Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Bacterial replisomes

Zhi-Qiang Xu and Nicholas E Dixon

Bacterial replisomes are dynamic multiprotein DNA replication machines that are inherently difficult for structural studies. However, breakthroughs continue to come. The structures of *Escherichia coli* DNA polymerase III (core)–clamp–DNA subcomplexes solved by single-particle cryo-electron microscopy in both polymerization and proofreading modes and the discovery of the stochastic nature of the bacterial replisomes represent notable progress. The structures reveal an intricate interaction network in the polymerase–clamp subassembly, providing insights on how replisomes may work. Meantime, ensemble and single-molecule functional assays and fluorescence microscopy show that the bacterial replisomes can work in a decoupled and uncoordinated way, with polymerases quickly exchanging and both leading-strand and lagging-strand polymerases and the helicase working independently, contradictory to the elegant textbook view of a highly coordinated machine.

Address
Molecular Horizons and School of Chemistry and Molecular Bioscience, University of Wollongong, and Illawarra Health and Medical Research Institute, Wollongong, New South Wales 2522, Australia

Corresponding author: Dixon, Nicholas E (nickd@uow.edu.au)

Current Opinion in Structural Biology 2018, **53**:159–168

This review comes from a themed issue on **Catalysis and regulation**

Edited by **Hazel Holden** and **Alice Vrielink**

https://doi.org/10.1016/j.sbi.2018.09.006

Introduction

Genetic information of living organisms is stored in chromosomal DNA. To faithfully pass it on to the next generation, it is essential that DNA be copied with high efficiency and fidelity. All organisms from bacteria to humans use complex multi-protein molecular machines, called the replisomes, to achieve this feat. Although general functions and mechanisms of replisomes from different domains of life are similar, the components and mechanistic details can be distinct. Here, we focus on bacterial replisomes, particularly that from *Escherichia coli*.

Bacterial DNA replication can be divided into three stages: initiation, elongation and termination. Each stage requires a different set of proteins with highly coordinated activities [1]. The details of each stage as well as recent insights into the structures and functions of protein components or subcomplexes are discussed separately.

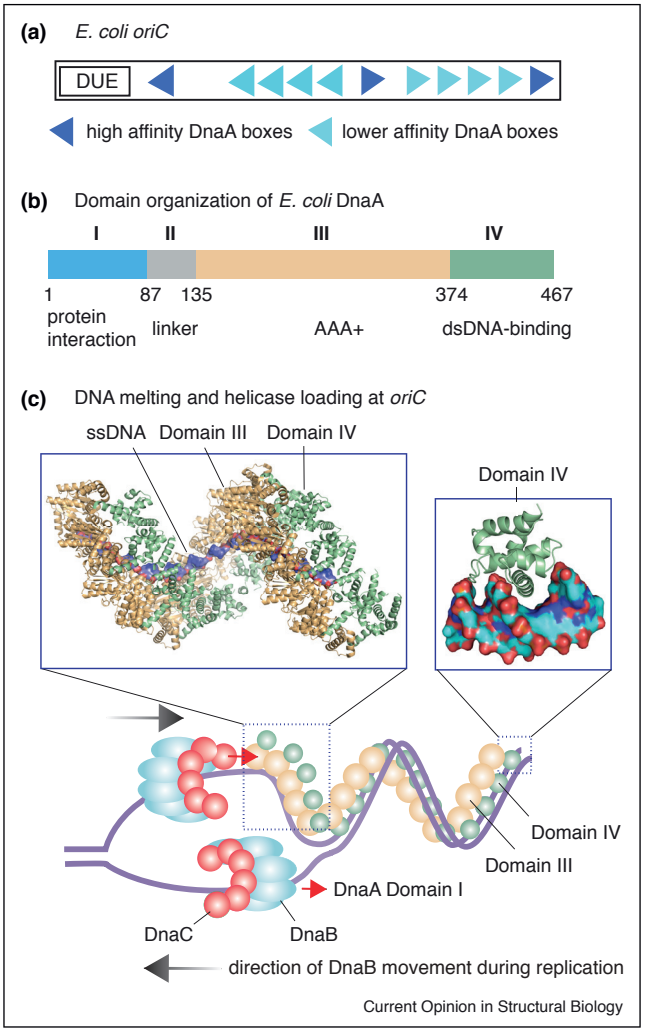
Initiation of DNA replication

Initiation of bacterial DNA replication is tightly controlled to ensure that the chromosome is duplicated once every cell division. Bacterial chromosomes are usually circular doubled-stranded (ds) DNA molecules with a single initiation locus called the replication origin, *oriC*. The *E. coli* chromosome is 4.6 Mb in size with a 250-bp *oriC*. Although there are significant variations in the length and organization of origins in different bacterial species, they are generally comprised of an array of 'DnaA boxes' for origin recognition by the initiator protein DnaA, together with an adjacent AT-rich DNA unwinding element (DUE) for strand separation [2] (Figure 1a). Recently, a string of repeating trinucleotides (5'-TA$^{G}/_{A}$) in the DNA unwinding region, termed DnaA-trio, was identified as an important element [3•].

