

functions. The refinement of de novo and comparative models provides a good test and application of the molecular dynamics methods widely used to simulate biological macromolecules (40).

Automated methods for deducing function from structure will be critical to obtaining functional insights from both predicted and experimentally determined structures. Considerable insight can be gained from structural comparison of a given structure with all other known protein structures using methods such as Dali (41), which can frequently detect structural relationships with functional significance that are not evident from sequence comparisons. Also promising are methods that match a structure against a library of structural motifs associated with different functions (42–44). For higher resolution models produced by comparative modeling methods, functional sites on proteins can potentially be identified and characterized by explicit ligand docking calculations. Finally, large-scale protein-protein docking calculations in years to come may contribute to the identification and characterization of protein interaction networks.

#### References

- M. A. Marti-Renom et al., *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291 (2000).
- S. F. Altschul et al., *Nucleic Acids Res.* **25**, 3389 (1997).
- A. E. Torda, *Curr. Opin. Struct. Biol.* **7**, 200 (1997).
- W. J. Browne, A. C. North, D. C. Phillips, *J. Mol. Biol.* **42**, 65 (1969).
- J. Greer, *J. Mol. Biol.* **153**, 1027 (1981).
- T. L. Blundell, B. L. Sibanda, M. J. Sternberg, J. M. Thornton, *Nature* **326**, 347 (1987).
- M. Levitt, *J. Mol. Biol.* **226**, 507 (1992).
- T. F. Havel, M. E. Snow, *J. Mol. Biol.* **217**, 1 (1991).
- A. Sali, T. L. Blundell, *J. Mol. Biol.* **234**, 779 (1993).
- A. Kolinski, M. R. Betancourt, D. Kihara, P. Rotkiewicz, J. Skolnick, *Proteins* **44**, 133 (2001).
- A. Fiser, R. K. G. Do, A. Sali, *Protein Sci.* **9**, 1753 (2000).
- M. J. Bower, F. E. Cohen, R. L. Dunbrack Jr., *J. Mol. Biol.* **267**, 1268 (1997).
- R. Samudrala, Y. Xia, E. Huang, M. Levitt, *Proteins* **3** (suppl.), 197 (1999).
- A. R. Ortiz, A. Kolinski, P. Rotkiewicz, B. Ilkowsky, J. Skolnick, *Proteins* **37**, 177 (1999).
- J. Pillardy et al., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2329 (2001).
- D. T. Jones, *Proteins* **1** (suppl.), 185 (1997).
- K. Simons, C. Strauss, D. Baker, *J. Mol. Biol.* **306**, 1191 (2000).
- R. Sanchez, A. Sali, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13597 (1998).
- P. Koehl, M. Levitt, *Nature Struct. Biol.* **6**, 108 (1999).
- M. A. Marti-Renom, M. S. Madhusudhan, A. Fiser, B. Rost, *Structure*, in press.
- L. Z. Xu, R. Sanchez, A. Sali, N. Heintz, *J. Biol. Chem.* **271**, 24711 (1996).
- R. Matsumoto, A. Sali, N. Ghildyal, M. Karplus, R. L. Stevens, *J. Biol. Chem.* **270**, 19524 (1995).
- C. M. T. Spahn et al., *Cell*, in press.
- See <http://predictioncenter.llnl.gov/casp4>.
- R. Bonneau et al., *Proteins: Structure, Function and Genetics*, in press.
- A. Gonzalez et al., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11221 (2000).
- G. Obmolova et al., *Nature* **407**, 703 (2000).
- R. Bonneau, J. Tsai, I. Rucinski, D. Baker, *J. Struct. Biol.* **134**, 186 (2001).
- A. E. Dann et al., *Nature* **412**, 86 (2001).
- D. Fischer, D. Eisenberg, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11929 (1997).
- N. Guex, A. Diemand, M. C. Peitsch, *Trends Biochem. Sci.* **24**, 364 (1999).
- See <http://guitar.rockefeller.edu/modbase/>.
- R. Sanchez et al., *Nature Struct. Biol.* **7** (suppl.), 986 (2000).
- R. Sanchez et al., *Nucleic Acids Res.* **28**, 250 (2000).
- S. K. Burley et al., *Nature Genet.* **23**, 151 (1999).
- D. Vitkup, E. Melamud, J. Moulton, C. Sander, *Nature Struct. Biol.* **8**, 559 (2001).
- A. Sali, *Nature Struct. Biol.* **5**, 1029 (1998).
- S. E. Brenner, *Nature Struct. Biol.* **7** (suppl.), 967 (2000).
- See <http://nysgrc.org/>.
- A. L. Brooks, M. Karplus, B. M. Pettit (Wiley, New York, 1988).
- L. Holm, C. Sander, *Trends Biochem. Sci.* **20**, 478 (1995).
- A. Zhang et al., *Protein Sci.* **8**, 1104 (1999).
- A. C. Wallace, N. Borkakoti, J. M. Thornton, *Protein Sci.* **6**, 2308 (1997).
- P. Aloy, E. Querol, F. X. Aviles, M. J. Sternberg, *J. Mol. Biol.* **311**, 395 (2001).
- R. Bonneau, D. Baker, *Annu. Rev. Biophys. Biomol. Struct.* **30**, 173 (2001).
- We are grateful to J. Frank and R. Beckmann for the picture of the ribosomal particle; M. A. Marti-Renom, R. Bonneau, and N. Eswar for help in preparing the figures; and members of our groups for many discussions about protein structure prediction. Supported by NIH/GM 54762 (A.S.), the Mathers Foundation (A.S.), a Merck Genome Research Award (A.S.), and the Howard Hughes Medical Institute (D.B.). A.S. is an Irma T. Hirsch Trust Career Scientist.

