a large array, that this coalescence is largely an effect of RNA polymerase I transcriptional activity and rRNA processing.

If the nucleolus is a paradigm for nuclear compartmentalization, then this compartmentalization is a byproduct of the initation of transcription, and the presumed consequent interactions of the proteins recruited to that site by the nascent RNA. Thus, comparison of the NPC and nucleolus illustrate two different ways for the genesis of a morphologically defined nuclear subcompartment: the NPC is a preassembled structure, whereas formation of the nucleolus is dynamic and depends upon underlying gene activity.

Compartmentalization of Factors Involved in RNA Biogenesis?

The principal reason to suggest that the eukaryotic nucleus may be functionally compartmentalized is based upon the observation that antibodies to specific nuclear antigens, typically splicing factors, are not homogeneously distributed within the nucleoplasm. The pattern receiving most attention is the 20–50 variably sized, ragged-edged "speckles" that are visualized by staining with antibodies to many splicing factors (Spector, 1993). This observation has been expanded by several recent technical advances, including: derivation of antibodies to newly discovered splicing factors; in situ hybridization protocols capable of distinguishing genes and their transcripts using fluorescence microscopy; and digital microscopy, which can increase detection of weak signals. These nonhomogeneous distributions have been suggested to reveal compartments of RNA biogenesis (e.g., splicing centers). In support of this hypothesis are studies reporting a spatial correlation between the speckles and genes that are actively transcribed or spliced (Spector, 1993; Xing et al., 1995; Huang and Spector, 1996).

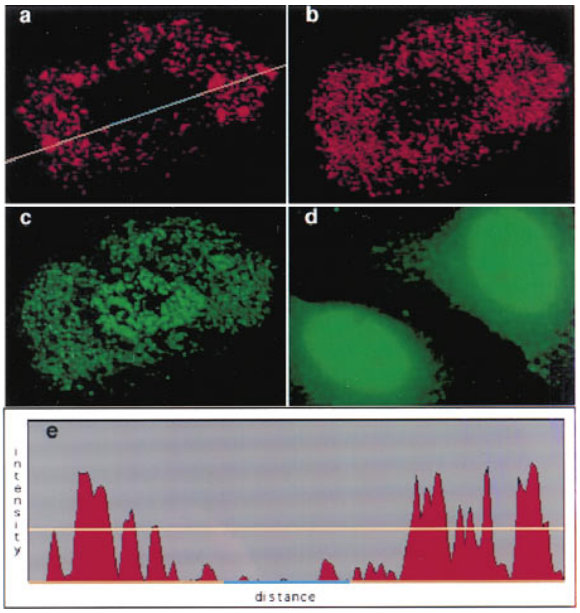

Figure 1. Digital Imaging of Diffuse and Concentrated Compartments in the Nucleus and Their Relationship to Function

The SC35 concentrated (a) and diffuse (b) compartments can be revealed by digital imaging. The histogram (e) shows a horizontal line representing the mean intensity of all nuclear SC35 signal (minus the nucleolus) where the digital image was thresholded after background subtraction along the selected line scan of the intensity levels in (a). The blue line indicates the region of the nucleolus that was used to set SC35 background levels. The signal above (a) and below (b) this threshold is shown in each image. Note that the speckles can be increased or decreased in size by the level at which the threshold is set, and the contrast enhanced by eliminating increasing amounts of the diffuse signal. Note also that the highest intensity is twice the mean intensity. The functional compartment (nascent RNA) is shown in (c), after microinjection of BrUTP. Note that the incorporation of BrUTP mimics the diffuse compartment of SC35, not the concentrated or speckled compartment. In contrast, (d) shows the diffusely localized transcription factor TATA box–binding protein (TBP). (a)–(c) reproduced from Fay et al. (1997).

