The Two Faces of Higher Eukaryotic DNA Replication Origins

Linskens, Maarten H.K.; Huberman, Joel A.

# The Two Faces of Higher Eukaryotic DNA Replication Origins

# Minireview

**Maarten H. K. Linskens and Joel A. Huberman**
Department of Molecular and Cellular Biology
Roswell Park Cancer Institute
Buffalo, New York 14263

Earlier this year, Vaughn et al. (1990) described convincing evidence for the surprising conclusion that the DNA replication origin located downstream of the dihydrofolate reductase (DHFR) gene in Chinese hamster cells consists of a broad initiation zone extending over 28 kb. However, the data presented by Burhans et al. (1990) in this issue suggest the contradictory conclusion that replication forks must emanate bidirectionally from a site that is no larger than 450 nucleotides. In this review, we hope to show that these apparently conflicting data are not necessarily irreconcilable. Instead, the two sets of data may illuminate two different faces of the higher eukaryotic chromosomal replication origin.

The existence of replication origins in mammalian cells was first suggested by DNA fiber autoradiography, which revealed that replication forks move bidirectionally outward from sites (origins) irregularly spaced along chromosomal DNA molecules. However, the resolution of fiber autoradiography (15–30 kb) is insufficient to provide information about events during initiation of replication or to determine whether origins correspond to specific nucleotide sequences. The subsequent mapping of replication origins in E. coli and prokaryotic and eukaryotic viruses (most notably SV40) to a few hundred nucleotides or less suggested that eukaryotic chromosomal origins might correspond to short stretches of defined nucleotide sequence. Indeed, chromosomal replication origins in the yeast S. cerevisiae correspond to specific DNA segments of a few hundred nucleotides or less (Linskens and Huberman, 1988).

The identification of replication origins in higher eukaryotes, however, has proven to be more difficult than in yeast (reviewed in Umek et al., 1989). The complexity of the higher eukaryotic genome has been a major obstacle. To overcome this difficulty, Hamlin's laboratory developed a strain of Chinese hamster ovary cells containing 500–1000 copies of 240 kb surrounding the DHFR gene. By determining which restriction fragments of the amplified domain are replicated earliest in S phase, Heintz and Hamlin (1982) identified a region downstream of the DHFR gene that appeared to contain a replication origin. This region is still the only stretch of mammalian chromosomal DNA has been proven to contain a replication origin.

High resolution, early S phase labeling (Burhans et al., 1986; Leu and Hamlin, 1989) and replication polarity analysis (Handeli et al., 1989) have suggested the existence of two separate origins about 22 kb apart within the DHFR downstream region. Figure 1 summarizes these data. Burhans et al. (1990) measured Okazaki fragment polarity by an independent technique to map the left-hand origin (Figure 1) to even higher resolution. Their data suggest

that this origin contains, in less than 450 nucleotides, a switch from leading to lagging strand synthesis similar to that observed at the SV40 origin (Hay and DePamphilis, 1982; Figure 1). They call this switch region an origin of bidirectional replication.

All these studies are consistent with the conclusion that downstream of the DHFR gene there are two closely spaced replication origins that are similar in size (and probably mechanism of initiation) to the classical E. coli, yeast, and SV40 origins. However, using the same two-dimensional (2D) gel electrophoresis methods that had been employed to map yeast replication origins, Vaughn et al. (1990) obtained results very different from those obtained with yeast. Their data appear to suggest that replication can initiate anywhere within a 28 kb zone containing the two putative origins and that, within this zone, replication forks can move in either direction.