#### REVIEW

## Making Sense of Eukaryotic DNA Replication Origins

David M. Gilbert

DNA replication is the process by which cells make one complete copy of their genetic information before cell division. In bacteria, readily identifiable DNA sequences constitute the start sites or origins of DNA replication. In eukaryotes, replication origins have been difficult to identify. In some systems, any DNA sequence can promote replication, but other systems require specific DNA sequences. Despite these disparities, the proteins that regulate replication are highly conserved from yeast to humans. The resolution may lie in a current model for once-per-cell-cycle regulation of eukaryotic replication that does not require defined origin sequences. This model implies that the specification of precise origins is a response to selective pressures that transcend those of once-per-cell-cycle replication, such as the coordination of replication with other chromosomal functions. Viewed in this context, the locations of origins may be an integral part of the functional organization of eukaryotic chromosomes.

Transmission of genetic information from one cell generation to the next requires the accurate and complete duplication of each DNA strand exactly once before each cell

division. Typically, this process begins with the binding of an "initiator" protein to a specific DNA sequence or "replicator." In response to the appropriate cellular signals, the initiator directs a local unwinding of the DNA double helix and recruits additional factors to initiate the process of DNA replication. This paradigm describes most of the currently tractable replication systems and,

although derived from prokaryotic and viral systems, there is no compelling reason to doubt that it will apply to all eukaryotic organisms. In fact, the proteins that regulate replication are highly conserved from yeast to humans, including the origin recognition complex (ORC), which binds directly to replication origin sequences in budding yeast (1, 2). However, in several eukaryotic replication systems, it appears that any DNA sequence can function as a replicator. Those outside the field are often perplexed as to how investigators of different eukaryotic systems can work with assumptions that range from very specific to completely random origin sequence recognition, yet all agree on the basic mechanism regulating DNA replication. This review summarizes our current understanding of eukaryotic replication origins and then presents some simple guidelines to help demystify these seemingly disparate observations, providing a framework for understanding eukaryotic origins that includes all existing data.

Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 East Adams Street, Syracuse, NY 13210, USA. E-mail: gilbertd@mail.upstate.edu

# Conserved Initiator, Divergent Replicator

The structure of replication origins that have been subjected to genetic analysis in various eukaryotic species is summarized in Fig. 1. In *Saccharomyces cerevisiae*, origins [autonomously replicating sequences, (ARS)] consist of an essential 11-base pair (bp) ARS consensus sequence (ACS) and several additional elements that contribute to initiation activity and are interchangeable between origins but are not conserved in sequence. Several single-base pair mutations in the ACS can abolish initiation activity (3, 4). *Saccharomyces cerevisiae* ORC binds in vivo to ARS elements throughout the cell cycle, and purified ORC binds specifically to the ACS in an ATP-dependent manner (2, 5). High-resolution mapping of ARS1 has delimited the initiation point to a single nucleotide adjacent to the ORC binding site (6). In *Schizosaccharomyces pombe*, origins are much larger and consist of multiple elements that contribute partially to origin activity (2, 7). Although these elements do not share a consensus sequence, they contain asymmetric AT stretches and can be replaced with artificial asymmetric AT stretches (e.g., A<sub>40</sub>) that reconstitute full origin activity (7). In vivo, *S. pombe* ORC associates with at least two separate fragments within the ARS 2004 (8).

In *Drosophila*, the elements required for amplification of the chorion genes have been analyzed extensively. Although the process of gene amplification involves re-replicating the same DNA segment many times within one cell cycle, initiation requires many of the same proteins as normal chromosomal replication does, including ORC (9). Amplification requires the 440-bp amplification control element (ACE), whereas the extent of amplification is stimulated by the presence of several amplification-enhancing elements (AERs), of which only the most origin-proximal (AER-d) is shown in Fig. 1. ORC associates with both ACE and AER-d in vivo, and purified *Drosophila* ORC exhibits preferential ATP-dependent binding to a fragment containing AER-d and three separate DNA fragments within the ACE (10). It is intriguing that replication initiates almost exclusively within the dispensable AER-d (11–13). Furthermore, when multiple tandem copies of the ACE are integrated at ectopic sites, they recruit additional ORCs to chromatin sites in the flanking DNA that are not otherwise occupied by ORC (10). The tandem ACEs promote amplification at these ectopic sites, but the primary initiation sites appear to reside in the flanking DNA (14, 15). Taken together, the ACE appears to contain several ORC binding sites that promote the interaction of ORC and the initiation of replication at specific adjacent sites.

