DNA Replication and Nuclear Organization: Prospects for a Soluble In Vitro System

Daniela S. Dimitrova and David M. Gilbert*
Department of Biochemistry and Molecular Biology, S.U.N.Y. Health Science Center, 750 East Adams Street, Syracuse, NY 13210

* Corresponding author.

Abstract: The role of nuclear structure in the replication of eukaryotic DNA has been the subject of debate for many decades. The recent demonstration that once-per-cell-cycle replication can take place in vitro without a nucleus, providing sufficiently high concentrations of replication factors are supplied, suggests that one role of the nucleus is to concentrate essential factors. This important finding has paved the way for the establishment of a purified biochemical system for replication of eukaryotic DNA. However, this soluble system, derived from Xenopus egg extracts, initiates replication within any DNA sequence and does not recapitulate the spatial and temporal regulation of DNA replication that is observed in most cells. In both Xenopus and Drosophila embryos, site-specific initiation of replication is not observed until after nuclei become transcriptionally active at the blastula stage of development. Furthermore, programmed changes in both the locations of origins and the time during S-phase at which sequences are replicated accompany key stages of metazoan development. Recent findings indicate that these changes correlate with changes in nuclear organization and that the spatial and temporal program for replication is established early in G1-phase when nuclei are structurally and functionally reorganized after mitosis.

Key Words: nucleus, chromosome domains, replication origins, replication timing, gene expression, development.

I. A GLIMPSE AT EUKARYOTIC DNA REPLICATION

A. At the Cytological Level: Replication Foci, Replicon Clusters, and Chromosomal Bands

Somatic cells compartmentalize their genome into segments that replicate according to a defined spatio-temporal program. Each segment contains one replicon or a cluster of several replicons that are activated in a coordinate manner (reviewed in Berezney et al., 1999). The boundaries of these segments coincide with the boundaries of chromomeric bands. In general, transcriptionally active domains (R-bands) replicate early and transcriptionally silent domains (G-bands) replicate late in S-phase (Hatton et al., 1988).

Discrete replication sites within the nucleus can be readily visualized by briefly pulse-labeling cells with nucleotide analogs and staining nuclei with fluorescent antibodies to the modified nucleotides (Dimitrova and Gilbert, 1999, and references therein). The resulting fluorescent replication foci are most well characterized in cultured mammalian cells. However, similar foci have been observed in plant cells, in Saccharomyces cerevisiae cell-free extracts, Drosophila larvae, amphibian egg extracts, and somatic cells (Berezney et al., 1999). During the first half of S-phase, replication takes place at 100 to 1000 sites located throughout the interior euchromatic regions, excluding the nucleoli. During mid-S, replication foci are fewer in number and are localized predominantly in the perinuclear and perinucleolar heterochromatic regions, whereas at the end of S-phase, replication occurs within a much lower number of large foci (often ring or horseshoe shaped), which likely represent long stretches of tandemly repeated sequences, such as LINES and satellite DNA. A number of replication proteins have been shown to dynamically colocalize with
replication sites throughout S-phase (Dimitrova et al., 1999, and references therein). It is likely that a significant fraction of replication foci encompass more than one replicon. Moreover, the DNA sequences replicated within a single focus retain their punctate appearance and spatial distribution in the nucleus throughout multiple cell cycles (reviewed in Berezney et al., 1999).

At the electron microscope level, replication sites appear as massive (0.1 μm in early S-phase and 0.25 to 0.7 μm in mid/late S) complexes of replication proteins, termed replication factories (Hozak et al., 1994), which are attached to a filamentous structural protein/RNA network known as the nuclear matrix (Ma et al., 1999). Radiolabeled nascent DNA originates from within and can be chased out of these factories. Replication factories continue to support DNA synthesis even after extraction procedures that remove greater than 90% of the nuclear protein (Hozak et al., 1994; Ma et al., 1999). Interestingly, the punctate pattern of fluorescently labeled replication foci is similarly stable to nuclear matrix-type extractions, but is lost when the matrix is destroyed (Ma et al., 1999). These observations provide a foundation for the prevailing view that DNA is spooled through immobile multiprotein complexes during replication. Yet, these complexes should not be regarded as static structures, because they are continuously disassembled and reassembled at new sites as the replication process spreads throughout the genome during S-phase (Dimitrova et al., 1999; Hozak et al., 1994).

