

Solution to the 50-year-old Okazaki-fragment problem

Peter M. Burgers^{a,1}

The antiparallel structure of double-stranded DNA, together with the known 5'→3' directionality of DNA polymerases, necessitates that the two DNA strands are replicated in opposite directions. The leading strand is synthesized in the same direction as the replication fork, whereas the lagging strand is replicated in the opposite direction. In a 1968 paper in PNAS, Reiji and Tsuneko Okazaki and colleagues (1) proposed that the lagging strand is replicated discontinuously in the form of small fragments that subsequently are matured into one continuous strand. A review of the current state of molecular biology published that year (2) named these small fragments "Okazaki fragments," as they have been called since. The seminal studies of Okazaki et al. (1) generated the textbook model of the semidiscontinuous replication fork, with a continuous leading strand and a discontinuous lagging strand. However, their original experimental results were not in accord with this model, and their studies suggested that all nascent DNA fragments are small. Was the leading strand also synthesized discontinuously, as was depicted in figure 1 of ref. 1 (Fig. 1C)? Fifty years after the landmark paper, this question has been answered by Cronan et al. (3), who show that the leading strand is indeed replicated continuously. However, this strand is fragmented due to ribonucleotide excision repair (RER) (4, 5). RER removes genomic ribonucleotides that are erroneously inserted by replicative DNA polymerases (Fig. 2A).

To detect the earliest lagging-strand pieces synthesized, before they could be matured into continuous DNA, Okazaki et al. (1) used pulse-labeling techniques in *Escherichia coli*, with ³H-thymidine pulses as brief as 5 s. Given that the DNA replication rate is 500 to 1,000 nt/s and an average Okazaki fragment is 1 to 2 kb in size, it only takes a few seconds to synthesize a single Okazaki fragment. The pulse-labeled DNA was denatured and size-fractionated on an alkaline sucrose gradient. Surprisingly, even though biochemical studies at the time suggested that the leading strand could be synthesized in a continuous fashion (6), virtually all of the labeled DNA

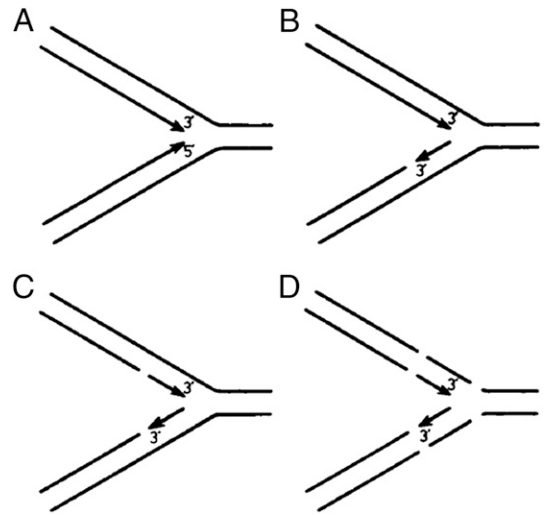


Fig. 1. (A–D) Models for the possible structure and reaction in the replicating region of DNA. Reprinted with permission from ref. 1.

fractionated near the top of the gradient, with a DNA length of 1 to 2 kb. Only a small fraction of the label appeared as a shoulder with a size of up to 10 kb, similarly to the result shown in Fig. 2B (black curve). Obviously, the data supported the model in Fig. 1C more than the idealized model in Fig. 1B for semidiscontinuous DNA replication. These curious results were not just an *E. coli* oddity or artifact; they extended to the gram-positive bacterium *Bacillus subtilis* and to bacteriophage T4 (7).

Over the next few years, the Okazaki team introduced several improvements and gained more advanced knowledge of the small fragments that were named after them. They showed that the pulse-labeled small fragments could be chased into large fragments upon addition of an excess of cold thymidine, indicating that they were true intermediates in the synthesis of chromosomal DNA. They discovered that the synthesis of Okazaki fragments is primed with a short RNA primer (8). They also showed that the

^aDepartment of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110

Author contributions: P.M.B. wrote the paper.

The author declares no conflict of interest.

Published under the PNAS license.

See companion article on page 1251 in issue 4 of volume 116.

¹Email: burgers@wustl.edu.

Published online February 15, 2019.

strand, and 90% of the large DNA mapped to the leading strand. These data are consistent with an earlier study by Okazaki et al. (16), in which a mutation in DNA polymerase I caused a defect in the joining of all small nascent fragments, because *E. coli* DNA polymerase I is required for RER as well as Okazaki fragment maturation (5).

