channel of fast-flowing ice embedded within the northeast sector of the ice sheet (*5*). Basal melt water, lubricating the interface between the ice and its bed, must be the cause of the enhanced flow.

With the exception of his descriptions of the physics, de Saussure's theories about subglacial water were arguably quite advanced. Water submerges the bed, thereby diminishing the basal drag that resists the gravitational driving stress. It also saturates and weakens subglacial tills, and under pressure, basal water can cause the overlying ice to float (*6–8*). Recently, another attribute of water has been added to the list: Melt water, flowing down the glacier from its source, may redistribute energy within the glacier system. The effectiveness of this redistribution is still under investigation (*9*), but it may be critically important to one of Earth's three remaining great ice masses, the West Antarctic Ice Sheet.

The temperature at the bottom of an ice sheet depends on the air temperature at the ice surface, the rate at which snow accumulates, the thickness of the ice, the geothermal gradient in the rock beneath the ice, and variations arising from horizontal motion of the ice. In general, the thinner the ice, the more likely it is that the geothermal gradient is accommodated by diffusion of heat through the ice and that the ice is frozen to its bed. But in West Antarctica, the fast-flowing ice streams are both thin and melted at the bed. And an increasing range of evidence suggests that this has been the case since at least the Last Glacial Maximum about 20,000 years ago (*10*, *11*). Weak, water-saturated marine sediments beneath the ice sheet are the cause of the fast flow, but this subglacial till would consolidate without sufficient water (*7*, *12*). The apparent contradiction of thin yet fast-flowing ice can only be resolved by recognizing the importance of melt water and its distribution at the base of the ice sheet. Melt-water thermodynamics is thus of great concern to glaciologists trying to predict future ice sheet retreat rates and corresponding changes in global sea level.

The notion of thin, fast-flowing West Antarctic ice streams throughout the Pleistocene (*10*, *11*) has implications beyond Antarctica. Many of the Northern Hemisphere ice sheets of recent glaciations were marine-based as well, and reconstructions of their thickness histories rely in part on fundamental assumptions about basal energy balance that may need reevaluation. Fahnestock *et al.*'s discovery of substantial basal melting in an unexpected place reminds us that basal melt water matters and that it may be important in places we never suspected.

References

1. H. B. de Saussure, *Voyages dans les alpes precèdés d'un essai sur l'histoire naturelle des environs de Genève*, vol. 1 (Saumel Fauche, Neuchâtel, Switzerland, 1779), section 535, pp. 378–379; translation from G. K. C. Clarke, Special issue commemorating the 50th anniversary of the International Glaciological Society, *J. Glaciol.* **2** (1988).
2. I. Joughin *et al.*, *Science* **286**, 283 (1999).
3. D. R. MacAyeal, *Paleoceanography* **8**, 775 (1993).
4. M. Fahnestock, W. Abdalati, I. Joughin, J. Brozena, P. Gogineni, *Science* **294**, 2338 (2001).
5. M. Fahnestock, R. Bindschadler, R. Kwok, K. Jezek, *Science* **262**, 1530 (1993).
6. R. Bindschadler, *J. Glaciol.* **29**, 3 (1983).
7. S. Tulaczyk, B. Kamb, H. F. Engelhardt, *J. Geophys. Res.* **105**, 483 (2000).
8. J. Weertman, *Rev. Geophys. Space Phys.* **10**, 287 (1972).
9. West Antarctic Ice Sheet Initiative Workshop 2001 (see http://igloo.gsfc.nasa.gov/wais).
10. R. P. Ackert *et al.*, *Science* **286**, 276 (1999).
11. J. B. Anderson, S. S. Shipp, *Antarct. Res. Ser.* **77**, 45 (2001).
12. B. Kamb, *Antarct. Res. Ser.* **77**, 157 (2001).

