Published in final edited form as:

Science. 2019 February 22; 363(6429): 814–815. doi:10.1126/science.aaw6265.

DNA replication from two different worlds

Huilin Li¹ and Michael E. O'Donnell²

¹Structural Biology Program, Van Andel Research Institute, Grand Rapids, Michigan, USA

²The Rockefeller University and HHMI, New York, New York, USA

Replication of the DNA genome is performed by a replisome composed of numerous proteins that function as a complex. An article in the current issue reveals an atomic model of the replisome of T7 bacteriophage using cryoEM (1). T7 phage is a smart choice for EM studies because it is the most streamlined replisome known (2). A functional replisome requires, at minimum, a helicase to unwind the duplex, two copies of DNA polymerase (Pol) for the two DNA strands, and a primase that forms RNA primers for DNA polymerases to extend. In T7 phage the helicase, gene protein 4 (gp4), forms a hexamer ring, a common feature of all replicative helicases (3, 4). However, unlike other helicases the T7 gp4 helicase contains a domain with the primase activity located behind the helicase during replication fork progression (2). The T7 DNA polymerase, gp5, forms a 1:1 complex with the 12-kDa bacterial host thioredoxin that increases processivity (2). Hence, only three different proteins comprise the T7 replisome. In contrast, the replisome of *Escherichia coli* - the bacterial host for T7 - contains a dozen different proteins, including a sliding clamp that tethers it to DNA and a clamp loader pentamer that loads clamps onto DNA (5). Despite the different complexity, the proteins of the T7 replisome are homologous to proteins of its *E. coli* host, and thus are representative of the bacterial core replisome.

Multi-protein complexes carry out each step of the “central dogma” of genetic information flow: 1) The replisome complex duplicates DNA, 2) the RNA polymerase holoenzyme performs transcription of DNA to RNA, and 3) the ribosome performs translation of RNA to proteins. The multiple component sequences of ribosomes (mostly RNA) and of RNA polymerase holoenzyme are homologous among Bacteria, Archaea, and Eukarya, and thus evolved from a common ancestor cell. In contrast, the polymerase, helicase, and primase of bacteria share no homology to their eukaryotic counterparts, implying that these replisome enzymes evolved twice, independently, after the evolutionary split of bacteria and eukaryotes (6, 7). The primordial cell possibly used a simpler process of replication, or used an RNA genome. Present day cells from all three domains of life, have duplex DNA genomes and a replisome that duplicates both strands simultaneously, a very complex task considering all cells use 5' activated dNTPs which imposes 3'–5' unidirectional elongation by DNA polymerases. The difficulty lies in the DNA structure. The two strands of DNA are antiparallel, and while one strand (leading strand) can be made continuously, the antiparallel strand (i.e. the lagging strand) is made discontinuously as a series of fragments, the Okazaki fragments. This “semi-discontinuous replication” is shared by all cells.

The report in this issue on the T7 replisome cryoEM structure (1) reveals the arrangement of the two Pols in a single replisome, in which they sandwich the DNA helicase, is asymmetric;

one Pol is on top of the helicase and one Pol is below (see Figure). This architecture is unlike illustrations in textbooks that show both polymerases behind the helicase. Surprisingly the asymmetric arrangement of two Pols that sandwich the DNA helicase was also demonstrated by EM for the eukaryotic replisome (8), albeit at lower resolution than the T7 study. Thus while “worlds apart” in terms of their independent evolution, the core elements of bacterial and eukaryotic replisomes both contain a helicase between two polymerases, although the eukaryotic replisome requires a trimeric scaffolding factor (Ctf4) to help tether the top polymerase to the helicase.

Further evidence, besides non-homologous sequences of core enzymes, that replisomes of bacteria and eukaryotes followed distinct evolutionary paths is the fact that replicative helicases of bacteria and eukaryotes travel in opposite directions on DNA (4, 5). Thus, bacterial helicases like T7 gp4 encircle the lagging strand and translocate 5'–3', while the eukaryotic helicase, referred to as CMG by its founder (9), encircles the leading strand and translocates 3'–5'. Because all replicative helicases bind DNA and translocate in the same direction, the opposing translocation directionality must have arisen from the opposite ATP hydrolysis sequence (clockwise vs anti-clockwise) in the helicase rings (10, 11). In a most interesting twist, the top Pol functions on the opposite strand in the two cell types: the polymerase residing at the top of the T7 helicase replicates the leading strand, while the polymerase at the top of eukaryotic CMG helicase replicates the lagging strand (see illustration). It remains a mystery why this “mirror” arrangement evolved, but they both share a pragmatic logic for replisome function. All replicative helicases split the duplex at their leading edge, with one strand going through the middle of the helicase ring and the other strand deflected off the top of the ring (4). Hence, the presence of a polymerase at the top of the helicase provides immediate action in duplicating the separated strand at the top of the helicase.