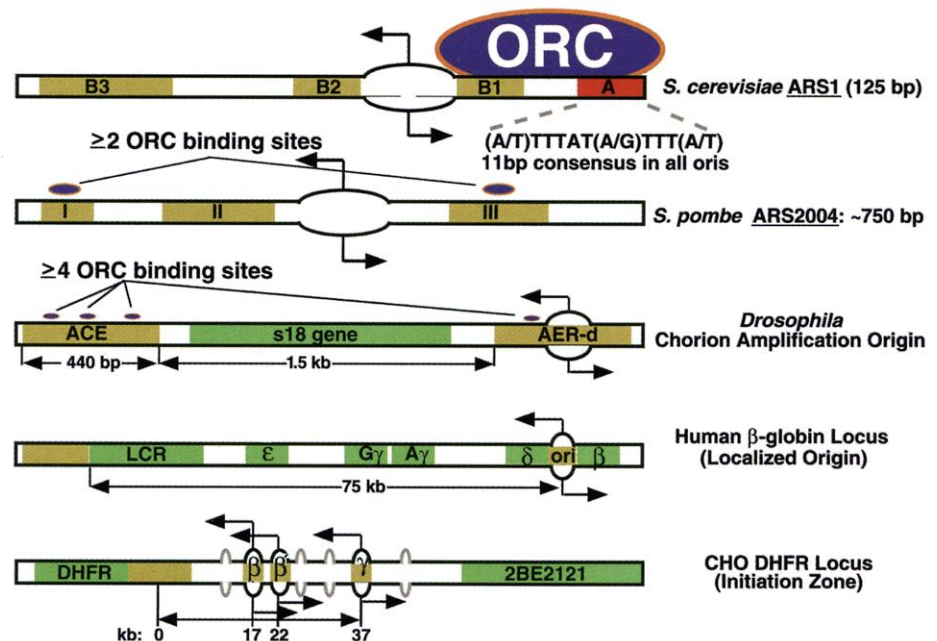
Replication origins in multicellular organ-

isms (metazoa) generally conform to one of two patterns. At some loci, initiation sites are localized to within a few kilobases. At other loci, multiple dispersed origins can be identified throughout "initiation zones" of 10 to 50 kb. Preliminary genetic dissection has been carried out at two of these loci, one representative of each class (Fig. 1). At the human  $\beta$ -globin locus, replication initiates within a few kilobases located between the adult  $\delta$ - and  $\beta$ -globin genes. However, deletions of sequences greater than 50 kb from the origin, as well as deletions within the initiation site itself, abolish the activity of this origin (16, 17). When the  $\beta$ -globin origin is transferred to an ectopic site, it can direct site-specific initiation of replication, and this activity is dependent on specific segments of DNA within the 8-kb transferred fragment (18). Other loci where initiation sites appear to be confined to within a few kilobases have been identified [e.g., (19, 20)]. In fact, at the lamin B2 locus in human cells, a single nucleotide demarcates the major transition between leading and lagging DNA synthesis (20).

The Chinese hamster ovary (CHO) dihydrofolate reductase (DHFR) locus is representative of the second class of metazoan origins. The pattern of initiation sites at this locus has been the subject of much controversy. On the one hand, high-resolution mapping of the locations of small nascent DNA strands revealed only a single origin throughout a 6.1-kb fragment (21). On the other hand, two-dimensional gel electrophoresis

(2D gel) analysis of replication intermediates identified DNA structures representing replication bubbles throughout the entire 55-kb intergenic zone between the DHFR gene and an adjacent gene (2BE2121) of unknown function (22). These results can be reconciled by considering the nature of the data obtained with each origin mapping technique. The 2D-gel method can search a larger area for initiation activity but cannot accurately discern the number or precise location of initiation sites within a fragment. By contrast, small nascent strand detection analyzes a focused area in detail. When the latter method was extended to cover an additional 6 kb of DNA, a second initiation site was revealed approximately 5 kb from the first (23). The 2D-gel method predicts that many more such sites will eventually be identified, constituting an initiation zone. Similar broad initiation zones have been identified at other metazoan loci (24–26).

Genetic analysis at the DHFR locus suggests the existence of specific elements that influence origin activity. When the most active origin is deleted (ori- $\beta$ ), adjacent replication origins retain or increase their activity (27). However, deletion of sequences near the 3' end of the DHFR gene renders the entire locus inactive for early S phase-initiation activity. Like the  $\beta$ -globin origin, DHFR ori- $\beta$  retains initiation activity when moved to ectopic sites, and deletions of specific sequences within ori- $\beta$  can influence this activity (28). Thus, origins found in both broad and localized initiation regions contain specific sequences favorable for



**Fig. 1.** Structure of genetically dissected replication origins in eukaryotes. Consensus ORC binding sites are indicated in red. Additional sequences important for origin activity are shown in brown. Transcription units and regulatory sequences are shown in green. Sites of ORC binding, where known, are indicated. Sites of initiation of replication are indicated with a bidirectional arrow passing through a bubble.

initiating DNA replication. However, the number and distribution of origins varies considerably at different loci.