PERSPECTIVES: DNA REPLICATION

Genomic Views of Genome Duplication

Bruce Stillman

Genomewide views of biological processes offer new perspectives on the organization of chromosomes and the replication of their DNA. In a pair of recent *Science* papers (see page 2357 of this issue), two groups report different, but complementary, global approaches for mapping the distribution of origins of DNA replication (*oris*) throughout the genome of the budding yeast *Saccharomyces cerevisiae* (*1*, *2*). DNA replication simultaneously proceeds along discrete sections of the chromosomes, and the *oris* are the DNA sequences where replication begins. Both techniques not only reveal the position of many yeast genome *oris*, but also provide a wealth of information about the way in which the genome is duplicated. The results also hint that the proteins that initiate DNA replication may be important in nonreplication events as well.

Enhanced online at www.sciencemag.org/cgi/content/full/294/5550/2301

The author is at Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. E-mail: stillman@cshl.org

S. cerevisiae was the first eukaryote to have its genome sequenced. Scientists have probed the sophisticated genome of this well-researched eukaryote to learn how whole chromosomes are replicated. The 16 chromosomes of the budding yeast contain multiple *oris*. Two decades ago, small pieces of yeast DNA called autonomously replicating sequences (ARSs), when inserted into bacterial plasmids, were found to promote the replication of these circular DNAs during each S phase of the cell cycle (*3*). Later, many, but not all, of these ARSs were found to correspond to *oris* in yeast chromosomes.

The detailed structures of individual ARSs and *oris* were determined a decade ago. Simultaneously came the discovery of the origin recognition complex (ORC), a six-subunit protein complex that binds to *oris* and coordinates the assembly of a pre-replication complex (pre-RC) at each *ori* (*3*). The pre-RC contains, among other components, the six different, but related, mini-chromosome maintenance (MCM) proteins (*3*) (see the top figure on the next page). We have learned much about how DNA is replicated in the budding yeast (*3*), including the distribution of *oris* in two of its chromosomes (*4*).

Brewer and Fangman and their colleagues teamed up with Davis and his group (*1*) to undertake the first genomewide analysis of *oris*. With high-density DNA microarrays, they determined exactly when each region of the yeast genome replicated (at a resolution of about 10 kilobase pairs) (*1*). To do this, the investigators modified one of the most elegant biology experiments ever undertaken—the famous Meselson and Stahl experiment that proved semiconservative DNA replication (that is, each DNA strand is a template for the new strand being synthesized). Unfortunately, the yeast lacks the enzyme thymidine kinase and hence its DNA cannot be labeled with the modified nucleotide bromodeoxyuridine during replication. So, instead, the investigators prelabeled budding yeast DNA with the isotopes ^{15}N and ^{13}C (the resulting DNA strands being called heavy heavy, HH). Yeast cells arrested in late G_1 phase of the cell cycle were synchronously released from this block in the presence of two other isotopes, ^{14}N and ^{12}C (light medium). At various time points during S phase, DNA was isolated—both DNA that replicated semiconservatively (called heavy light, HL) and DNA that remained unreplicated (HH). By calculating

the percentage of HL DNA at specific locations in the genome using whole-genome DNA microarrays, the sites of *oris* could be identified. Furthermore, it was possible to predict rates of movement of the replication forks (the points where the parent DNA strands separate) and the locations of regions where replicons (units of replication from a single *ori*) terminate. A total of 332 *oris* in the genome were identified with this method.

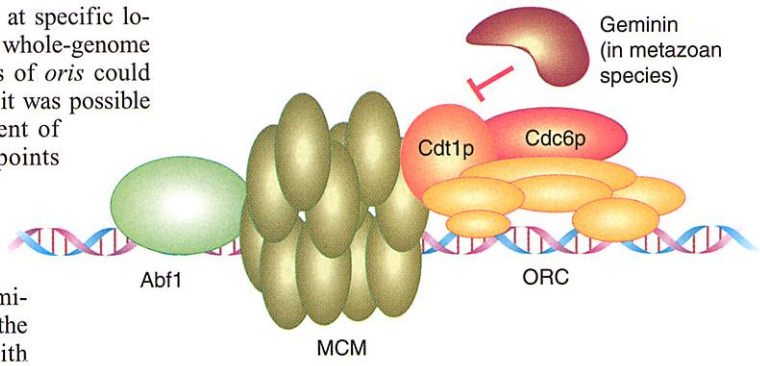
Taking a different tack, Wyrick *et al.* (*2*) identified sites in the yeast genome that bound the ORC or individual MCM proteins. Using a version of the chromatin immunoprecipitation (CHIP) method, short DNA fragments that coprecipitated with antibodies directed against ORC or MCM proteins were mapped to chromosomes with high-density whole-genome microarrays. This strategy was used previously to map the chromosomal locations of sequence-specific DNA binding transcription factors (*5*). Assuming that the simultaneous localization of MCM proteins and the ORC marked the sites of *oris*, these investigators came up with a total of 429 potential *ori* sites, which they called proposed ARSs (pro-ARSs).