B. At the Molecular Level: Replication Origins and Proteins

Replication initiates at multiple sites (replication origins) along the eukaryotic chromosomes. The most well-characterized eukaryotic replication origins are those from the yeast *Saccharomyces cerevisiae*, where specific "autonomously replicating sequences" (ARS elements) were identified based on their ability to direct autonomous replication when cloned into plasmid vectors. ARS elements have been genetically dissected and critical base pairs for initiation both in plasmids and in chromosomes have been identified. Using two-dimensional gel electrophoresis (2D gels) capable of distinguishing replication intermediates, such as bubbles and forks, ARS elements were shown to be the replication initiation sites both in ARS-containing plasmids and within yeast chromosomes (reviewed in Fangman and Brewer, 1991). Recently, a high-resolution PCR-based approach revealed that the 5' ends of nascent DNA strands map to a single nucleotide site within ARS1 (Bielinsky and Gerbi, 1999).

In general, DNA replication in metazoan chromosomes also initiates within specific chromosomal loci. Using a variety of methods that measure the abundance and distribution of nascent DNA strands, initiation sites for replication have been mapped to within 0.5 to 11 kb at about 30 different loci. However, with a single exception (Toledo et al., 1998), whenever 2D gel methods have been applied to map origins at these same loci, replication bubble structures are detected throughout large (several tens of kb) regions, and replication forks can be found moving in both directions throughout these same regions (Gilbert, 1998). It is important to understand that these results do not suggest replication initiates at random sites within metazoan chromosomes. While it is clear that replication is confined to specific origin loci, the size of these loci is the subject of much debate. The emerging scenario is that some origins consist of one or more efficient sites of initiation surrounded by several less efficient potential sites, whereas other origins are focused to a single, relatively short chromosomal segment.

The knowledge of specific origin sequences allowed for the identification of proteins that interact with yeast origins and their homologues in higher eukaryotes (reviewed in Leatherwood, 1998). Studies in both yeasts and metazoa have led to a global view for the assembly of prereplication complexes (pre-RCs) at the sites where replication will initiate. First is a complex of six proteins (the origin recognition complex, ORC), which is bound to yeast origins throughout the cell cycle. ORC recruits other pre-RC proteins during late mitosis and G1-phase. The assembly/disassembly of pre-RCs is regulated by oscillating levels of S-phase-promoting kinase (SPK) activity. Shortly after mitosis, SPK activities are ac-
tively eliminated by the destruction of their cyclin components. During this period of low SPK, Cdc6 associates with ORC-bound chromatin sites, followed by loading of another six-member family of proteins, the Mcm 2-7 complex, and possibly other pre-RC proteins. Subsequently, a rise in SPK activity triggers entry into S-phase and the initiation of replication, accompanied by the disassembly of pre-RCs. High levels of SPK activity, which persist until the end of mitosis, are required to initiate replication throughout S-phase and also serve to prevent the formation of new pre-RCs. Hence, once-per-cell cycle regulation of DNA replication is elegantly orchestrated by the temporal separation of mutually exclusive periods of high and low SPK activity. However, it is important to note that there is nothing inherent in this regulatory network that requires replication to initiate at specific DNA sequences or at a specific number of sites. In fact, as discussed below, replication initiates at random sites on DNA templates introduced into Xenopus egg extracts—the system that has provided most of the available information on the composition and assembly of metazoan pre-RCs.