In evaluating these apparently contradictory data, it is important to realize that the techniques used by both groups are based on premises tested in other systems, but probe distinctly different aspects of replicating molecules. 2D gel techniques (used by Vaughn et al., 1990) identify nonlinear molecules in a population and to a certain extent describe the shape of those molecules. One technique can also be used to determine apparent directions of replication fork movement. Because 2D gel techniques detect replication intermediates based on their abundance in a population of molecules, they are biased toward intermediates that accumulate (Brewer and Fangman, 1988; Linskens and Huberman, 1988). In other words, they tend to detect slower stages of replication. In addition, the nonlinear molecules detected by the 2D gels are not necessarily all replication intermediates; they may be produced by recombination or other processes.

In contrast, Okazaki fragment polarity mapping (used by Burhans et al., 1990) identifies in which direction leading and lagging nascent strands are created on a template. Incorporation of radioactively and density labeled precursors allows identification of newly synthesized DNA. However, newly synthesized DNA is not exclusively involved in replication; repair and recombination may also involve DNA synthesis. In addition, the efficiency of label uptake can play a critical role: long-lived replication intermediates may be poorly labeled in short pulses. This method may therefore be biased toward faster stages of replication.
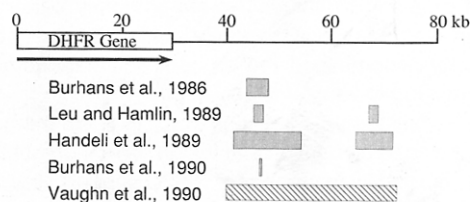


Figure 1. Origin Localization near the DHFR Gene in Chinese Hamster Cells

Based on these considerations, it is likely that the two sets of data shed light on disparate aspects of initiation at mammalian replication origins. Indeed, it is possible to integrate both sets of results into a single model, which can also serve as a working hypothesis for further testing. As a note of caution, this model is by no means the only possible way of reconciling the current data.

As shown in Figure 2, we propose that a mammalian replication origin contains an origin of bidirectional replication, as suggested by Burhans et al. (1990). We also propose, however, that replication does not initiate solely at this origin. Instead, initiation takes place throughout a large zone as proposed by Vaughn et al. (1990). Why a large zone should be required for initiation is not clear, but one possibility is that packing hundreds of origins into small regions of the nucleus imposes structural constraints on parental-strand unwinding that can only be resolved by unwinding at multiple sites throughout a large zone (reviewed in Laskey et al., 1989).

Electron microscopic studies of DNA isolated from proliferating mammalian cells have revealed that molecules containing clusters of microbubbles (small bubbles of several hundred basepairs) are much more abundant than molecules containing minibubbles (bubbles in the 1–10 kb size range), macrobubbles (>10 kb), or forks that would be produced by fragmentation of mini- or macrobubbles (Micheli et al., 1982). We propose that the first stage of initiation involves unwinding of parental strands to form microbubbles at random positions throughout the initiation zone; this is followed by a pause in the initiation process (to account for the abundance of microbubbles observed). Microbubbles are shown without nascent strands in Figure 2, because electron microscopic studies do not specify whether they contain nascent strands (Benbow et al., 1985). Owing to their small size, microbubbles may not render DNA molecules sufficiently nonlinear to be discriminated from linear molecules by 2D gel analysis.

How microbubble clusters are converted to minibubbles (large enough to be detected in 2D gels) is unknown, but the data of Burhans et al. (1990) suggest that Okazaki fragments must be generated on only one strand of each minibubble (except for the minibubble located over the origin of bidirectional replication). This asymmetry of Okazaki fragments would be a natural consequence of minibubble expansion if each minibubble (with the exception noted) were to expand unidirectionally outward from the origin of bidirectional replication (Figure 2). The data that Vaughn et al. (1990) interpreted as indicating that replication forks move in both directions throughout the initiation zone are equally consistent with this model, provided the stationary (inner) forks of the minibubbles are randomly located. By 2D gels, a population of randomly located stationary forks cannot be distinguished from a population of randomly located moving forks.