Of interest was the finding that 75 *ori* sites bound three different MCM proteins, but binding of the ORC to these sites was not detected. It is possible that antibody access to the ORC was masked at some *oris*—as has been found for ORCs bound to chromatin in G_1 phase (*6*). Alternatively, loading of MCM proteins onto the DNA may be independent of ORCs. We need to know whether other pre-RC proteins are localized at these MCM-bound sites. Binding of MCM proteins at these sites may have little to do with DNA replication, and these proteins may be involved in another capacity such as gene transcription. MCM proteins are known to bind to transcription factors (*7*).

Wyrick *et al.* (*2*) found 12 ORC binding sites that did not correspond to sites where MCM proteins bound. This suggests that the ORCs at these sites may be doing something other than initiating DNA replication. For example, the ORC interacts with transcriptional silencers in yeast, and binds to heterochromatin in the fruit fly (*8*). It would not be surprising if the ORC also binds to heterochromatin at discrete sites in the yeast genome.

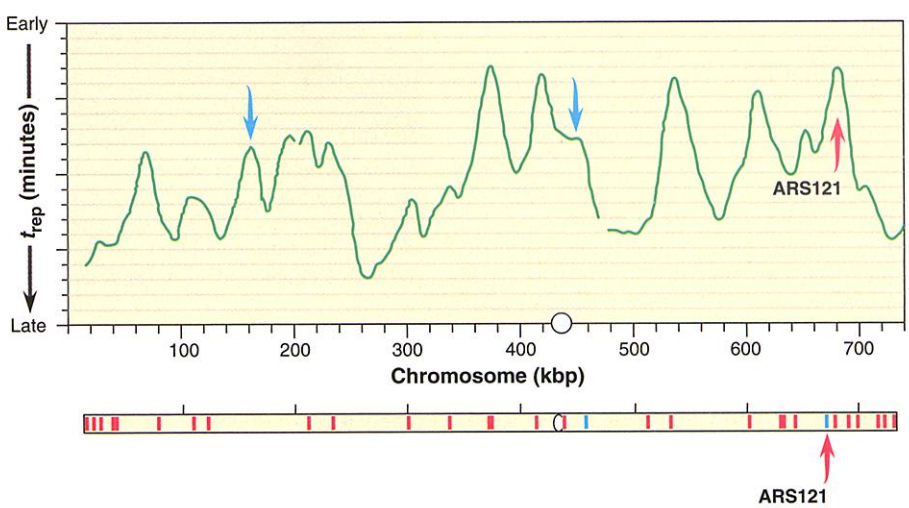
A potential next step would be to measure replication fork movement by performing genomewide CHIP-array experiments with proteins, such as Cdc45p, that associate with *oris* and then to track the replication fork at specific times during S phase (*6*, *9*). Such experiments would complement the replication-fork predictions made with the density shift method (*1*).

How do the two genomewide techniques for predicting the locations of DNA replication origins compare? In an initial analysis, I compared data for yeast chromosome X (ten) from both papers (see the bottom figure). At first glance, the two methods show considerable overlap in the predicted locations of origins of DNA replication. The replication density shift method (*1*) used a window of 10 kilobase pairs and, thus, is expected to miss *oris* that are close together; these would be detected easily by the high-resolution CHIP-array method (*2*). Indeed, many *oris* (pro-ARSs) that are close together—especially those near telomeres (the ends of chromosomes)—do not appear as multiple peaks in the replication timing data. Furthermore, not all pre-RCs "fire" during a single S phase, and thus many *oris* "fire" with weak efficiency within a yeast cell population. Perhaps this explains the greater number of pro-ARS peaks compared with *ori* peaks detected with the replication timing method.