Perhaps the most enigmatic aspect of the field is that, in many eukaryotic systems, replication seems to initiate within any DNA sequence. It appears that any cloned plasmid DNA will replicate autonomously in *Caenorhabditis elegans* (29) and *Paramecium* (30). In cultured animal cells, systematic searches for ARS elements analogous to those successfully carried out in yeast have generally failed to identify specific sequences that confer a significant replication advantage when reintroduced into cells (31). In one study with human cells, virtually every DNA fragment greater than 15 kb promoted autonomous and once-per-cell-cycle replication with equal efficiency, and initiation sites were distributed throughout the plasmid sequences (32). Similar results were obtained in cultured *Drosophila* cells (33). In *Xenopus* and *Drosophila* embryos, any DNA sequence will efficiently replicate once per cell cycle up to the blastula stage of development whether microinjected into embryos or introduced into egg extracts (34). Likewise, replication of embryonic chromosomes appears to initiate within any DNA sequence and does not become

focused to specific sites until the midblastula transition, when transcription and differentiation commence (35, 36). Despite a random origin site selection, in both *Xenopus* and *Drosophila* extracts, initiation of replication requires the ATP-dependent DNA binding activity of ORC (2, 37). How can such a precisely regulated process be carried out without the requirement for specific start sites? The explanation is revealed in the mechanism by which DNA replication is coordinated with the various phases of the cell cycle.

### Regulated Replication Is Independent of Sequence

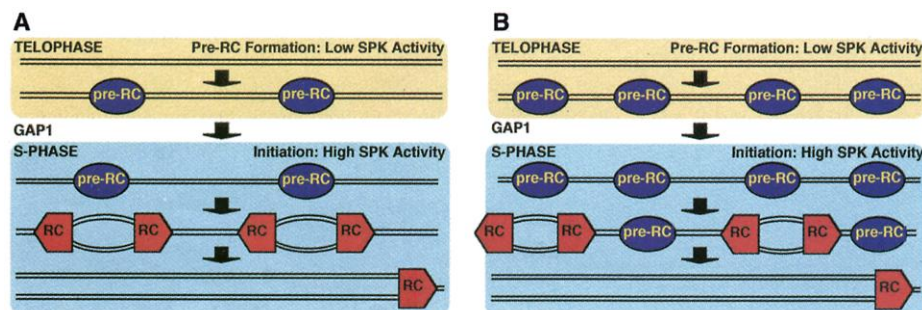
In all eukaryotic systems that have been amenable to study, replication is regulated by the assembly of a prereplication complex (pre-RC) of highly conserved proteins at ORC-bound DNA sites shortly after metaphase. After cells pass through the R point or START, a sharp rise in the activities of S phase-promoting kinases (SPK: Cdc7/Dbf4 and B-type cyclin-Cdk) triggers the conversion of the pre-RC to an active replication complex. These high levels of Cdk activity (38), as well as a protein called geminin (1, 39, 40), persist from the onset of S phase

through metaphase and prevent the assembly of new pre-RCs. Both of these activities are destroyed by proteolysis during anaphase, allowing pre-RCs to reassemble. Hence, mutually exclusive periods of the cell cycle that promote either pre-RC formation or initiation ensure that replication can only initiate once per cell cycle. This model (Fig. 2A) does not invoke any requirement for specific origin sequences to accomplish accurate duplication of the genome.

**The origin spacing problem.** Although initiation once per cell cycle does not require specific sequences, the positions of origins cannot be distributed randomly, as this would run the risk that some origins might be too far apart to complete replication of the intervening DNA within the length of a single S phase (41). One way to solve this problem is to direct ORC to specific DNA sequences spaced at appropriate intervals. However, the origin spacing problem can also be solved without the need for specific sequence recognition. Indeed, rapid replication in early *Xenopus* development is accomplished by initiating replication at sites that appear random with respect to sequence but are regularly spaced every 9 to 12 kb (41). The mechanism that establishes this regular origin spacing is unknown but, under conditions where chromatin is saturated with pre-RCs, any mechanism that prevents more than one pre-RC from assembling or firing per 10 kb would produce the observed spacing.