In bacteria, the deflected strand at the top of the helicase is the leading strand, extended continuously by the top polymerase. The primase domain of gp4 helicase-primase is below the helicase and hands the primer to the lagging strand gp5 polymerase for Okazaki fragment extension. In eukaryotes the deflected strand is the antiparallel lagging strand, requiring priming by Pol α -primase, a four-subunit enzyme containing a primase and DNA polymerase that is held more firmly to the helicase by a scaffolding trimer called Ctf4 (8, 12–14). Pol α -primase generates a RNA/DNA primer of about 25 nucleotides and then hands the primer to the lagging strand polymerase δ (13). In eukaryotes the leading and lagging polymerases are encoded by different genes, and the leading strand DNA polymerase ϵ (13) is located at the bottom of the helicase (8).

The above description is a gross oversimplification of replisome action, and has only focused on the relative organization of polymerases, primase and helicase. In reality, replisomes are amazingly complex, even the stripped down T7 replisome, not to mention the three dozen or more proteins that replicate eukaryotic cellular chromosomes. For additional information, the reader is directed to recent reviews (3, 12–15), and the article on the T7 replisome in this issue of *Science* (1).

References

1. Gao Y et al., Operating principles of a replisome revealed by cryo-EM. *Science*, (2019).
2. Kulczyk AW, Richardson CC, The Replication System of Bacteriophage T7. *Enzymes* 39, 89–136 (2016). [PubMed: 27241928]
3. Bleichert F, Botchan MR, Berger JM, Mechanisms for initiating cellular DNA replication. *Science* 355, (2017).
4. O'Donnell ME, Li H, The ring-shaped hexameric helicases that function at DNA replication forks. *Nat Struct Mol Biol* 25, 122–130 (2018). [PubMed: 29379175]
5. O'Donnell M, Langston L, Stillman B, Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb Perspect Biol* 5, :a010108 (2013). [PubMed: 23818497]
6. Leipe DD, Aravind L, Koonin EV, Did DNA replication evolve twice independently? *Nucleic Acids Res* 27, 3389–3401 (1999). [PubMed: 10446225]
7. Yao NY, O'Donnell ME, Evolution of replication machines. *Crit Rev Biochem Mol Biol* 51, 135–149 (2016). [PubMed: 27160337]
8. Sun J et al., The architecture of a eukaryotic replisome. *Nat Struct Mol Biol* 22, 976–982 (2015). [PubMed: 26524492]
9. Moyer SE, Lewis PW, Botchan MR, Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A* 103, 10236–10241 (2006). [PubMed: 16798881]
10. Georgescu R et al., Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proc Natl Acad Sci U S A* 114, E697–E706 (2017). [PubMed: 28096349]
11. Thomsen ND, Berger JM, Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* 139, 523–534 (2009). [PubMed: 19879839]
12. Bell SP, Labib K, Chromosome Duplication in *Saccharomyces cerevisiae*. *Genetics* 203, 1027–1067 (2016). [PubMed: 27384026]
13. Burgers PMJ, Kunkel TA, Eukaryotic DNA Replication Fork. *Annu Rev Biochem* 86, 417–438 (2017). [PubMed: 28301743]
14. Pellegrini L, Costa A, New Insights into the Mechanism of DNA Duplication by the Eukaryotic Replisome. *Trends Biochem Sci* 41, 859–871 (2016). [PubMed: 27555051]
15. On KF, Jaremko M, Stillman B, Joshua-Tor L, A structural view of the initiators for chromosome replication. *Curr Opin Struct Biol* 53, 131–139 (2018). [PubMed: 30218786]

Figure. Asymmetric organization of core enzymes in replisomes.

Replisomes require the core enzymes, helicase, primase and DNA polymerases. While these are the only proteins in a T17 replisome, Cellular replisomes of both bacteria and eukaryotes require many more proteins not shown here. The figure illustrates the juxtaposition of these core enzymes in T7 phage (left) and the eukaryote, *Saccharomyces cerevisiae* (right). In eukaryotes the Ctf4 trimer (grey) helps the top Pol to associate with the CMG helicase. See text for details.