The last stage in the initiation process is the merger of minibubbles by elongation and fusion into a macrobubble. The 2D gel results (Vaughn et al., 1990) suggest structures consistent with the fusion process ("triangular smears"). Once formed, macrobubbles would expand bidirectionally
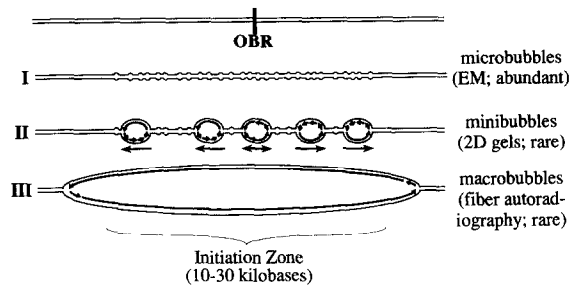


Figure 2. Model of Initiation of Replication in Higher Eukaryotic Cells
OBR, origin of bidirectional replication.

and give rise to the patterns detected by fiber autoradiography. Rapid expansion would explain the rarity of macrobubbles in electron microscopic studies.

Two additional observations on replication in higher eukaryotes indicate a requirement for long stretches of DNA for certain types of initiation. First, 2D gel analyses of chorion gene amplification during Drosophila oogenesis suggest that initiation takes place at multiple locations in a zone spanning several kilobases (Delidakis and Kafatos, 1989; Heck and Spradling, 1990). Second, numerous human DNA fragments can substitute for the dyad symmetry component of the Epstein–Barr virus latent replication origin, oriP, if the virus-encoded EBNA1 protein is provided in trans (Krysan et al., 1989). The ability of a human DNA fragment to permit efficient autonomous plasmid replication in certain types of human cells is more a function of its size than its primary sequence; most randomly selected fragments larger than 10 kb permit efficient replication (Yates, 1990). These data suggest a long stretch of human DNA is required for EBNA1-assisted autonomous replication, reminiscent of the apparent requirement for a broad initiation zone in mammalian chromosomal origins.

Although the proposed model is compatible with most experimental observations on mammalian replication origins, it may be inconsistent with two observations and fails to deal with a third. The frequency of initiation is apparently uniform throughout the initiation zone (Vaughn et al., 1990); the model may require a higher initiation frequency near the origin of bidirectional replication because one of the minibubbles must span the origin of bidirectional replication. The second observation possibly inconsistent with the model is that the shortest nascent leading strands appear to become progressively longer at increasing distances from the origin of bidirectional replication (Vassilev et al., 1990). However, this apparent conflict may not be serious because the nascent strand size measurements were performed over a stretch of only 5 kb.

The observation for which the model provides no explanation is the abundance of single-stranded DNA detected by electron microscopy in proliferating cells. Molecules that are totally single-stranded or with single-stranded gaps or tails are usually more abundant than any other type of possible replication intermediate, including microbubbles (Benbow et al., 1985; Micheli et al., 1982). The

abundance of single-stranded molecules and the absence of classical replication forks in DNA from Xenopus embryos led Gaudette and Benbow (1986) to conclude that, at least in such embryos, replication proceeds by synthesis of new DNA on already separated parental strands without formation of classical forks. There is no direct evidence for such strand separation that is not fork-driven in higher eukaryotic replication. However, there is a strong correlation between replication frequency and the amount of single-stranded DNA detected by electron microscopy; we suspect that a complete description of DNA replication in higher eukaryotic cells will eventually include a role for single-stranded DNA.

The model in Figure 2 suggests that a replication origin must contain at least three elements: a point or region in which the polarity of Okazaki fragments switches (called an origin of bidirectional replication by Burhans et al. [1990]); the region required for the initial unwinding and strand synthesis events and subsequent establishment of a stable, bidirectionally expanding macrobubble (referred to as an initiation zone by Vaughn et al. [1990], and perhaps analogous to the easily unwound stretch [Kowalski and Eddy, 1989; Umek et al., 1989] found in E. coli and yeast origins); and the sequence (still elusive in mammalian cells) recognized by initiation proteins, sometimes referred to as *ori*. It is possible that the origin of bidirectional replication colocalizes with *ori*. All three elements appear to be confined within a few hundred nucleotides in the simple origins of E. coli, SV40, and yeast.