We suggest that speckles are fluctuations of concentrations of various factors associated with RNA biogenesis that may be enhanced (or exaggerated) using digital imaging microscopy to create artificial boundaries by thresholding the signal to enhance contrast. An example of this is found for the essential splicing factor SC35, the now standard marker for speckles. In actuality, the SC35 signal in speckles is only twice that of the nucleoplasmic average (Fay et al., 1997). In other words, SC35 is ubiquitously distributed throughout the nucleoplasm but is more dense in some regions than in others. Figure 1 illustrates how a simple change in the arbitrarily set threshold parameter can dramatically alter the apparent relationship between the speckles and surrounding nucleoplasm.

Several studies have shown that the SC35 localization pattern can be highly variable. For example, the staining pattern is dependent upon the concentration of antibody used; at lower antibody concentrations, the distribution of splicing factors appeared much more homogeneous (Neugebauer and Roth, 1997). In another recent study, the nuclear distribution of splicing factors, as well as factors involved in transcription and polyadenylation, was shown to be dependent upon the physiological state of the cell; in cells with high transcriptional activity, the distribution of these factors was much more diffuse than the typical speckled pattern (Zeng et al., 1997).

Is RNA Splicing Compartmentalized?

Thus, despite variations in concentration, splicing factors are in actuality present throughout the nucleoplasm and thus available to all pre-mRNAs. Therefore, the nonhomogeneous distribution of splicing factors does not indicate whether or not splicing is compartmentalized: functional experiments are required. One study (Zhang et al., 1994) tested whether splicing activity was confined to speckles by infecting cells with adenovirus (Ad2) to generate a collection of viral transcription units randomly distributed throughout the nucleoplasm. The viral products of these transcription units were then detected by in situ hybridization using probes that could distinguish spliced from unspliced RNA. If certain nucleoplasmic regions contained an insufficient concentration of splicing factors to support splicing (i.e., the region between the speckles), then these sites would contain unspliced but not spliced Ad2 RNA. However, this never occured; all viral genomes were found to be transcribed, and the resultant RNA spliced irrespective of nucleoplasmic location, which was often distant from speckles.

Significantly, if the level of Ad2 transcription was allowed to become extremely high, the distribution of speckles changed and became coincident with the viral genomes (Jimenez-Garcia and Spector, 1993; Bridge et al., 1995). Thus, high levels of transcription can alter the apparent spatial relationship between genes and speckles. In this regard, studies reporting a coincidence between speckles and active genes analyzed only endogenous genes with very high transcriptional activity (e.g., actin and fibronectin), or transfected genes that are expressed also at high levels (see also below).

Is RNA Polymerase II–Directed Transcription Compartmentalized?

Analysis of the localization of RNA polymerase II–directed transcription provides an independent assessment as to whether gene expression is subcompartmentalized in the nucleus. Three sets of results indicate that the transcriptional activity of the nucleus is not confined to particular regions, but rather occurs thoughout the nucleoplasm. First, RNA polymerase II (see, for example, Zeng et al., 1997) and its associated transcription factors (see Figure 1d) assume a random nucleoplasmic distribution, with nucleolar exclusion. Second, when global RNA polymerase II–directed transcription was analyzed by incorporation of BrUTP, it was spread randomly throughout the nucleoplasm as thousands of punctate signals, but the speckles were not preferentially labeled and thus were transcriptionally neutral (Jackson et al., 1993; Wansink et al., 1993; Fay et al., 1997; see Figure 1c). Finally, Ad2 transcription units distributed randomly throughout the nucleoplasm were all found to be transcriptionally active, and therefore had access to transcription factors (Zhang et al., 1994).

These considerations indicate that transcription activity is much more widespread than the 40–50 speckles typically observed in the nucleoplasm of the studied mammalian cells. Thus, for RNA processing to occur

exclusively in speckles would demand that the primary transcript move from the its site of synthesis to the speckles. But this possibility is inconsistent with a variety of other results. Most importantly, both ultrastructural and biochemical data show that at least for some pre-mRNAs, splicing occurs cotranscriptionally, indicating that transcription and splicing components are in proximity. This notion has been strengthened by recent experiments demonstrating physical interactions between the largest subunit of RNA polymerase II, and components of the splicing and polyadenylation apparatus (Steinmetz, 1997 and references therein). Taken together, these observations indicate that splicing and polyadenylation factors target active genes, distributed throughout the nucleoplasm.