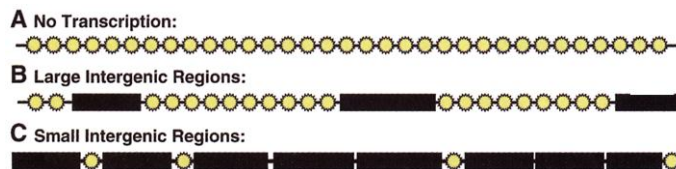
**More pre-RCs than needed.** An alternative means to solve the origin spacing problem is to assemble more pre-RCs than are necessary (Figs. 2B and 3), which appears to be the case in many eukaryotic systems. In budding yeast, many ARS elements function efficiently to promote the autonomous replication of plasmid DNA but do not normally function as origins in the chromosome or are utilized significantly less than once per cell cycle (42–44). However, pre-RCs are assembled on both active and silent origins (45). In *Xenopus* egg extracts, as the concentration of sperm nuclei is increased, the number of active origins per nucleus decreases but the number of ORC- and Mcm-DNA complexes assembled per nucleus remains constant (46). In mammalian cells, pre-RCs are assembled during telophase but pre-RC assembly is not sufficient to specify which sites will function as replication origins (47). Together, these results suggest that the number of pre-RCs assembled in eukaryotic cells exceeds the number of origins activated in each cell cycle; additional factors must select which of these pre-RCs will initiate replication.

Potentially, extraneous pre-RCs could effect the reduplication of portions of the genome if they were to persist on daughter DNA strands after replication. However, evidence suggests that pre-RCs are destroyed by the passage of



**Fig. 2.** Once-per-cell-cycle genome duplication is independent of the positions or density of replication origins. (A) Duplication of DNA exactly once-per-cell division is achieved with two mutually exclusive periods of the cell cycle during which either pre-RCs can be assembled but replication cannot initiate or replication can initiate but pre-RCs cannot be assembled. Regardless of where pre-RCs assemble, DNA is completely replicated and cannot be re-replicated until after cell division. (B) The assembly of extraneous pre-RCs can ensure that the entire segment of DNA is replicated in a timely fashion, without the need for specific DNA sequences to optimize origin spacing. Any pre-RCs that do not initiate are destroyed by passage of the replication fork and so cannot re-initiate on the replicated strands.

**Fig. 3.** Exclusion of pre-RCs from specific regions could create the need for origin focusing mechanisms. In this model, passage of the transcription apparatus before replication depletes transcription units (black boxes) of pre-RCs (yellow stars). (A) Without transcription, there is no selective pressure to focus initiation to specific sites as the assembly of pre-RCs at many sites ensures timely genome replication. (B) When transcription units are sparse, the assembly of multiple pre-RCs at many sites within the intergenic regions is sufficient to accomplish genome replication. (C) When intergenic regions are few and far between, specific DNA sequence recognition elements are required to ensure that at least one pre-RC is assembled at intervals appropriate to accomplish the timely replication of the DNA segment.





replication forks from active origins (Fig. 2B). Silent or infrequently utilized budding yeast origins become activated when replication forks are prevented from passing through them (48, 49). In both budding and fission yeast, when two or more origins are found in close proximity only one appears to be utilized within any given S phase (50–54). The mechanism that determines the frequency with which a given pre-RC will be activated is not known, however, chromosomal context and epigenetic elements clearly play a role. When identical ARS elements were placed in close proximity, one of them was utilized more frequently than the other, with the determinants of origin preference localized to flanking DNA sequences (52, 54). Mutations that disrupt silent chromatin at telomeres in budding yeast activate a normally silent telomeric origin (55) and manipulations that alter the positioning of nucleosomes on origins can directly influence the efficiency with which origins fire (56). Hence, each pre-RC has a potential to initiate replication that is influenced by a combination of local chromatin structure and the probability that it will be activated before the passage of a replication fork from adjacent origins.

### Origin Choice Is Developmentally Regulated

The gradual transition from random to specific origin site selection after the midblastula transition (MBT) during *Xenopus* (36) and *Drosophila* (35) development provides a dramatic example of origin choice during development. One possible explanation for this transition is that the higher concentration of ORC in preblastula embryos could result in a more relaxed binding of ORC to DNA, which would then become site-specific as the ORC:DNA ratio decreases after the MBT. Both purified (10) and recombinant (57) *Drosophila* ORC bind preferentially to specific DNA segments found near origins of gene amplification. In one case (57), the resolution was sufficient to conclude that replication initiates at the border of the ORC binding site, as in budding yeast (6). So far, the only sequence motifs shared by these ORC binding segments are short asymmetric AT stretches, reminiscent but not as prominent as those found in budding and fission yeast origins. Hence, it is presently difficult to predict the extent to which the affinity of ORC for specific DNA sequences, at any ORC:DNA ratio, can account for the specificity of initiation found in metazoan chromosomes. Changes in chromosome architecture that take place at the MBT, including chromatin condensation, the appearance of histone H1, cellular differentiation, and the gradual onset of transcription could also play a role in focusing initiation to specific sites (36, 58). Differences in origin specificity have also

been observed between cell lines from different tissues. In murine non-B cells, the entire IgH locus is replicated from a single replication fork that proceeds gradually through the locus from an origin located downstream of the constant region genes (59). However, in pre-B cells, the entire locus is duplicated during the first hour of S phase, indicating that one or more additional origins are activated.