The complexities of duplication. The proteins that form the pre-replication complex (pre-RC) required for the initiation of DNA replication. The ORC binds to origins of replication (*oris*) in the chromosomes and establishes docking sites for the other protein components, such as MCM proteins, of the pre-RC. In metazoan species, geminin, which is degraded during mitosis, inhibits the activity of Cdt1, which is necessary for binding of MCM proteins to the origins of replication.

Even more surprising is the existence of obvious peaks in the replication timing data, which signal early *oris* that are not located anywhere near the pro-ARS peaks of the CHIP-array data (see the blue arrows, bottom figure). Using data from yeast chromosomes III and VI where the *oris* and ARSs have been well characterized, Wyrick *et al.* (*2*) estimated that they detected 88% of the known ARSs. Thus, the *oris* marked with blue arrows (see the bottom figure) might represent the 12% that escaped detection. However, initiation of DNA replication may occur in an ORC- and MCM-independent manner, for example, at sites of induced DNA breaks.



Twin peaks. Comparison of the replication density shift and CHIP-array genomic techniques for identifying potential origins of DNA replication (*oris*) in yeast chromosome X (ten). (**Top**) Graph shows the time of DNA replication (t_{rep}): origins coincide with peaks, termination regions appear as valleys. The earliest firing origins are the highest peaks [adapted from (*1*)]. (**Bottom**) The positions of predicted origins of DNA replication based on fragments that bind to the ORC, to MCM proteins, or to both [adapted from (*2*)]. Red bars represent proposed ARSs (called pro-ARSs) and blue bars indicate known *oris* that were not detected. ARS121 (red arrow), an origin of DNA replication on yeast chromosome X, has been well characterized. Blue arrows highlight potential *oris* detected by the replication density shift method that were not detected by the CHIP-array method.

These anomalies should be investigated further to discern how replication is initiated at these sites.

Both groups fine-tuned their techniques with the well-characterized *S. cerevisiae* genome, and both used knowledge of existing *oris* to enhance their data analysis. Detailed information about the replication of yeast chromosomes III and VI made this possible. How, then, might these technologies be applied to discover *oris* in, for example, human chromosomes where the location of *oris* has not been well researched? It is clear that neither technique alone would convincingly point to the location of *oris* in a naïve genome, but when combined, sites of overlap would strongly suggest (but not prove) potential *ori* locations. For the chromosomes of metazoan species, we need better techniques to map the genetic elements that determine where *oris* occur, and better antibody reagents for chromatin precipitation.

Recently, CHIP experiments with anti-ORC and anti-MCM antibodies suggested the locations of pre-RC components near known sites of replication initiation in the Epstein-Barr virus genome (*10*). For mammalian genomes, ordered bacterial artificial chromosome (BAC) arrays or representation arrays will prove useful for mapping *oris* with both the replication density shift and CHIP-array methods.

References

1. M. K. Raghuraman *et al.*, *Science* **294**, 115 (2001).
2. J. J. Wyrick *et al.*, *Science* **294**, 2357 (2001).
3. T. J. Kelly, G. W. Brown, *Annu. Rev. Biochem.* **69**, 829 (2000).
4. A. Poloumienko *et al.*, *Mol. Biol. Cell* **12**, 3317 (2001).
5. B. Ren *et al.*, *Science* **290**, 2306 (2000).
6. O. M. Aparicio, D. M. Weinstein, S. P. Bell, *Cell* **91**, 59 (1997).
7. J. J. Zhang *et al.*, *EMBO J.* **17**, 6963 (1998).
8. D. T. Pak *et al.*, *Cell* **91**, 311 (1997).
9. L. Zou, B. Stillman, *Science* **280**, 593 (1998).
10. B. Chaudhuri *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10085 (2001).