Speckles: a Result, Not a Cause, of Gene Expression

The above considerations indicate that transcription and RNA processing can occur throughout the nucleoplasm. Thus, studies reporting a correlation between speckles and actively transcribed genes or intron-containing pre-mRNAs remain to be explained (e.g., Xing et al., 1995). First, we note that these studies do not find a 100% association between active genes and speckles, in agreement with the conclusion that gene activity is not obligatorily compartmentalized (Zhang et al., 1994). Differences in the degree of association may vary depending on the threshold of signal chosen for imaging, the location of the gene relative to the speckle, or the physiological state of the cells. Evidence suggests that any apparent associations observed between highly active genes and the speckles is the result, not the cause, of gene expression. Specifically, an actively transcribed intron-containing gene will recruit a high concentration of splicing factors to its vicinity (Huang and Spector, 1996). When splicing is completed, the splicing factor dissociates from the pre-mRNA and diffusion away from the gene is probably limited by the extraordinarily high concentration of macromolecules in the nucleus. Most importantly, many pre-mRNA splicing factors can engage in networks of protein–protein interactions (Wu and Maniatis, 1993), perhaps nucleating the formation of an aggregate (speckle) near the active gene. It is also possible that incompletely spliced RNAs and excised introns become associated with speckles because of their bound splicing factors.

Several pieces of evidence are consistent with this model. First, as described above, the association of speckles with Ad2 transcription units was dependent upon the level of transcription, an association occuring only at high levels of transcription (Jiminez-Garcia and Spector, 1993; Bridge et al., 1995). Second, as noted above, the studies reporting an association between speckles and gene activity have analyzed only highly active genes, a clearly nonrepresentative set. The choice of highly active genes is in large part due to limitations in sensitivity imposed by target (RNA) abundance. Likewise, a correlation between speckles and intron-containing genes can be explained by the fact that only intron-containing pre-mRNAs will recruit splicing factors (Huang and Spector, 1996).

The model described above implies that the speckles function, at least in part, as a reservoir that supplies factors to highly active genes. Again, several observations fit nicely with this idea. For example, Zeng et al. (1997) show that upon increased transcriptional activity, the speckles diminish and splicing factors become diffusely distributed throughout the nucleoplasm. Using a GFP-tagged version of the splicing factor ASF/SF2, Misteli et al. (1997) demonstrated movement of the factor from the speckles toward sites of active transcription. These authors further showed that upon treatment of cells with α-amanitin, the speckles rounded up, which would be expected because the lack of transcription eliminates binding sites for the factors. This redistribution of splicing factors from the speckles can be explained by a purely passive mechanism: splicing factors are called upon as needed and released from the speckle by mass action. The redistribution from speckles may also be more actively regulated: several studies have shown that phosphorylation of splicing factors leads to dispersion of the speckles. In either case, a proximal reservoir of splicing factors may facilitate processing of highly transcribed genes.

Nuclear Matrix

In contrast to this view of nuclear structures as dynamic and dependent on transcriptional activity, a "preexisting organization" view implies that some structural components superimpose an initial spatial organization upon the nucleus. For this to occur, it seems likely that such components would be stably anchored. Such a compartment has been suggested by the "nuclear matrix" or "nucleoskeleton," a fibrillar remnant resulting from the exhaustive extraction of nuclear components (Nickerson et al., 1997). In contrast to the other compartments, this nuclear structure is undefined structurally, biochemically, or genetically in a way that distinguishes it from the nucleoplasm in general. Proteins have been isolated from this nuclear matrix residue and cloned. In almost all cases, they are hnRNPs or pol II (Mattern et al., 1997). Because these proteins are not known to be structural, this would suggest that some of the matrix contains components that are made insoluble by the isolation procedure. When antibodies are raised to matrix components, the resulting immunofluorescence is invariably punctate, not fibrillar, suggesting that structural polymers of the matrix, if they exist, are very elusive.