There are several means by which origin specification could be influenced or even regulated during development. Changes in chromatin structure that accompany key stages of development could influence both the sequences to which ORC will preferentially bind and the efficiency with which pre-RCs are activated. The onset of differentiation could also result in the expression of accessory factors that interact with ORC and target it to specific sites. For example, the ORC4 subunit in *S. pombe* contains an NH<sub>2</sub>-terminal extension with nine copies of the HMG-I(Y) related AT-hook motif, which is known to mediate binding to the minor groove of AT tracts (60). In metazoa, interaction with accessory factors such as HMG proteins could focus ORC to specific DNA sequences as, for example, HMG proteins increase the site-specific binding of steroid receptors to their cognate sites (61). In fact, human ORC subunits coimmunoprecipitate with many unidentified proteins (62) and are targeted to the Epstein-Barr virus (EBV) replication origin, apparently through an interaction with the viral origin-binding protein EBNA (63–65). Furthermore, extracts from differentiated *Drosophila* tissue culture cells contain activities that appear to increase the selectivity of ORC for specific DNA fragments (66). Assuming that there is a selective advantage to regulating the positions of origins during development, utilizing ancillary factors to regulate the specificity of ORC could have circumvented the need to evolve different initiators for different tissues.

### Chromosomal Traffic Control

The fact that origins are localized in most species suggests that there is some selective pressure to initiate replication at particular sites. However, since once-per-cell-cycle replication per se does not require specific DNA sequences, the selective pressure must derive from considerations other than genome duplication. One potential source for this selective pressure is the need to coordinate transcription with replication (67). Indeed, transcription inhibits the autonomous replication of plasmids in human cells (68). There are at least two ways in which transcription and replication could be mutually antagonistic. First, head-on collisions of the replication and transcription machinery would create the need for both apparatuses to share the same

DNA template temporarily. Bacterial genes are heavily biased toward an orientation that places the polarity of transcription and replication in the same direction (69, 70). In both budding yeast and humans, a physical barrier in the 3' region of the ribosomal genes prevents replication forks from traveling in a direction that opposes RNA polymerase (25, 71, 72). Also, replication forks stall when they oppose the direction of yeast tRNA transcription but not when the tRNA genes are defective in transcription (73). However, there is some evidence in mammalian cells to suggest that transcription and replication do not take place simultaneously on the same DNA segments [reviewed in (74)], suggesting that polymerase collisions may not generally occur in mammalian cells.

A second antagonistic effect of replication and transcription could take place before replication, if pre-RCs are disrupted by passage of the transcription machinery. Although direct evidence for this mechanism is lacking, and exceptional cases exist (origins within transcription units and localized origins within transcriptionally silent regions), the explanatory power is intriguing (Fig. 3). In *Xenopus* embryos, there is no transcriptional activity before the MBT, so there is no selective pressure to place origins at particular locations. In contrast, the vast majority of the budding yeast genome is transcribed, and the locations of replication origins are almost exclusively restricted to intergenic regions (67). In this context, there would be a strong selective advantage for evolving specific DNA sequences that focus initiation to intergenic sites to ensure proper origin spacing (Fig. 3). This logic can also be applied to understand the different patterns of initiation at individual metazoan loci. Most solitary origin sites have been identified within loci containing multiple genes (18–20). By contrast, broad initiation zones consisting of multiple inefficient origins are observed at loci where there are large intergenic regions (22, 24, 25).

At present, it is difficult to predict the significance of transcription to origin localization. Clearly, we need to determine the sites where pre-RCs assemble during telophase and whether the onset of transcription after mitosis does, indeed, displace pre-RCs from transcription units. New applications for microarray analysis that reveal the genome-wide locations of DNA bound proteins have already been developed in budding yeast (75), and the answers to some of these questions in this simple model system should be forthcoming. In cases where replication does initiate within transcription units, it will be important to determine whether transcription is activated after replication. Finally, the coordination of transcription with replication provides a useful working model but other roles for origin placement should be considered. For example, ORC appears to play a central role in the assembly of

heterochromatin (76) and chromosome condensation during mitosis (77, 78). The location of replication origins may be important to organize chromosomes for sister chromatid cohesion and/or chromosome condensation.

What is clear is that the need to duplicate the genome once per cell cycle does not itself impose selective pressure to initiate replication at specific sites. In fact, it could be argued that a more degenerate origin recognition system would favor rapid genome evolution, allowing more flexible rearrangement of sequences without risking large gaps of DNA without an origin, and allowing for the modulation of origin sites during development. With this revelation, the difficulties in identifying consensus origin sequences in metazoa should come as no surprise. The more pressing question becomes why origins are localized at all in most eukaryotic systems. The answer to this question may be different for different loci. Although this indeed complicates the analysis of eukaryotic origins, it should by no means discourage investigators from entertaining creative hypotheses and pursuing directions that will likely reveal new insights into the organization of complex genomes.