New techniques promise to help clarify the complex structure of the higher eukaryotic replication origin. Handeli et al. (1989) have developed a transfection assay that with refinement should identify genetically essential elements of origins. In addition, the polymerase chain reaction permits the detection and size measurement of nascent strands produced by replication of unique genes in unsynchronized mammalian cells (Vassilev et al., 1990). Ligation-mediated polymerase chain reaction (Mueller and Wold, 1989) may permit localization of the 5' ends of nascent strands and possibly in vivo genomic footprinting of initiation proteins bound to *ori* sequences, both at single-nucleotide resolution.

Tests of this or any other model for higher eukaryotic replication origins will require identification of all predicted replication intermediates. Each intermediate should be characterized with respect to structure, abundance, turnover rate, and length and 5' end position of nascent strands. This information, combined with information on the proteins that interact with the origin, will provide a detailed understanding of initiation of replication and its regulation.

## References

Benbow, R. M., Gaudette, M. F., Hines, P. J., and Shioda, M. (1985). In Initiation of DNA Replication in Eukaryotes, Volume 1, A. L. Boynton and H. L. Leffert, eds. (New York: Academic Press), pp. 449–483.

Brewer, B. J., and Fangman, W. L. (1988). Cell 55, 637–643.

Burhans, W. C., Selegue, J. E., and Heintz, N. H. (1986). Proc. Natl. Acad. Sci. USA 83, 7790–7794.

Burhans, W. C., Vassilev, L. T., Caddle, M. S., Heintz, N. H., and DePamphilis, M. L. (1990). Cell 62, this issue.

Delidakis, C., and Kafatos, F. C. (1989). EMBO J. 8, 891–901.

Gaudette, M. F., and Benbow, R. M. (1986). Proc. Natl. Acad. Sci. USA 83, 5953–5957.

Handeli, S., Klar, A., Meuth, M., and Cedar, H. (1989). Cell 57, 909–920.

Hay, R. T., and DePamphilis, M. L. (1982). Cell 28, 767–779.

Heck, M. M. S., and Spradling, A. C. (1990). J. Cell Biol. 110, 903–914.

Heintz, N. H., and Hamlin, J. L. (1982). Proc. Natl. Acad. Sci. USA 79, 4083–4087.

Kowalski, D., and Eddy, M. J. (1989). EMBO J. 8, 4335–4344.

Krysan, P. J., Haase, S. B., and Calos, M. P. (1989). Mol. Cell. Biol. 9, 1026–1033.

Laskey, R. A., Fairman, M. P., and Blow, J. J. (1989). Science 246, 609–614.

Leu, T.-H., and Hamlin, J. L. (1989). Mol. Cell. Biol. 9, 523–531.

Linskens, M. H. K., and Huberman, J. A. (1988). Mol. Cell. Biol. 8, 4927–4935.

Micheli, G., Baldari, C. T., Carr, M. T., Cello, G. D., and Buongiorno-Nardelli, M. (1982). Exp. Cell Res. 137, 127–140.

Mueller, P. R., and Wold, B. (1989). Science 246, 780–786.

Umek, R. M., Linskens, M. H. K., Kowalski, D., and Huberman, J. A. (1989). Biochim. Biophys. Acta 1007, 1–14.

Vassilev, L. T., Burhans, W. C., and DePamphilis, M. L. (1990). Mol. Cell. Biol., in press.

Vaughn, J. P., Dijkwel, P. A., and Hamlin, J. L. (1990). Cell 61, 1075–1087.

Yates, J. (1990). In Viruses That Affect the Immune System, H. Fan, J. Coffin, and N. Rosenberg, eds. (Washington, D.C.: American Society for Microbiology), in press.