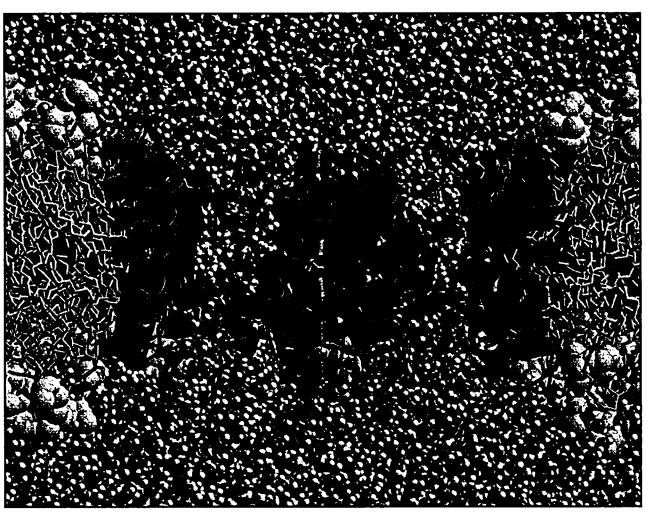
PERSPECTIVES: BIOINFORMATICS

Reality Simulation—Observe While It Happens

Herman J. C. Berendsen

Since the first simulation of the dynamics of a small protein in 1976 (*1*, *2*), considerable insights have been gained into the dynamics of biological macromolecules and membranes on time scales of nanoseconds. But progress toward simulations on biologically relevant time scales has been slow (*3*, *4*). Decades of method development, together with rapid growth in computational power, are now beginning to enable simulation of complex biological processes on realistic time scales, as demonstrated by de Groot and Grubmüller on page 2353 of this issue (*5*).

Biomolecular processes present three problems for simulation on the atomic level. First, they involve hundreds of thousands of atoms, often in intricate interactions that are difficult to simplify. Second, they span a wide range of time scales: Primary events (for example, in vision or photosynthesis) occur within picoseconds, enzymatic and regulatory processes take milliseconds, and protein folding and structural reorganizations may exceed seconds. Third, the small driving forces that cause molecular changes result from large, opposing energetic effects. This requires careful fine-tuning of the force fields that describe interatomic interactions. Current simulations are limited to system sizes of about 100,000 atoms, time scales of about 100 ns, and classical dynamics with simple, pair-additive, interactions. Quantum mechanics can also be incorporated, but at the expense of system size or time scale.



Water permeation through aquaporin-1. From a 10-ns molecular dynamics simulation (*5*). During the simulation, water molecules (red/white, shown as an overlay of 100 snapshots) permeate the four pores of the tetramer (blue), which is embedded in a lipid bilayer (yellow: hydrophilic head groups; green: hydrophobic tails).

The future will see rapid progress in all three areas. Within existing limitations, however, most biological processes cannot be simulated in "real time." Even for processes that do occur on a 100-ns time scale, it is not sufficient to simulate an event—say, the folding of a peptide into a helix—once. Statistics must be collected under varying conditions. In practice, one can either simulate processes that do occur on the time scale attainable for simulations with full atomic detail, or one can resort to approximations in system size, detail of description, or accuracy of dynamics. The latter approach is the only possible way to proceed for most processes but does involve approximations that need thorough testing. There is also a danger that the chosen approximations bias the results.

De Groot and Grubmüller describe a state-of-the-art example of the first approach (*5*). They have simulated a biological process that occurs on the nanosecond time scale: the motion of water through the water-specific membrane-channel protein, aquaporin-1. Aquaporins are essential for maintaining osmotic balance in cells. They are embedded in cell membranes and transport water molecules at rates of more than 10^9 molecules per second. Other molecules or ions, including hydronium ions and protons, cannot pass through the channel. Because of the high permeation rate, several complete permeation events can be observed in a simulation.

De Groot and Grubmüller's system consists of about 100,000 atoms, which form an aquaporin tetramer embedded in a bilayer of lipid molecules and surrounded by water molecules (see the figure). They use a full lattice summation method for computing electrostatic interactions, which is computationally expensive but essential for reliable results, and use one of the fastest (if not the fastest) molecular dynamics programs for biomolecular simulation (*6*).

The authors observe 16 full permeation events within 10 ns, in reasonable agree-

The author is emeritus professor at the Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 4, 9747 AG Groningen, Netherlands. E-mail: berendsen@chem.rug.nl