# How *Xenopus Laevis* Replicates DNA Reliably even though Its Origins of Replication are Located and Initiated Stochastically

John Bechhoefer* and Brandon Marshall
*Department of Physics, Simon Fraser University, Burnaby, B.C., V5A 1S6, Canada*

enough. In fact, the situation is more subtle. Even when all origins are initiated at the beginning of *S* phase, it is possible to replicate with arbitrary reliability simply by having enough origins. While it is true that there will be a few unusually long gaps that will set the replication time, these gaps may be reduced arbitrarily if one starts with enough replication origins. We thus propose an alternate way of viewing the random-completion problem: Instead of fixing the number of origins and looking at the replication times for different strategies, we fix a time $^{**}$ at which either a cell has finished replication or it dies. Since evolution selects on the basis of mortality, the replication parameters [$I(\ )$, , the number of potential origins, etc.] should be a consequence of this selection, and not vice versa. Choosing $^{**}$ to be the cell-cycle time (25) minutes and allowing a failure rate of $10^{-}$, we calculate, for various forms of $I(\ )$, the replication parameters required to meet the reliability constraint. (Our results depend only logarithmically on the failure rate.)

In order to compare with experiment, we must confront a further problem. While the *in vivo* replication time is estimated to be 20 minutes, the *in vitro* experiments require nearly twice this time to replicate. We must thus make additional assumptions to translate the *in vitro* experimental results to the *in vivo* situation. In fact, we can do this with one simple assumption. In earlier studies, it was assumed that the replication fork velocity is constant throughout *S* phase. The original analysis of the *in vitro* *Xenopus* data thus estimated an average forkv5nueh5culat.o5a9903basi da0F3 1 scn(S)97333D[(v5]it)-(add.1(celln)-16 -)hof

In Fig. 4, we summarize the results of these investigations. The dashed line at the top gives the fork density required to make the delta-function $I(\ )$ meet the reliability constraint. The solid curve represents the fork density required for power-law initiations. As we anticipated, the curve has a minimum (between $=1$ and 2). The fine-dashed line, which lies close to the minimum value of the power-law case, is the experimental maximum fork density [12]. Finally, the broad-dashed line gives the optimal fork density ($1/\ ^{**}$).

Although the optimal fork density is lower than that observed, it clearly does not represent a physiologically possible case. It is unrealistic to expect the perfect coordination implied by the delta function at the beginning of $S$ phase. More serious, at the end of $S$ phase, Eq. (7) implies that the rate of initiation diverges, along with the total number of activated origins. Still, we note that the qualitative shape of the curve shares the quadratically increasing form of the experimental result. More generally, it would be surprising if the initiation program were identical to the optimum (even if one were to limit the space of functions to those that are physiologically achievable). We note that the minimum is clearly broad: there is little difference in required fork density between a linear and a quadratic $I(\ )$. The main point is that there are some strategies—most notably the initiation of all origins at the beginning of $S$ phase—that are clearly bad, and these differ from the observed $I(\ )$.

In conclusion, we have calculated the distribution of replication times $\rho$ for the stochastic limit of replication, where origins are placed randomly and initiate stochastically at a rate $I(\ )$. Choosing an $I(\ )$ that increases with time narrows $\rho$ and increases the reliability of replication. Using the known mortality rates and length of the cell cycle, we gave a quantitative interpretation to the random-completion problem and showed that one can meet the reliability constraint using an arbitrary $I(\ )$. Different $I(\ )$ functions demand different resources from the cell. Measuring this resource use by the maximum required fork density, we show that the experimentally observed form of $I(\ )$ is close to optimum. In the future, it would be interesting to consider the effects of any regularity in origin spacing. While we have shown that reliable replication may be achieved even in the worst case of random spacing of origins, there is evidence for some regularity. It would also be interesting to measure the replication-time distribution directly. While determining the time at which the last base (of $\times 10$ ) replicates is unrealistic, one might be able to determine when a given fraction (e.g., 90 or 95%) of origins have replicated. It is straightforward to generalize the methods presented here to determine the distribution of times required to reach a given replication fraction.

We thank O. Hyrien, N. Rhind, R. Harland, J. Herrick, and S. Jun for helpful discussions. This work was supported by NSERC (Canada) and the Univ. de Rennes, 1 (France). J. B. thanks Z. Gueroui for the invitation to Rennes.

*Electronic mail: johnb@sfu.ca

[1] J. J. Blow, EMBO J. **20**, 3293 (2001).
[2] O. Hyrien and M. Méchali, EMBO J. **12**, 4511 (1993).
[3] J. J. Blow, P. J. Gillespie, D. Francis, and D. A. Jackson, J. Cell Biol. **152**, 15 (2001).
[4] O. Hyrien, K. Marheineke, and A. Goldar, Bioessays (1984-) / BioEssays **25**, 116 (2003).
[5] T. A. Prokhorova, K. Mowrer, C. H. Gilbert, and J. C. Walter, Proc. Natl. Acad. Sci. U.S.A. **100**, 13241 (2003).