There are several possible explanations for the matrix. One possibility is that the nuclear matrix results from the excessively nonphysiological extraction during which high concentrations of RNA binding proteins and RNA polymerize to make strands, or precipitate into large clusters. Another possibility is that fibrillar components of the matrix may be lamins or lamin-like proteins extending into the nucleoplasm from the inner surface of the nuclear envelope. A functionally more relevant possibility may be that they are composed of intranuclear filaments extending from the nuclear pore (Cordes et al., 1997). Finally, evidence suggests that specific DNA segments can form "matrix-associated regions" (MARs) within the interphase nucleus. Transcribed and replicating DNA is thought to loop out of these regions into the interchromatin spaces (Strouboulis and Wolffe, 1996). These structural components may be remnants of the scaffold structure that attaches the base of the DNA loops in metaphase chromosomes

(Bickmore and Oghene, 1996). Although organization of the genome into these functional and nonfunctional regions may be considered a higher order structure, whether this is related to any consequent "compartmentalization" is unknown.

Perspective

A critical examination of the literature does not support the idea that any pre-existing nuclear organization directs gene expression, but rather supports a model in which the expression of genes directs an apparent reorganization of nuclear components. Other nonhomogeneous distributions, such as coiled bodies or PML nuclear bodies (Strouboulis and Wolffe, 1996), may also arise by virtue of the principles enumerated here. Of course, it remains possible that pre-existing organized elements within the nucleus will eventually come to light, possibly in the form of interphase chromatin organization, or of the nucleoplasmic components of the NPC. In any case, it will be important to state the hypotheses and models clearly so that their predictions can be tested.

Selected Reading

Bickmore, W.A., and Oghene, K. (1996). Cell *84*, 95–104.

Bridge, E., Xia, D.X., Carmo-Fonseca, M., Cardinali, B., Lamond, A.I., and Pettersson, U. (1995). J. Virol. *69*, 281–290.

Cordes, V.C., Reidenbach, S., Rackwitz, H.R., and Franke, W.W. (1997). J. Cell Biol. *136*, 515–529.

Fay, F.S., Taneja, K.L., Shenoy, S., Lifshitz, L., and Singer, R.H. (1997). Exp. Cell Res. *231*, 27–37.

Huang, S., and Spector, D.L. (1996). J. Cell Biol. *133*, 719–732.

Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). EMBO J. *12*, 1059–1065.

Jimenez-Garcia, L.F., and Spector, D.L. (1993). Cell *73*, 47–59.

Mattern, K.A., van Goethem, R.E., de Jong, L., and van Driel, R. (1997). J. Cell. Biochem. *65*, 42–52.

Misteli, T., Caceres, J.F., and Spector, D.L. (1997). Nature *387*, 523–527.

Neugebauer, K.M., and Roth, M.B. (1997). Genes Dev. *11*, 1148–1159.

Nickerson, J.A., Krockmalnic, G., Wan, K.M., and Penman, S. (1997). Proc. Natl. Acad. Sci. USA *94*, 4446–4450.

Nierras, C.R., Liebman, S.W., and Warner, J.R. (1997). Chromosoma *105*, 444–451.

Spector, D.L. (1993). Annu. Rev. Cell Biol. *9*, 265–315.

Steinmetz, E.J. (1997). Cell *89*, 491–494.

Strouboulis, J., and Wolffe, A.P. (1996). J. Cell. Sci. *109*, 1991–2000.

Wansink, D.G., Schul, W., van der Kraan, I., van Steensel, B., van Driel, R., and de Jong, L.J. (1993). J. Cell Biol. *122*, 283–293.

Wu, J.Y., and Maniatis, T. (1993). Cell *75*, 1061–1070.

Xing, Y., Johnson, C.V., Moen, P.T. Jr., McNeil, J.A., and Lawrence, J. (1995). J. Cell Biol. *131*, 1635–1647.

Zeng, C., Kim, E., Warren, S.L., and Berget, S.M. (1997). EMBO J. *16*, 1401–1412.

Zhang, G., Taneja, K.L., Singer, R.H., and Green, M.R. (1994). Nature *372*, 809–812.