The new study by Cronan et al. (3) shows that the incorporation of ribonucleotides and their subsequent excision by RER is a frequent event. Cronan et al. assess the frequency of this event by measuring the number of ribonucleotides present in a plasmid propagated in an *rnhB* strain. From the fraction of the isolated plasmid that was nicked by RNase H2 and using the reasonable assumption that ribonucleotide insertion is random, Cronan et al. use the Poisson equation to calculate the misinsertion frequency at one ribonucleotide per 8 to 9 kb, or about 500 per genome. By 1 order of magnitude, RER in bacteria appears to be the most frequently employed DNA repair pathway, as it is in eukaryotes.

While it is evident from the current study that the leading strand is synthesized largely continuously, the question still

remains as to how continuous. Replication forks tend to stall for a variety of reasons, and cells have several mechanisms to restart replication forks (17). Some of these restart mechanisms use the available 3' terminus for restart, after recombination or fork remodeling, but a de novo restart by RNA priming is also a possible mechanism. The latter would show up as a fragmentation event on the leading strand. In their study of the size distributions of leading strands in the multiple-pathway excision repair-defective strain (RER⁻ BER⁻ NER⁻ MMR⁻), Cronan et al. (3) provide evidence that, indeed, the leading strand is not completely continuous. Whether these discontinuities are the result of still another repair pathway, the result of replication-fork restart, or the result of the stochastic, unprovoked activity of DNA primase on the leading strand needs further investigation.

Acknowledgments

I thank Andrei Kuzminov and colleagues for sharing their results prior to publication. This research is supported by the National Institutes of Health Grant R35-GM118129.

- 1 Okazaki R, Okazaki T, Sakabe K, Sugimoto K, Sugino A (1968) Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc Natl Acad Sci USA* 59:598–605.
- 2 Anonymous (1968) Molecular biology comes of age. *Nature* 219:825–829.
- 3 Cronan GE, Kouzminova EA, Kuzminov A (2019) Near-continuously synthesized leading strands in *Escherichia coli* are broken by ribonucleotide excision. *Proc Natl Acad Sci USA* 116:1251–1260.
- 4 Sparks JL, et al. (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* 47:980–986.
- 5 Vaisman A, et al. (2014) Investigating the mechanisms of ribonucleotide excision repair in *Escherichia coli*. *Mutat Res* 761:21–33.
- 6 Goulian M, Kornberg A, Sinsheimer RL (1967) Enzymatic synthesis of DNA, XXIV. Synthesis of infectious phage phi-X174 DNA. *Proc Natl Acad Sci USA* 58:2321–2328.
- 7 Okazaki R, et al. (1968) In vivo mechanism of DNA chain growth. *Cold Spring Harb Symp Quant Biol* 33:129–143.
- 8 Sugino A, Hirose S, Okazaki R (1972) RNA-linked nascent DNA fragments in *Escherichia coli*. *Proc Natl Acad Sci USA* 69:1863–1867.
- 9 Kitani T, Yoda K, Okazaki T (1984) Discontinuous DNA replication of *Drosophila melanogaster* is primed by octaribonucleotide primer. *Mol Cell Biol* 4:1591–1596.
- 10 Okazaki T (2017) Days weaving the lagging strand synthesis of DNA—A personal recollection of the discovery of Okazaki fragments and studies on discontinuous replication mechanism. *Proc Jpn Acad, Ser B, Phys Biol Sci* 93:322–338.
- 11 Tye BK, Chien J, Lehman IR, Duncan BK, Warner HR (1978) Uracil incorporation: A source of pulse-labeled DNA fragments in the replication of the *Escherichia coli* chromosome. *Proc Natl Acad Sci USA* 75:233–237.
- 12 Amado L, Kuzminov A (2013) Low-molecular-weight DNA replication intermediates in *Escherichia coli*: Mechanism of formation and strand specificity. *J Mol Biol* 425:4177–4191.
- 13 Berg P, Fancher H, Chamberlin M (1963) The synthesis of mixed polynucleotides containing ribo- and deoxyribonucleotides by purified preparations of DNA polymerase from *Escherichia coli*. *Informational Macromolecules*, eds Vogel HJ, Bryson V, Lampen JO (Academic Press, New York), pp 467–483.
- 14 Nick McElhinny SA, et al. (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci USA* 107:4949–4954.
- 15 Buckstein MH, He J, Rubin H (2008) Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *J Bacteriol* 190:718–726.
- 16 Okazaki R, Arisawa M, Sugino A (1971) Slow joining of newly replicated DNA chains in DNA polymerase I-deficient *Escherichia coli* mutants. *Proc Natl Acad Sci USA* 68:2954–2957.
- 17 Marians KJ (2018) Lesion bypass and the reactivation of stalled replication forks. *Annu Rev Biochem* 87:217–238.