II. REQUIREMENTS FOR NUCLEAR STRUCTURE IN EUKARYOTIC IN VITRO REPLICATION SYSTEMS

For decades investigators have tried unsuccessfully to develop a soluble system for the initiation of chromosomal DNA replication. This could be explained by a requirement to precisely reproduce all aspects of DNA organization typical for eukaryotic nuclei. This deficiency of in vitro systems can be bypassed by using preassembled G1-phase nuclei as a substrate.

To date, most in vitro replication studies have been performed with cell-free extracts from Xenopus eggs. Naked DNA or Xenopus sperm chromatin is first assembled into a functional nucleus before undergoing a single round of semiconservative DNA replication (Blow and Laskey, 1986). Importantly, replication of naked DNA or chromatin templates in the egg extracts proceeds without any preference for specific initiation sites (Gilbert et al., 1995, and references therein), similar to the in vivo situation in early Xenopus embryos. However, when intact Chinese hamster ovary (CHO) nuclei are used as a substrate, extracts initiate replication specifically at physiological origin sites within the dihydrofolate reductase (DHFR) locus (Gilbert et al., 1995), providing that nuclei are prepared after a specific point in G1-phase, termed the origin decision point (ODP; Wu and Gilbert, 1996).

Nuclear structure is an essential requirement for replication in this system. Disruptions in nuclear integrity lead to nonspecific initiation, inefficient replication, or even complete loss of replication competence. For example, inhibiting nuclear transport prevents replication of both Xenopus sperm chromatin and mammalian nuclei (Cox, 1992). Exposure of mammalian nuclei to nonionic detergents virtually eliminates their ability to replicate in Xenopus egg extracts (Dimitrova and Gilbert, 1998; Gilbert et al., 1995). Moreover, permeabilization of the CHO nuclear membrane eliminates site-specific initiation even under conditions that preserve high replication efficiency (Dimitrova and Gilbert, 1998). Perturbation of nuclear lamina assembly results in an altered distribution of imported replication factors (Spann et al., 1997) and a failure to initiate DNA replication (Ellis et al., 1997). Because nuclei lacking a lamina do not assemble a normal nucleoskeleton, it has been proposed that the absence of replication foci in nuclei assembled in Xenopus lamin B3-depleted extracts arises because these nuclei lack a properly organized nuclear matrix (Zhang et al., 1996). Finally, repereplication of G2-phase nuclei in Xenopus egg extract requires permeabilization of the nuclear membrane, either by experimental reagents or naturally during mitosis (Blow and Laskey, 1988; Leno et al., 1992). Because permeabilized G2 nuclei do not initiate replication if they are resealed with purified membranes prior to introduction into extract (Coverley et al., 1993), the intact membrane appears to prevent reinitiation by serving as a barrier to an essential positive replication factor(s).

Recently, two new cell-free systems capable of initiating chromosomal DNA replication have been developed. One system employs a soluble nuclear extract from yeast cells synchronized in S-phase (Pasero et al., 1997). With intact nuclei
as a substrate, replication is moderately efficient (≤ 20% input DNA replicated), but complete duplication of the genome has not been achieved yet. Importantly, replication is dependent on functional ORC within the substrate nuclei and initiates specifically at ARS elements. By contrast, with naked DNA as a substrate, these same extracts initiate replication inefficiently, in an ORC- and Cdc6-independent manner, and without any preference for ARS elements (Duncker et al., 1999). These results provide strong support for the view that a physiological nuclear environment is indispensable for the recapitulation of regulated genome replication. Another in vitro system was developed for the replication of nuclei in S-phase cytosolic extracts from mammalian cells (Krude et al., 1997; Stoeber et al., 1998). The use of nuclei prepared at various stages of the cell cycle demonstrated that only mammalian nuclei that have assembled functional pre-RCs are competent to replicate in these extracts. So far, replication is very inefficient (~5% of input DNA replicated) and the issue of preferential initiation within specific origin loci has not been addressed. Nonetheless, anticipated improvements in this system are certain to provide valuable insight into many aspects of mammalian DNA replication.