DnaA has four domains (Figure 1b). The protein interaction domain 1 interacts with protein partners, including the replicative helicase DnaB, and domain 2 is a flexible linker. Domain 3 is the AAA+ ATPase domain, which mediates DnaA oligomerization and binding to single-stranded (ss) DNA [4]. Domain 4 is the dsDNA-recognition domain that binds to DnaA boxes via a helix-turn-helix motif [5] (Figure 1c). Both ATP-bound and ADP-bound DnaA can bind to high-affinity DnaA boxes, but only ATP-DnaA binds to lower affinity boxes and oligomerizes to form a helical filament on *oriC* [6,7] (Figure 1a, c). DNA wrapping around the DnaA filament causes torsional strain in the DUE, contributing to DNA melting [8,9]. The DnaA filament then extends beyond the DnaA boxes with the AAA+ domain interacting with DnaA-trio. This sequesters and stretches one strand of the DUE, facilitating DNA melting and bubble formation for helicase loading [4] (Figure 1c).

After forming a DNA bubble, two DnaB helicase hexamers are recruited and loaded onto each of the separated ssDNA strands as $DnaB_6$–$(DnaC)_6$ complexes. Binding of the helicase loader DnaC inhibits the ATPase and helicase activities of DnaB and traps it like an open right-handed lockwasher, ready to be loaded onto ssDNA [10,11]. DnaC is a homolog of DnaA. Its AAA+ domain interacts with the AAA+ domain of DnaA at the end of the filament and serves as an adaptor to load one DnaB–DnaC

Figure 1



Schematic representation of the initiation of bacterial DNA replication. **(a)** *E. coli oriC*, showing DnaA boxes and the AT-rich DNA unwinding element (DUE). The DnaA boxes contain 9 base pairs with consensus sequence 5'-TTATNCACA (6). The high-affinity DnaA boxes are colored in dark blue and lower affinity boxes in light blue. **(b)** Domain organization of *E. coli* DnaA replication initiator protein. **(c)** DNA melting at *oriC* and loading of the $DnaB_6$–$(DnaC)_6$ helicase–loader complex onto the DNA bubble. Lower schematic: ATP-bound DnaA binds to DnaA-boxes via Domain IV, thereby promoting dsDNA to wrap around the DnaA filament, causing torsional strain to the dsDNA [8,9]. Meantime, Domain III of DnaA binds to one of the two ssDNA strands of DUE and stretches the strand. These interactions cause the AT-rich DUE to melt, forming a bubble [4]. At the same time, binding of DnaC traps DnaB like an open lockwasher, to enable its loading onto ssDNA [10]. DnaC interacts with DnaA at the end of the filament and serves as an adaptor to load one DnaB–DnaC complex [12]. It is not known if closing of DnaB around ssDNA to form a hexameric ring occurs before or concomitantly with dissociation of DnaC. Domain I of DnaA interacts with the N-terminal domain of DnaB, helping to load another DnaB–DnaC on the complementary strand [2]. Upper insets: The helical filament of DnaA formed by Domains III (light orange) and IV (pale green) of *Aquifex aeolicus* DnaA (PDB: 3R8F [4]) and Domain IV of *E. coli* DnaA (pale green) bound to dsDNA (PDB: 1J1V [5]). The ssDNA binds in the middle of the DnaA filament via interactions with the AAA+ Domain III of DnaA.

complex onto the strand that DnaA is stretching [12]. Domain I of DnaA interacts with DnaB of the other DnaB–DnaC complex, helping to load it on the complementary strand [2] (Figure 1c).

In Gram-positive bacteria, such as *Bacillus subtilis*, the hexameric replicative helicase DnaC (counterpart of DnaB) is believed to be assembled from individual subunits with the assistance of the helicase loader DnaI and two others proteins, DnaD and DnaB [6]. In *Helicobacter pylori*, a bacterium with no identified helicase loader, DnaB assembles as a head-to-head double hexamer, which later separates into two hexameric helicases [13].

Next, the DnaG primase interacts with the N-terminal collar of $DnaB_6$, stimulating DnaC dissociation [14]. The two DnaB hexamers later move to the apices of the bubble to form two replication forks moving in opposite directions. DnaG recognizes specific priming sites (preferentially 5'-CTG) to produce a leading-strand RNA primer for DNA elongation, and to repeatedly prime Okazaki-fragment (OF) synthesis on the lagging strand.

Elongation stage of DNA replication

DNA contains two antiparallel strands that have been thought to be replicated simultaneously by the same replisome. The leading strand is replicated continuously, while the lagging strand is synthesized as short Okazaki fragments. RNA primers of OFs are replaced by DNA by gap filling and nick translation by DNA polymerase I, and the nicks are sealed by DNA ligase.

In *E. coli*, the major replicative polymerase is the Pol III holoenzyme (HE) comprised of 10 different proteins organized into three functionally distinct but physically interconnected assemblies: the $\alpha\varepsilon\theta$ polymerase core, the β_2 sliding clamp and the $\delta\tau_n\gamma_{3-n}\delta'\psi\chi$ clamp loader complex [1] (Figure 2a). In the polymerase core, α is the polymerase subunit, ε the 3'–5' proofreading exonuclease and θ is a small subunit that stabilizes ε. After a RNA primer is made by DnaG, the β_2 clamp is loaded onto the primer terminus by the clamp loader. The α and ε subunits separately bind the clamp, each via a short linear clamp-binding motif (CBM) to the two symmetrically related CBM-binding pockets of β_2. Tethered to the clamp, Pol III is able to synthesize DNA at high speed ($\sim$1000 Nt/s) and with much higher processivity (>150 kb) [1,15].

Bacterial replisomes are highly flexible and mobile machines, their dynamics being mediated and controlled by a network of protein–protein interactions of different strengths. Many of the replication proteins are either conformationally flexible or contain flexible or unstructured regions, making structural studies by X-ray crystallography or NMR difficult. However, through decades of efforts, structures of all *E. coli* replication proteins or their