# References and Notes

1. J. F. Diffley, *Curr. Biol.* **11**, R367 (2001).
2. T. J. Kelly, G. W. Brown, *Annu. Rev. Biochem.* **69**, 829 (2000).
3. Y. Marahrens, B. Stillman, *Science* **255**, 817 (1992).
4. J. Van Houten, C. S. Newlon, *Mol. Cell. Biol.* **10**, 3917 (1990).
5. R. D. Klemm, S. P. Bell, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8361 (2001).
6. A. K. Bielinsky, S. A. Gerbi, *Mol. Cell* **3**, 477 (1999).
7. Y. Okuno, H. Satoh, M. Sekiguchi, H. Masukata, *Mol. Cell. Biol.* **19**, 6699 (1999).
8. Y. Ogawa, T. Takahashi, H. Masukata, *Mol. Cell. Biol.* **19**, 7228 (1999).
9. A. C. Spradling, *Genes Dev.* **13**, 2619 (1999).
10. R. J. Austin, T. L. Orr-Weaver, S. P. Bell, *Genes Dev.* **13**, 2639 (1999).
11. C. Delidakis, F. C. Kafatos, *EMBO J.* **8**, 891 (1989).
12. M. M. Heck, A. C. Spradling, *J. Cell Biol.* **110**, 903 (1990).
13. L. Lu, H. Zhang, J. Tower, *Genes Dev.* **15**, 134 (2001).
14. D. de Cicco, A. C. Spradling, *Cell* **38**, 45 (1984).
15. J. L. Carminati, C. G. Johnston, T. L. Orr-Weaver, *Mol. Cell. Biol.* **12**, 2444 (1992).
16. D. M. Cimbora et al., *Mol. Cell. Biol.* **20**, 5581 (2000).
17. M. Aladjem et al., *Science* **270**, 815 (1995).
18. M. I. Aladjem, L. W. Rodewald, J. L. Kolman, G. M. Wahl, *Science* **281**, 1005 (1998).
19. F. Toledo et al., *Nucleic Acids Res.* **26**, 2313 (1998).
20. G. Abdurashidova et al., *Science* **287**, 2023 (2000).
21. C. Pelizon, S. Diviacco, A. Falaschi, M. Giacca, *Mol. Cell. Biol.* **16**, 5358 (1996).
22. P. A. Dijkwel, J. L. Hamlin, *Mol. Cell. Biol.* **15**, 3023 (1995).
23. T. Kobayashi, T. Rein, M. DePamphilis, *Mol. Cell. Biol.* **18**, 3266 (1998).
24. S. Ina, T. Sasaki, Y. Yokota, T. Shinomiya, *Chromosoma* **109**, 551 (2001).
25. R. D. Little, T. H. Platt, C. L. Schildkraut, *Mol. Cell. Biol.* **13**, 6600 (1993).
26. P. A. Dijkwel, L. D. Mesner, V. V. Levenson, J. d'Anna, J. L. Hamlin, *Exp. Cell. Res.* **256**, 150 (2000).
27. R. F. Kaleja et al., *Mol. Cell* **2**, 797 (1998).
28. A. L. Altman, E. Fanning, *Mol. Cell. Biol.* **21**, 1098 (2001).
29. C. C. Mello, J. M. Kramer, D. Stinchcomb, V. Ambros, *EMBO J.* **10**, 3959 (1991).
30. C. S. Kim, J. R. Preer, B. Polisky, *Dev. Genet.* **13**, 97 (1992).
31. H. Masukata, H. Satoh, C. Obuse, T. Okazaki, *Mol. Biol. Cell* **4**, 1121 (1993).
32. P. J. Krysan, J. G. Smith, M. P. Calos, *Mol. Cell. Biol.* **13**, 2688 (1993).
33. J. G. Smith, M. P. Calos, *Chromosoma* **103**, 597 (1995).
34. D. Coverley, R. A. Laskey, *Annu. Rev. Biochem.* **63**, 745 (1994).
35. T. Sasaki, T. Sawado, M. Yamaguchi, T. Shinomiya, *Mol. Cell. Biol.* **19**, 547 (1999).
36. O. Hyrien, C. Maric, M. Mechali, *Science* **270**, 994 (1995).
37. I. Chesnokov, M. R. Botchan, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
38. E. Noton, J. F. Diffley, *Mol. Cell* **5**, 85 (2000).
39. J. A. Wohlschlegel et al., *Science* **290**, 2309 (2000).
40. S. Tada, A. Li, D. Maiorano, M. Mechali, J. J. Blow, *Nature Cell Biol.* **3**, 107 (2001).
41. J. J. Blow, P. J. Gillespie, D. Francis, D. A. Jackson, *J. Cell Biol.* **152**, 15 (2001).
42. M. Yamashita et al., *Genes Cells* **2**, 655 (1997).
43. K. L. Friedman, B. J. Brewer, W. L. Fangman, *Genes Cells* **2**, 667 (1997).
44. A. Poloumienkol, A. Dershowitz, D. Jitakshi, C. S. Newlon, *Mol. Biol. Cell*, in press.
45. C. Santocanale, J. Diffley, *EMBO J.* **15**, 6671 (1996).
46. J. Walter, J. Newport, *Science* **275**, 993 (1997).
47. Y. Okuno, A. J. McNairn, N. den Elzen, J. Pines, D. M. Gilbert, *EMBO J.* **20**, 4263 (2001).
48. Y. Wang, M. Vujcic, D. Kowalski, *Mol. Cell. Biol.* **21**, 4938 (2001).
49. C. Santocanale, K. Sharma, J. F. Diffley, *Genes Dev.* **13**, 2360 (1999).
50. J. G. Wohlgemuth, G. H. Bulboacă, M. Moghadam, M. S. Caddle, M. P. Calos, *Mol. Biol. Cell* **5**, 839 (1994).
51. D. Dubey, J. Zhu, D. Carlson, K. Sharma, J. Huberman, *EMBO J.* **13**, 3638 (1994).
52. B. J. Brewer, W. L. Fangman, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3418 (1994).
53. B. J. Brewer, W. L. Fangman, *Science* **262**, 1728 (1993).
54. Y. Marahrens, B. Stillman, *EMBO J.* **13**, 3395 (1994).
55. J. B. Stevenson, D. E. Gottschling, *Genes Dev.* **13**, 146 (1999).
56. J. R. Lipford, S. P. Bell, *Mol. Cell* **7**, 21 (2001).
57. A.-K. Bielinsky et al., *Curr. Biol.*, in press.
58. D. M. Gilbert, H. Miyazawa, M. L. DePamphilis, *Mol. Cell. Biol.* **15**, 2942 (1995).
59. O. V. Ermakova et al., *Mol. Cell* **3**, 321 (1999).
60. R. Y. Chuang, T. J. Kelly, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2656 (1999).
61. V. Boonyaratankornkit et al., *Mol. Cell. Biol.* **18**, 4471 (1998).
62. S. K. Dhar, A. Dutta, *J. Biol. Chem.* **275**, 34983 (2000).
63. B. Chaudhuri, H. Xu, I. Todorov, A. Dutta, J. L. Yates, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10085 (2001).
64. S. K. Dhar et al., *Cell* **106**, 287 (2001).
65. A. Schepers et al., *EMBO J.* **20**, 4588 (2001).
66. E. L. Beall, M. R. Botchan, personal communication.
67. B. J. Brewer, *Curr. Opin. Genet. Dev.* **4**, 196 (1994).
68. S. B. Haase, S. S. Heinzel, M. P. Calos, *Mol. Cell. Biol.* **14**, 2516 (1994).
69. B. J. Brewer, *Cell* **53**, 679 (1988).
70. E. R. Tillier, R. A. Collins, *Nature Genet.* **26**, 195 (2000).
71. B. J. Brewer, D. Lockshon, W. L. Fangman, *Cell* **71**, 267 (1992).
72. J. K. Gerber et al., *Cell* **90**, 559 (1997).
73. A. M. Deshpande, C. S. Newlon, *Science* **272**, 1030 (1996).
74. P. Cook, *Science* **281**, 1466 (1998).
75. B. Ren et al., *Science* **290**, 2306 (2000).
76. D. Pak et al., *Cell* **91**, 311 (1997).
77. M. Loupart, S. Krause, M. S. Heck, *Curr. Biol.* **10**, 1547 (2000).
78. M. F. Pflumm, M. R. Botchan, *Development* **128**, 1697 (2001).