III. THE NUCLEUS AS A CONCENTRATING DEVICE

Perhaps the strongest challenge to the concept for a requirement of nuclear structure for DNA replication comes from the recent development of a soluble system capable of completing once-per-cell-cycle replication of chromatin or naked DNA templates (Walter et al., 1998). To test the assumption that one of the functions of the nuclear envelope is to provide an intranuclear environment (availability of replication factors and appropriate concentrations of regulatory activities) permissive for DNA replication, Newport and colleagues prepared highly concentrated nucleoplasmic fraction from Xenopus egg extracts. Next, they demonstrated that a single complete round of semiconservative replication can be achieved by initial incubation of either sperm chromatin or naked DNA in membrane-free Xenopus egg cytosol (allowing assembly of pre-RCs in the absence of nuclear formation), followed by the addition of concentrated nucleosol to trigger initiation of replication. These investigators propose that the nuclear envelope is required to concentrate SPK activity (through active import) to levels that promote initiation and prevent untimely pre-RC assembly. Hence, once-per-cell-cycle replication itself does not require specific DNA sequences or higher-order nuclear structure, but only alternating periods of high and low SPK activity. Of course, it is still formerly possible that the nucleus performs a role beyond simply concentrating proteins. Active nucleosolic extract may require a processing step taking place within structurally sound nuclei. Regardless, the development of this completely soluble, highly efficient, and manipulable system has opened a new avenue to evaluate the molecular mechanisms regulating once-per-cell-cycle replication.

IV. THE NUCLEUS AS MORE THAN A CONCENTRATING DEVICE

While the soluble in vitro system described above may accurately reflect replication in rapid early cleavage embryos, it is unlikely that such a system can recapitulate the complexity of replication within differentiated somatic cells. The Xenopus and Drosophila embryonic nucleus is transcriptionally silent and appears to be organized exclusively to accommodate the requirements of rapid replication: replication proteins are in vast excess, initiation sites are selected at random and are closely spaced, and replication does not follow a defined spatio-temporal order. After the mid-blastula stage of development in insects and amphibians, there is a gradual focusing of replication initiation to fewer, more specific, origin sites (Hyrien et al., 1995; Sasaki et al., 1999). In Xenopus, this transition coincides with the onset of transcription, the appearance of histone H1, increase in chromosome loop and replicon sizes, changes in the composition of nuclear lamin proteins, changes in the spatial distribution of replication proteins, and a gradual
increase of the length of the cell cycle (Gilbert et al., 1995, and references therein). These features seem to be unique to organisms with rapid early cleavage cycles, because other embryonic systems differ in their replication characteristics. For example, early mouse embryos have much longer cell cycles, are transcriptionally active, and display dynamic spatio-temporal replication patterns that accompany changes in chromatin organization (Ferreira and Carmo-Fonseca, 1997). Furthermore, replication of viral DNA microinjected into mouse embryos requires specific origin sequences (Wirak et al., 1985). Hence, an in vitro system derived from Xenopus embryos inherently lacks features of nuclear organization that likely influence patterns of genome replication in most eukaryotic cells.

Recent evidence (discussed below) indicates that the various functions of nuclei (transcription, RNA processing, transport, DNA repair, and recombination) must be coordinated with replication. Regulatory proteins not only need to be concentrated within the nucleus, but they also need to be directed to the precise chromosomal sites where they are to perform their function. The influence of these complexities will need to be addressed on the molecular level before a soluble replication system capable of reproducing physiological control of replication can be developed. Below, we outline aspects of DNA replication that are influenced by other nuclear activities and discuss to what extent these aspects may be amenable to in vitro studies.

A. Coordination of Transcription and Replication

The relationship between transcription and replication is complex. DNA replication creates a window of opportunity to introduce changes in chromatin structure and thus activate or repress gene expression (Gilbert, 1986). Experimental evidence suggests that ongoing transcription can stimulate replication origin function by creating negative supercoiling upstream of the transcribed region that facilitates local DNA unwinding (Ohba et al., 1996). Furthermore, it has been proposed that transcription factors can activate cellular DNA replication through direct interactions with components of the replication machinery (Dutta, 1993) or by chromatin remodeling, which makes origin regions more accessible for binding of replication proteins (Hu et al., 1999). However, the presence of active replication origins in transcriptionally silent chromosomal regions (Kitsberg et al., 1993) argues that active transcription is not a universal requirement for origin function.

Duplication and expression of the genome must be precisely coordinated, so that they do not interfere with each other. In prokaryotes, the direction of transcription is the same as that of DNA replication, apparently to prevent head-on collisions of the transcription and replication machineries (Brewer, 1988). An additional mechanism may have evolved in higher eukaryotes; in mammalian cells, evidence suggests that transcription within most chromosomal domains is temporarily shut down during replication (Wansink et al., 1994; Wei et al., 1998). Despite this temporal separation of transcription and replication, most replication origins map to intergenic regions (Brewer, 1994). Furthermore, metazoan origins tend to be more circumscribed when located within a narrow intergenic region (Dimitrova et al., 1996; Toledo et al., 1998), compared with origins found in large intergenic regions (Wang et al., 1998). This preferential exclusion of origins from transcription units implies that transcription interferes with origin function even though replication and transcription do not occur simultaneously. One way to account for this is if the interference takes place during non-replicative phases of the cell cycle. In fact, it has been shown that the act of transcription through an origin inhibits initiation potential (Tanaka et al., 1994, and references therein). Further support for this hypothesis is provided by the observation that when general transcription is inhibited in CHO cells during G1-phase, replication initiation is no longer focused to the origin locus in the DHFR domain (Dimitrova and Gilbert, unpublished observations). Perhaps the passage of the huge transcription machinery destabilizes pre-RCs. Early in G1-phase, prior to the rise in SPK activity, pre-RCs may be dynamically reassembled at many potential sites, with pre-RCs located between transcription units having the
longest half-lives. After the rise in SPK activity, pre-RC assembly is no longer possible, and collision of a pre-RC with transcription would inactivate the origin. Continuous reassembly of the multiprotein pre-RCs may be too inefficient for more streamlined genomes such as yeast and viruses, which therefore have evolved specific sequences to direct the assembly of the pre-RCs to stable locations between genes (Brewer, 1994).

This model cannot explain the positions of all origins. First, examples have been found where replication initiates within transcription units. At least in some of these cases, efficient origin function might be achieved by restricting transcription to cell cycle times following disassembly of pre-RCs. For example, transcription of the human lamin B2 gene occurs during S-phase (Biamonti et al., 1992). Thus, the assembly of initiation complexes at the replication origin located at the 3'-end of this gene would not be affected by transcription during G1-phase and pre-RCs will remain stable until the origin is activated during the first minutes of S-phase (Biamonti et al., 1992). In cases where origins are located within genes that appear to be transcribed during G1-phase, specific sequences may somehow stabilize the pre-RC to the disruptive effect of the transcription machinery. Finally, highly localized origins have been mapped within transcriptionally silent genomic regions (Kitsberg et al., 1993). Again, specific DNA sequences or other features of chromatin structure could be responsible for focusing initiation to short chromosomal regions. Indeed, specific DNA sequences that appear to be more proficient at potentiating replication origin function in mammalian cells have just started to be characterized (Aladjem et al., 1998).

In conclusion, in vitro replication systems that do not take into account the influence of transcription are unlikely to recapitulate the physiological replication patterns at many metazoan origins. Introduction of appropriate transcription factors into soluble DNA replication systems could potentially lead to site-specific initiation in vitro with purified DNA templates. Successful reconstitution of transcriptionally active chromatin in Xenopus egg extracts starting from cloned DNA templates has already been reported (Barton and Emerson, 1994).

B. Nuclear Positioning and Replication Timing

Data rapidly accumulating in the literature reveal a highly organized cell nucleus with a distinct substructure (in fact, 1998 witnessed the first Cold Spring Harbor Meeting devoted to nuclear structure). The picture that emerges is one in which the positions of sequences within the nucleus reflects, at least in part, their transcriptional potential. Chromosomes are organized into distinct territories, with regulatory factors confined to inter- and intrachromosomal channels that reduce the effective volume and enhance the assembly of large macromolecular transcription and splicing complexes (Bridger et al., 1998; Ferreira et al., 1997; Zink et al., 1998). Active genes appear to be localized preferentially on the periphery of chromosome territories to facilitate contact with complexes that process and transport RNA (Hendzel et al., 1998; Kurz et al., 1996). Studies in different organisms have shown that transcriptional silencing is accompanied by changes in the position of the silenced genes in the nucleus (Brown et al., 1999; Csík and Henikoff, 1998; Dernburg et al., 1996). As discussed above, replication foci form stable units of chromosome structure (replication domains), distinguishable both during mitosis and interphase. Following mitosis, these units are directed to their characteristic positions within the newly formed nucleus over the course of the first 1 to 2 hours of the mammalian G1-phase. Importantly, the completion of replication domain repositioning at this early stage of the cell cycle is coincident with the establishment of a temporal order for their replication (Dimitrova and Gilbert, 1999). Together with the already strong correlation that has been established between replication timing and gene expression, these data suggest that the position of a chromosomal domain within the nucleus is a reflection of its functional activity and dictates its replication timing.

Much information pertaining to the organization of microenvironments within the nucleus is required before attempting to recapitulate these aspects of nuclear structure in a soluble system. In fact, it is difficult to imagine how these complexities of nuclear organization could be reproduced.
starting with cloned DNA templates. They may be more amenable to a top down approach that begins with whole nuclei as a substrate. In fact, the first cell-free system that can recapitulate the replication timing program using mammalian nuclei has been reported recently (Dimitrova and Gilbert, 1999). This should now allow the identification of mechanisms that dictate replication timing.

C. Developmental Regulation of Replication

Developmental switches in replication control are not limited to early *Xenopus* development. Developmentally regulated genes often replicate late in most tissues and early in cell types in which they are expressed (Hatton et al., 1988). The most extensively studied gene region in which a switch in replication pattern accompanies the activation of transcriptional potential is the human β-globin locus. At this locus, over 200 kb of DNA is early replicating and DNase I sensitive in human erythroleukemia cells, but late replicating and DNase I resistant in cultured lymphocytes and in HeLa cells (Dhar et al., 1988; Epner et al., 1988). The immunoglobulin genes provide another interesting example. In nonlymphoid cell lines, variable, joining, and constant heavy chain genes replicate at a time during S-phase proportional to their distance from a downstream origin (Ermakova et al., 1999). After transcriptional activation in pre-B cells, the entire locus is replicated from different origins very early in S-phase (Brown et al., 1987). The fact that different patterns of replication are observed in cells derived from different tissues provides compelling evidence that the temporal order of replication is influenced by epigenetic factors that are specific for each cell type and will be difficult to recapitulate through the use of a single model system.

V. CONCLUSIONS

The last decade of this century has witnessed a great deal of progress in the understanding of what regulates once-per-cell-cycle initiation of replication, including the development of the first soluble system that can initiate chromosomal DNA replication. By the same token, a great deal of recent evidence supports the view that spatial and temporal patterns of replication are strongly influenced by nuclear structure. So far, specific replication patterns have been observed only in living cells and cell-free systems that use whole nuclei as a template. While the influence of transcription on the ability of specific sequences to function as origins may well be within our grasp, it seems that there are levels of nuclear organization for which our knowledge is still in the preliminary stages. Furthermore, because both the sites of initiation of replication and the relative order in which origins are fired are developmentally regulated, no one soluble in vitro system will be capable of revealing the regulation of every origin.

ACKNOWLEDGMENTS

We apologize to all those whose work could not be referenced due to space limitations. The work in our lab is supported by grants to D.M.G. from the March of Dimes, American Cancer Society and the NIH.

REFERENCES