I would like to thank K. Friedman, S. Fiering, M. Aladjem, J. Blow, M. Schmitt, M. Botchan, and S. Gerbi for helpful criticisms of this essay and H. Masukata, S. Gerbi, and M. Botchan for sharing unpublished information. Work in my laboratory is supported by NIH grant GM-57233-01, NSF grant MCB-0077507, and American Cancer Society grant RPG-97-098-04-CCG.

## Enhance your AAAS membership with the Science Online advantage.

- **Full text Science**—research papers and news articles with hyperlinks from citations to related abstracts in other journals before you receive *Science* in the mail.
- **ScienceNOW**—succinct, daily briefings, of the hottest scientific, medical, and technological news.
- **Science's Next Wave**—career advice, topical forums, discussion groups, and expanded news written by today's brightest young scientists across the world.

Science ONLINE

- **Research Alerts**—sends you an e-mail alert every time a *Science* research report comes out in the discipline, or by a specific author, citation, or keyword of your choice.
- **Science's Professional Network**—lists hundreds of job openings and funding sources worldwide that are quickly and easily searchable by discipline, position, organization, and region.
- **Electronic Marketplace**—provides new product information from the world's leading science manufacturers and suppliers, all at a click of your mouse.

All the information you need....in one convenient location.

Visit Science Online at <http://www.scienceonline.org>, call 202-326-6417, or e-mail [membership2@aaas.org](mailto:membership2@aaas.org) for more information.

AAAS is also proud to announce site-wide institutional subscriptions to Science Online. Contact your subscription agent or AAAS for details.

 AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE