8-2020

Analysis Of The Role Of Dnaa Recognition Site Position In Directing Assembly Of The Pre-Replicative Complex In Escherichia Coli

Christopher Aaron Czerwonka

ANALYSIS OF THE ROLE OF DNAA RECOGNITION SITE POSITION IN DIRECTING ASSEMBLY OF THE PRE-REPLICATIVE COMPLEX IN *ESCHERICHIA COLI*

by

CHRISTOPHER AARON CZERWONKA

B.S., Walla Walla University

A dissertation submitted to the Department of Ocean Engineering and Marine Sciences of Florida Institute of Technology in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
In
BIOLOGICAL SCIENCES

Melbourne, FL
August 2020

ANALYSIS OF THE ROLE OF DNAA RECOGNITION SITE POSITION IN DIRECTING ASSEMBLY OF THE PRE-REPLICATIVE COMPLEX IN *ESCHERICHIA COLI*

A DISSERTATION

By

CHRISTOPHER AARON CZERWONKA

Approved as to style and content by:

_____

Alan C. Leonard, Ph.D., Chairperson
Professor
Department of Biomedical and Chemical Engineering and Sciences

_____

Julia E. Grimwade, Ph.D., Member
Professor
Department of Biomedical and Chemical Engineering and Sciences

_____

David J. Carroll, Ph.D., Member
Associate Professor
Department of Biomedical and Chemical Engineering and Sciences

_____

Charles D. Polson, Ph.D., Member
Associate Professor
Department of Biomedical and Chemical Engineering and Sciences

_____

Debasis Mitra, Ph.D., Member
Professor
Department of Computer Science

_____

Richard B. Aronson, Ph.D.,
Professor and Head
Department of Ocean Engineering and Marine Sciences

ABSTRACT

ANALYSIS OF THE ROLE OF DNAA RECOGNITION SITE POSITION IN DIRECTING ASSEMBLY OF THE PRE-REPLICATIVE COMPLEX IN *ESCHERICHIA COLI*

by Christopher Aaron Czerwonka, B.S., Walla Walla University

Chairperson of Advisory Committee: Alan C. Leonard, Ph.D.

Prior to initiating new rounds of DNA replication, all cells assemble pre-replicative complexes (pre-RC) consisting of initiator proteins and regions of DNA termed origins of replication. In the bacterial model system, *E. coli*, multiple copies of the initiator protein, DnaA, assemble into a complex with the unique chromosomal replication origin, *oriC*, to produce a pre-RC that unwinds the DNA helix, preparing the origin for the new replication forks. Nine base pair (bp) low affinity DnaA recognition sites are distinctly arrayed in each half of *oriC* and both arrays are bounded by high affinity recognition sites. Although it is known that DnaA does not interact at low affinity site arrays without assistance from DnaA occupying a proximally positioned high affinity site, it remains unclear why this particular arrangement of high and low affinity sites exists. Furthermore, despite the fact that DnaA is a highly conserved protein among all bacteria, the numbers

and arrangements of high and low affinity sites in *oriC* varies considerably among different bacterial types. In order to understand the reason for this diversity, more information about *E. coli oriC* geography and its relationship to pre-RC assembly and origin function is required.

In the studies reported here, two specific questions about *E. coli oriC* geography were addressed: 1) are all three high affinity sites required for *oriC* function, and 2) is the nucleotide spacing between recognition sites an inflexible feature. To answer these questions, mutagenized versions of *E. coli oriC* on chromosomes were evaluated for function *in vivo*, and DNA fragments containing different spacing and arrangements of high and low affinity recognition sites were examined by electrophoretic mobility shift assays (EMSA), high resolution DNA footprints, and DNA unwinding assays.

When inserted into *E. coli* as replacements for the wild-type chromosomal origin, mutant versions of *oriC* lacking any two high affinity recognition sites (R1, R2, or R4) were inactive, but origins carrying only one inactivated high affinity DnaA recognition site retained function no matter which site was altered. This finding contradicts previous reports that the R1 site is essential for *oriC* function, although loss of either R1 or R4 binding caused reduced origin activity and loss of initiation synchrony. Although it appears that the requirement for high affinity DnaA binding is not stringent for origin activity, a novel phenotype for high affinity site-deficient mutants was identified. Supercoiled *oriC* DNA is not

normally unwound at low levels of DnaA, but spontaneous DNA unwinding of every high affinity binding mutant was detected under these conditions. A model is proposed whereby *oriC* DNA is normally constrained by DnaA during the cell cycle to prevent spontaneous initiations and constraint requires the cooperative interactions among the distantly spaced DnaA bound to high affinity sites as an origin recognition complex. These interactions, presumably through the proper positioning of domain I-domain I contacts, would generate multiple DNA loops, analogous to the wrapping of DNA around nucleosomes in eukaryotes. Further contraction and stabilization of these loops by additional DnaA occupying arrayed low affinity sites would then produce the torsional stress required to separate DNA strands within the A-T rich DNA Unwinding Element (DUE) at the left side of *oriC*.

Electrophoretic mobility shift assays (EMSA) with DNA fragments containing differently spaced DnaA recognition sites revealed that very small changes in base pair spacing between low-low or high-low affinity recognition sites resulted in dramatic changes in cooperative DnaA binding required for site occupation. Optimal binding at any two sites was detected at 2 bp spacing, but was diminished as single bases were added. Surprisingly, reducing the 2 bp spacing also diminished DnaA binding suggesting that some structural interference exists between adjacent DnaA molecules when they are brought too close together.

These observations are consistent with specific positioning of DnaA recognition sites in E. coli *oriC* for correct origin function and even slight alterations in these positions have an impact. While it remains difficult to predict where DnaA recognition sites will be placed in the replication origin of any particular bacteria type, the knowledge gained from the studies presented here establishes a rule set that include the requirement for the arrangement of high affinity DnaA recognition sites to, at an early stage in the cell cycle, topologically constrain the *oriC* in an origin recognition complex and the requirement for additional recognition sites that convert the constrained complex into one that imparts sufficient torsional stress to unwind *oriC*.

DEDICATION

This dissertation is dedicated to my wife, Kelly Czerwonka. Without her love, support, and endless encouragement this would not have been possible.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Alan Leonard for all his support and mentoring. His words of advice pushed me to grow, not only as a scientist, but as a person and for that I am eternally grateful. A big thank you to Dr. Julia Grimwade for the many thoughtful discussions and guidance in designing experiments. Thank you to my committee members Dr. Charles Polson, Dr. David Carroll and Dr. Debasis Mitra for their advice and suggestion throughout the process.

A thank you to all the graduate and under-graduate students in the Leonard/Grimwade labs for fostering an environment of learning and comradery. A special thank you to Tania Rozgaja, Mansi Vora, and Gulpreet Kaur for all their help, support and friendship.

I would like to thank my parents and grandparents for their love and support throughout my research, even when they had no idea what I was talking about.

I would like to express my gratitude to the National Institutes of Health for funding this research. A thank you to the Florida Institute of Technology Biology Department teaching assistantship program for the opportunity to instruct lab courses.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

OVERVIEW

Chromosome replication must be precisely timed within every cell cycle to ensure complete duplication and equal distribution of genetic information before cell division. To accomplish this task, all cells focus regulatory mechanisms on the step-wise assembly of DNA-protein complexes that trigger new rounds of DNA replication, and the stages of assembly are well conserved in all cell types (Duderstadt and Berger 2008, Masai et al. 2010, Diffley 2010, Kawakami and Katayama 2010, Katayama et al. 2010). First, initiator proteins are assembled into complexes that recognize regions of DNA termed origins of replication (Zakrzewska-Czerwinskaet et al. 2007, Mott and Berger 2007, Katayama et al. 2010, Diffley 2011, Leonard and Grimwade 2011, Parker et al. 2017). These Origin Recognition Complexes (ORC) persist throughout the cell cycle and form a stable foundation for the recruitment of additional initiator proteins (Samitt et al. 1989, Nievera et al. 2006, Miller et al. 2009, Kaur et al. 2014, Leonard and Grimwade 2015). The recruited proteins assemble into a pre-Replication Complex (pre-RC) with the ability to load DNA helicase onto single-stranded DNA as the first step in preparing the DNA for new replication forks. Following DNA helicase loading, and the onset of new DNA synthesis, spurious reassembly of pre-RCs is blocked to prohibit unscheduled re-initiations and ORCs are reset for the next cell cycle. In bacteria, most of our understanding of these events and their associated

regulatory mechanisms come from studies in *Escherichia coli* (Leonard and Grimwade 2011, 2015), whose circular chromosome replicates bi-directionally from a unique origin of replication (*oriC*) with a defined nucleotide sequence. Every new round of replication initiation requires the accumulation of about 20 molecules of initiator protein, DnaA, whose activity is regulated by bound ATP (see below and McGarry et al. 2004, Kawakami et al. 2005, Grimwade et al. 2007, Saxena et al. 2011, 2013a&b, Rozgaja et al. 2011, Kaur et al. 2014). DnaA-ATP molecules interact specifically with double-stranded *oriC* DNA at both high and low affinity 9 base pair (bp) recognition sites, and at exactly the correct time, complete assembly of a pre-RC that is able to unwind *oriC* DNA and assist with DNA helicase loading onto the single strands (Sutton et al. 1998, Abe et al. 2007, Mott et al. 2008, Ozaki and Katayama 2012, Ozaki et al. 2012a). Although the positioning of each recognition site within *oriC* is believed to play a role in ordered pre-RC assembly, it remains unclear why a particular arrangement is used and how this arrangement leads to a pre-RC capable of mechanically unwinding *oriC* DNA.

The goal of this study was to better understand why DnaA recognition sites are located at specific *E. coli oriC* positions and learn how these arrangements produce the properly ordered assembly of a functional pre-RC. It is important to note that the arrangement of DnaA recognition sites is highly diverse among the *oriC*s of different bacterial types, but DnaA is a highly conserved protein among all eubacteria (Løbner-Olesen et al. 1989, Mott and Berger 2007, Leonard and Méchali 2013, Wolanski et al. 2015). Thus, the information obtained in this study will be

useful for comparative analysis of pre-RC assembly, particularly among bacteria of medical interest, and may reveal interesting new targets for the rational design of antimicrobials.

OVERVIEW OF THE BACTERIAL INITIATOR PROTEIN DNAA

The most widely accepted models for a regulatory mechanism that couples the initiation step to cellular growth rate (Løbner-Olesen et al. 1989) are based on the accumulation of DnaA, a member of the AAA^+ (ATPases Associated with various cellular Activities) superfamily of proteins whose activity is regulated by the hydrolysis of ATP (Sekimizu et al. 1987, Messer et al. 1999, Erzberger et al. 2002, Snider et al. 2008). DnaA has a high affinity for ATP (K_D= 30nM) and ADP (K_D =100nM), but only DnaA-ATP is active at the replication initiation step (Fuller and Kornberg 1983, Sekimizu et al. 1987). Overexpression of DnaA causes *E. coli* to initiate chromosome replication early in the cell cycle and at a smaller cell mass, while under-expression leads to late initiation at a larger cell mass (Atlung et al. 1987, Pierucci et al. 1989, Løbner-Olesen et al. 1989). Based on these results, it appears that a threshold amount of DnaA-ATP is required during the bacterial cell cycle to trigger chromosome replication.

The *E. coli* DnaA protein (52 kDa) is divided into four structural domains (Figure 1A). Domain I (N-terminus amino acids 1-79) is required for DnaA-DnaA oligomerization (Weigel et al. 1999, Abe et al. 2007) and association with other replication proteins like DnaB (replicative helicase), DnaC (helicase loader), and

DiaA (DnaA-initiator-associating protein) which stimulates assembly of the pre-RC. Although the details are not yet understood, DiaA appears to regulate initiation by a novel mechanism, in which DiaA tetramers likely bind to multiple DnaA molecules and stimulate the assembly of specific ATP–DnaA–oriC complexes. (Sutton et al. 1998, Seitz et al. 2000, Ishida et al. 2004, Su'etsugu et al. 2005, Keyamura et al. 2007). Substituting the highly conserved tryptophan at position six with an alanine (W6A) inactivated the oligomerization activity and prevented interaction with the DnaC-DnaB replicative complex (Felczka et al. 2005, Zawilak-Pawlik et al. 2017).

The role of Domain II (amino acids 80-135) is not fully known but is generally believed to be a flexible linker separating Domains I and III. Domain I oligomeric interactions between adjacent DnaA molecules in *oriC* are required during pre-RC assembly (Weigel et al. 1999, Simmons et al. 2003, Kawakami et al. 2005, Miller et al. 2009, Rozgaja et al. 2011, Zawilak-Pawlik et al. 2017) and it was suggested that the length of the flexible linker may determine the distance over which these interactions can take place (Zawilak-Pawlik et al. 2005, 2017). Domain II is the least conserved domain, varying significantly in amino acid sequence and length among bacterial types (Messer et al. 1999). Some mutations that shorten the length of *E. coli* Domain II are viable but perturb initiation timing of fast-growing cells (Nozaki and Ogawa 2008). All sections of the domain can be deleted, but not at the same time. A minimum 21-27 amino acids separating Domains I and III are required for replication function (Nozaki and Ogawa 2008).

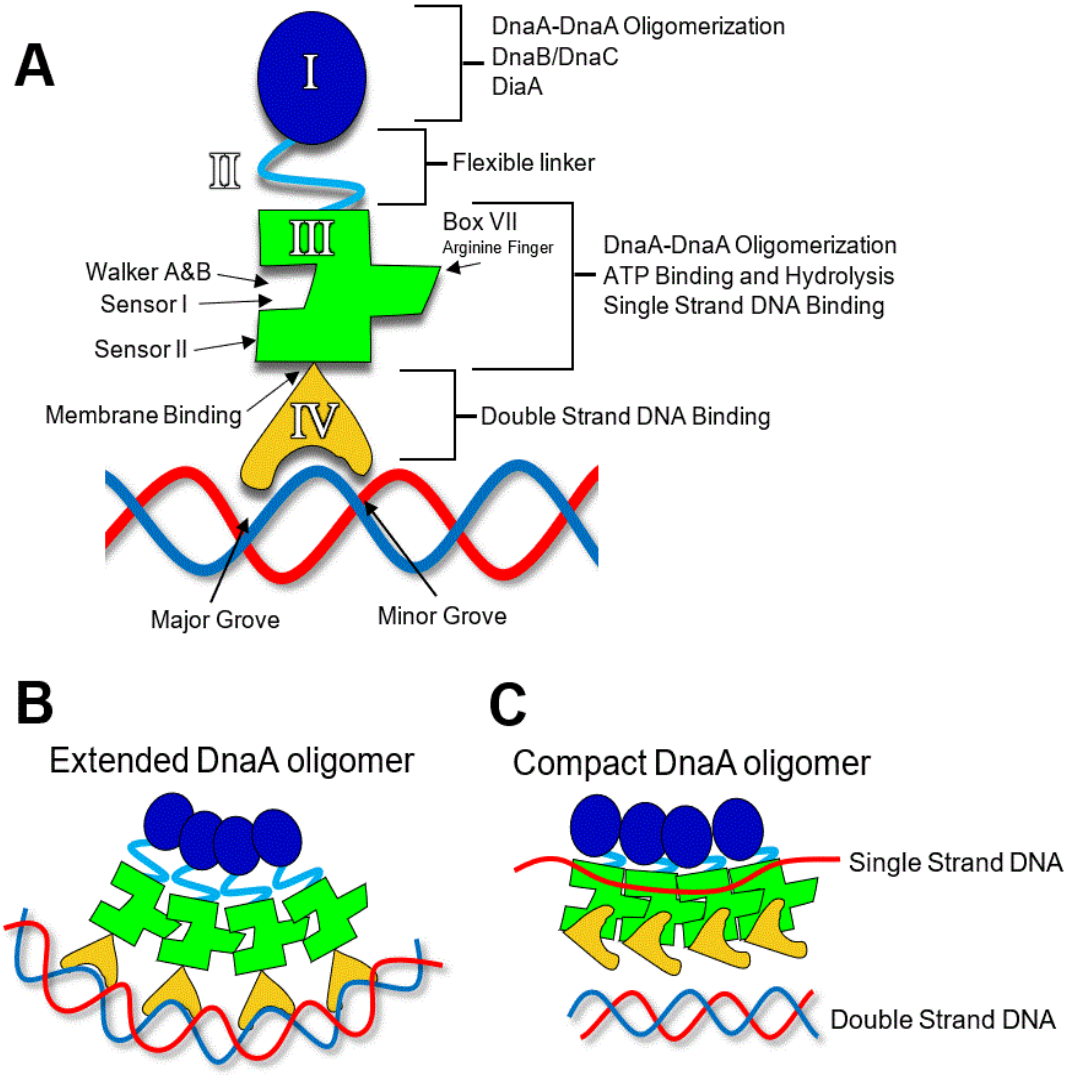

Figure 1. A) Structural domains of DnaA. Domain I (dark blue), site of self-oligomerization and interacts with other replication proteins like the DnaC/DnaB complex and DiaA. Domain II (light blue), a flexible linker whose length varies among bacterial types. Domain III (green), the AAA+ domain with Walker A&B and Sensor I&II motifs for ATP binding and hydrolysis, and a Box VII motif essential for self-oligomerization. Domain III also has hydrophobic and basic motifs for binding single stranded DNA. Domain IV (yellow), the double strand DNA binding domain contains a Helix-Turn-Helix motif that interacts with 9 bp recognition sites. B) Illustration of DnaA head-to-tail oligomerization along double strand DNA. C) Illustration of a compact single stranded binding DnaA oligomer with Domain IV tuned up to interact with Domain III. (Illustrations are not to scale)

Domain III (amino acids 136-347) is the site of ATP binding and hydrolysis, as well as DnaA oligomerization. The structural motifs of the domain are arranged in such a way to form a fold in which Adenine nucleotide binds. The binding fold is made up of Walker A and B motifs, which interact directly with ATP/ADP, along with a Sensor I motif that enhances the affinity for ATP/ADP binding (Neuwald et al., 1999, Erzberger et al. 2002, Nishida et al. 2002, Iyer et al. 2004, Kawakami et al. 2006, Ozaki et al. 2012b). A Sensor II motif, near the C-terminus of the domain, monitors the phosphorylation state of the bound nucleotide and supports hydrolysis (Nishida et al. 2002). Opposite the nucleotide binding fold is a Box VII structure with a protruding Arginine finger (Erzberger et al. 2002, Felczak and Kaguni 2004). Occupation of some low affinity recognition sites in *oriC* is restricted to DnaA-ATP (Margulies and Kaguni 1996, Speck et al. 1999, McGarry et al. 2004, Rozgaja et al. 2011). While it has not been firmly established that Domain III-dependent oligomerization is required for binding to double-stranded DNA, a head-to-tail oligomer model exists with the Box VII Arginine finger of one DnaA inserted into the nucleotide binding fold of the adjacent DnaA (Figure 1B). Domain III also contains hydrophobic and basic motifs that allow interaction with single-stranded DNA (Ozaki et al. 2008, Duderstadt et al. 2011). Binding to single-stranded DNA does not appear to be as nucleotide sequence specific as is the case for double-stranded interactions (see below, and Richardson et al. 2016), and the structure of the single strand binding oligomer is more compact with Domain IV turned up and inward to interact with Domain III

(Duderstadt et al. 2010) (Figure 1, compare B and C). At the junction between Domains III and IV lie amino acids that interact with the cell membrane in a process that stimulates the exchange of ADP for ATP (see below) (Sekimizu and Kornberg 1988, Garner and Crooke 1996).

Domain IV (amino acids 348-467) is the double strand DNA binding domain. A Helix-Turn-Helix (HTH) motif within the domain interacts with specific nucleotides in both the major and minor groves of double helix DNA (Fujikawa et al. 2003). DnaA recognition sites consist of nine consecutive nucleotides with the consensus sequence of 5'-TTA/TTNCACA-3'. Sites with consensus or near consensus sequences usually have high affinity for DnaA, while sites that deviate by 2 or more base pairs typically have lower affinities (see below). Since the recognition sequence is not palindromic, DnaA interactions with its contacts are directional. For this reason, the orientation of *oriC* recognition sites (shown by arrows in Figure 3) determines the direction of DnaA oligomerization along *oriC* (see below). High affinity DnaA binding causes a bend in the DNA helix of approximately 40 degrees (Roth and Messer 1995, Schaper and Messer 1995, Sutton and Kaguni 1997, Fujikawa et al. 2003) which may cause torsional stress on supercoiled DNA.

REGULATING DNAA-ATP LEVELS IN *E. COLI* CELLS

The overall amount of DnaA does not fluctuate much during *E. coli*'s cell cycle, instead it is the ratio of DnaA-ATP/DnaA-ADP that changes (Sekimizu et al.

1988a, Kurokawa et al. 1999). At the time of initiation, 80% of freely available DnaA is in the active ATP-bound form. These levels drop sharply after initiation to roughly 30% of the available total (Kurokawa et al. 1999) due to the conversion of DnaA-ATP to DnaA-ADP by a post-initiation activity associated with the replication forks. Simultaneously, any freely available DnaA-ATP is blocked from interacting with *oriC* during the post-initiation period. Also, autoregulation of the *dnaA* gene (Braun et al. 1985, Atlung et al. 1985) prevents transcription. Synthesizes of new DnaA is required during each cell cycle to trigger new rounds of chromosome replication (McMacken et al. 1987). Newly formed DnaA rapidly becomes DnaA-ATP due to high levels of ATP in bacterial cells. These mechanisms are described below and diagramed in Figure 2.

INACTIVATION OF DNAA-ATP BY RIDA

Inactivation of DnaA by the replication fork is regulated by the protein Hda (Homologous to DnaA) which stimulates the intrinsic ATPase activity of DnaA in a process termed Regulatory Inactivation of DnaA (RIDA) (Katayama et al. 1998, Kato and Katayama 2001). Hda, a AAA+ protein, is active only when bound to ADP and associated with the β-clamp subunit of DNA polymerase III holoenzyme (Su'etsugu et al. 2004, Kim et al. 2017). As the replication fork moves along the DNA encountering bound DnaA, the arginine finger of Hda complexes with Domain III of this DnaA to stimulate hydrolysis of ATP to ADP (Katayama et al. 2010, Keyamura and Katayama 2011, Skarstad and Katayama 2013).

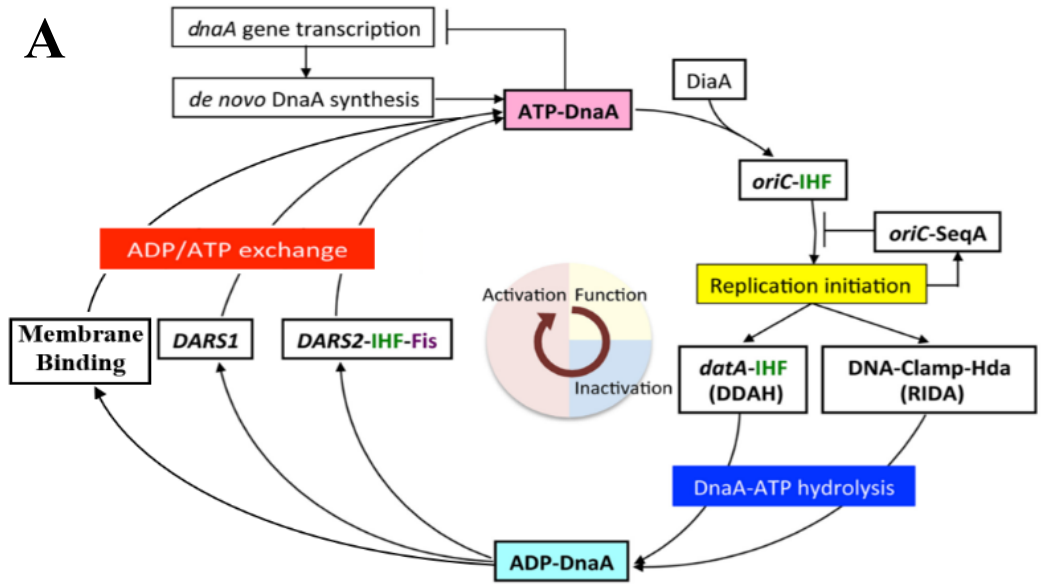

Figure 2. A) The DnaA-ATP/ADP cycle. Accumulation of active DnaA-ATP at *oriC*, aided by DiaA and IHF, leads to initiation of replication (light Yellow "Function" period). Post initiation, rebinding of DnaA to *oriC* is blocked by SeqA and DnaA is inactivated by hydrolysis of ATP to ADP through the Regulatory Inactivation of DnaA (RIDA) and *datA* dependent DnaA-ATP Hydrolysis (DDAH) pathways (light Blue "Inactivation" period). DnaA-ADP is reactivated through association with acidic phospholipids in the cell membrane and binding to DARS1&2 chromosomal loci which stimulate the exchange of ADP for ATP (light Red "Activation" period). Regenerated DnaA-ATP, in conjunction with newly synthesized DnaA, bind *oriC* to initiate the next round of replication. (diagram adapted from Katayama et al. 2017). B) Relative positions of key DnaA regulatory DNA regions on the *E. coli* chromosome.

SEQUESTRATION OF *ORIC*

The *E. coli* origin contains eleven GATC sequences which are the target of DNA adenine methyltransferase (DAM methylase) (Waldminghaus and Skarstad 2009). Most sites are in the DUE and left half of the origin. DAM methylase methylates the adenine nucleotides on both strands of the DNA (fully methylated), however, following a round of semiconservative replication, the DNA is hemimethylated. Hemimethylated GATCs are the target of the protein SeqA, whose binding sequesters *oriC*, blocking DnaA interactions (Campbell and Kleckner 1990, Torheim and Skarstad 1999, Kang et al. 2003, Nievera et al. 2006). This sequestration period persists roughly one third of the cell cycle before DAM methylase converts GATCs back to the fully methylated configuration (Lu et al. 1994). There are no GATC sequences within or nearby the high affinity DnaA recognition sites (R1, R2, and R4). These sites are not blocked during sequestration and newly replicated *oriC*s are rapidly reset to an early stage DnaA-*oriC* complex (Nievera et al. 2006 and see below). Post initiation transcription of *dnaA* is also blocked by binding of SeqA to GATC sites within the gene promoter region (Campbell and Kleckner 1990, Theisen et al. 1993), thereby turning off new DnaA synthesis for a significant portion of the cell cycle.

TITRATION AND INACTIVATION OF DNAA

About 300 high affinity DnaA recognition sites are dispersed throughout the *E. coli* genome (Roth and Messer 1998). As these sites are duplicated they become

potential binding sites for DnaA-ATP, thereby reducing the number of molecules available to interact with *oriC*. One unique location on the chromosome, termed *datA*, has a high capacity for DnaA binding, and may tie up 50-100 molecules of initiator protein (Kitagawa et al. 1996 and 1998) during each cell cycle. *DatA* is duplicated approximately 9 minutes after initiation (during sequestration) and could ensure that any stray unbound DnaA-ATP does not interact with *oriC* during this time period (Ogawa et al. 2002) (Figure 2B). In addition to titrating DnaA away from *oriC*, evidence suggest *datA* stimulates hydrolysis of DnaA-ATP to the inactive ADP form. *DatA* contains clusters of DnaA binding sites along with a recognition site for the DNA bending protein, IHF (Nozaki et al. 2009). Binding of IHF to negatively supercoiled *datA* forms a looped structure that promotes binding and oligomerization of DnaA. The *datA*-IHF-DnaA complex is sufficient to stimulate the intrinsic ATPase activity of DnaA (*datA*-dependent DnaA-ATP hydrolysis, DDAH) even in the absence of RIDA (Kasho and Katyama 2013, Kasho et al. 2017). Cells lacking either *datA* or IHF have increased levels of DnaA-ATP and over-initiate replication (Kitagawa et al. 1996, Nozaki et al. 2009, Kasho and Katayama, 2013).

REGENERATION OF DNAA-ATP

Under conditions of rapid bacterial growth, it appears that sufficient newly synthesized DnaA-ATP is not available and must be supplemented by regeneration from the ADP-bound form (Kurokawa et al. 1999). Two different regulatory

mechanisms are reported to be responsible for reactivation (Skarstad and Katayama 2013), one associated with DNA and the other associated with the cell membrane (Figure 2). Two chromosomal DNA regions, termed DnaA Reactivation Sequences (DARS1 and DARS2), contain clusters of high affinity DnaA recognition sites that promote the exchange of ADP for ATP. Both DARS sites contain a core of three back-to-back DnaA high affinity sites with two sites in the same orientation and one in the opposite orientation (Fujimitsu et al. 2009). The positioning of these recognition sites is vital for activity (Kasho et al. 2014). Although the mechanism remains unclear, interactions between DnaA molecules stimulates the release of ADP leaving an open nucleotide fold into which ATP quickly binds. In addition to the core high affinity sites, DARS2 has additional DnaA recognition sites along with binding sites for Fis and IHF proteins. The activity of DARS2, but not DARS1, is strongly enhanced by Fis and IHF binding (Fujimitsu et al. 2009, Kasho et al. 2014). These two loci are opposite each other on the chromosome, with DARS2 closer to *oriC* and DARS1 closer to the replication terminus, *terC* (Figure 2B). It remains unclear whether the use of these sites is restricted to particular times during the cell cycle, but loss of either locus decreases the cellular level of available DnaA-ATP (Fujimitsu and Katayama 2004, Fujimitsu et al. 2009, Kasho et al. 2014).

DnaA-ADP can also be regenerated through interactions with acidic phospholipids in the cellular membrane (Boeneman and Crooke 2005, Sekimizu and Kornberg 1988, Yung and Kornberg 1988, Crooke et al. 1992, Aranovich et al. 2006, Fingland et al. 2012, Saxena et al. 2013a). The mechanistic basis for this exchange has yet to be determined, but accumulation of DnaA to a threshold density at the membrane and oligomerization through Domain I-I interactions appear to be key factors in the ADP-ATP exchange (Heacock and Dowhan 1987, Aranovich et al. 2006, 2015).

FEATURES OF THE *E. COLI* ORIGIN OF REPLICATION

The *E. coli* origin of replication (*oriC*) consists of 245 base pairs located between nucleotides 3,923,767 to 3,924,012 on the K-12 genome map (Blattner et al. 1997). The origin is divided into two functional regions, an AT-rich DNA Unwinding Element (DUE) and an adjacent control region containing a series of high and low affinity DnaA binding sites (Figure 3). The DUE is made up of three 13 bp motifs (5' - GATCTNTTN$_{A/TA/TA/T}$G - 3′) and DnaA-dependent DNA strand separation is localized between the middle and right repeats (Bramhill and Kornberg 1988, Kowalski and Eddy 1989).

Adjacent to the DUE are eleven binding sites for the initiator protein DnaA. Three sites, with consensus or near consensus sequence, R1, R2 and R4, have high affinity for DnaA (K_D 4-20nM) (Figure 3) (Fuller et al. 1984, Matsui et al. 1985, Schaper and Messer 1995). The remaining DnaA sites; R5M, τ2, I1, I2, C1, I3, C2,

and C3 deviate from consensus by two or more nucleotides and have low affinity for DnaA (K_D >200nM) (McGarry et al. 2004, Kawakami et al. 2005, Rozgaja et al. 2011, Shimizu et al. 2016).

The three high affinity sites, R1, R2, and R4, are evenly spaced within the origin forming two gap regions in which low affinity sites are arranged in clusters with each site separated from their neighbors by precisely two base pairs. Within each clustered array the sites have the same binding orientation and the two arrays are opposite each other (Figure 3). The three high affinity sites, along with low affinity sites R5M and C1, bind DnaA-ATP and DnaA-ADP equivalently while all other low affinity sites show a four-fold preference for DnaA-ATP over DnaA-ADP (McGarry et al. 2004, Ozaki et al. 2008, Rozgaja et al. 2011). DnaA binds high affinity sites independently while binding to low affinity sites requires assistance in the form of cooperative interaction from DnaA bound to nearby high affinity site (Miller et al. 2009, Rozgaja et al. 2011). The distance between the low affinity arrays and proximal high affinity sites varies from as few as three base pairs (R4 to C1), up to 46 base pairs (R1 to R5M).

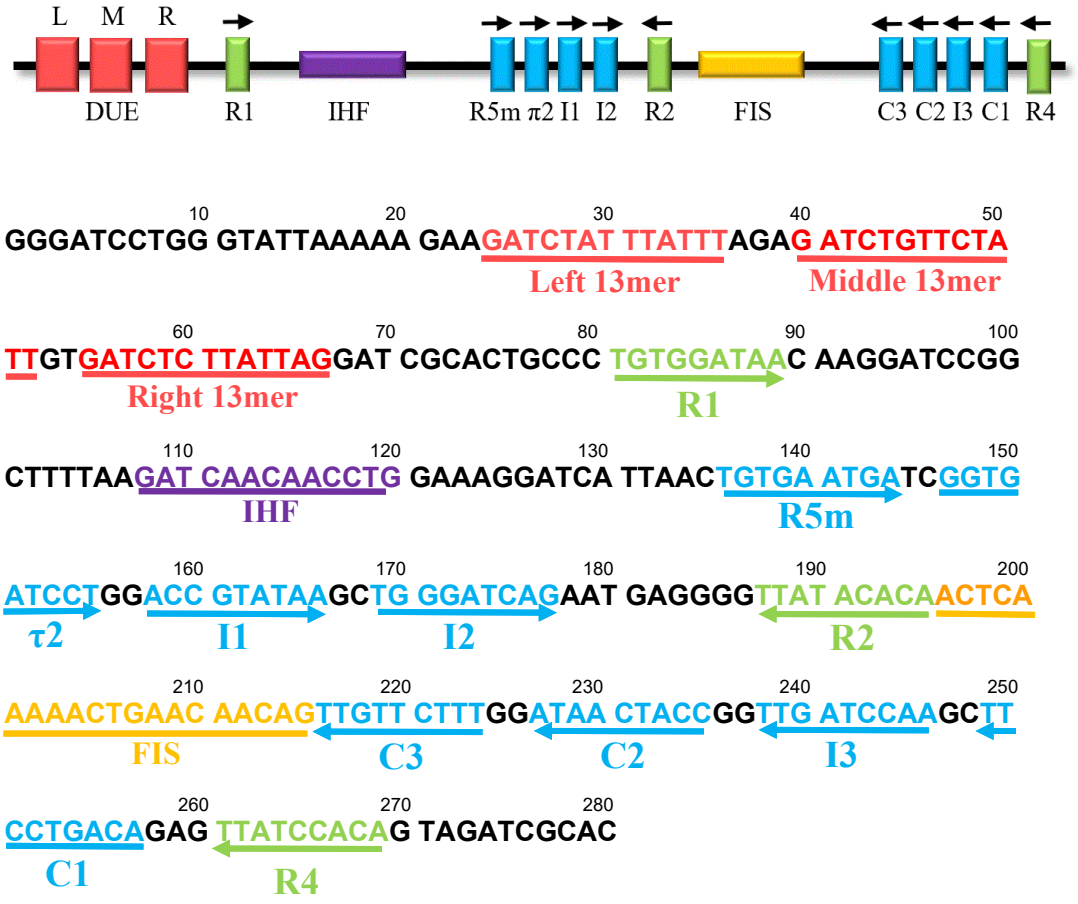

Figure 3. The *E. coli* origin of replication (*oriC*). High affinity DnaA binding sites are indicated in green, low affinity DnaA binding sites in blue. The Fis binding site is denoted in yellow and IHF in purple. The three AT-rich repeats that make up the DNA Unwinding Element (DUE) are shown in red. Arrows at each DnaA recognition site indicate relative binding orientation.

In addition to the DnaA binding sites in *oriC*, there are recognition sites for two accessory proteins, Fis (Factor for Inversion Stimulation) and IHF (Integration Host Factor) (Figure 3), both of which are able to bend *oriC* DNA and contribute to the regulation of DNA repair, recombination, replication, and gene expression. Fis recognizes the consensus sequence 5'-GNTCAAATTTTGANC-3' and binds to the right half of *oriC* inducing a bend of 40 to 90 degrees (Thompson and Landy 1988, Finkel and Johnson 1992, Pan et al. 1994). IHF binds the consensus sequence 5'-YAANNNNTTGATW-3' in the left half of *oriC* inducing a sharp bend of nearly 180 degrees (Gardner and Nash 1986, Leong et al. 1985, Gamas et al. 1987, Rice et al. 1996). Fis, but not IHF, is growth rate-regulated and associates with *oriC* only during rapid growth (Ali Azam et. al. 1999, Dillon and and Dorman 2010). The roles of Fis and IHF during pre-RC assembly are described below.

STAGED ASSEMBLY OF THE PRE-REPLICATIVE COMPLEX

Studies of *oriC*-DnaA interactions reveal that pre-RC assembly proceeds through at least two stages, each of which generates a distinctive subassembly (Kaguni and Kornberg 1984, Sekimizu et al. 1988b, Margulies and Kaguni 1996, McGarry et al. 2004, Rozgaja et al. 2011, Ozaki and Katayama 2012, Ozaki et al. 2012a). The three high affinity sites, R1, R2, and R4 are occupied throughout the cell cycle while the low affinity sites fill only briefly at the time of initiation (Cassler et al. 1995, Grimwade et al. 2000, Nievera et al. 2006, Miller et al. 2009, Rozgaja et al. 2011). The persistent DnaA bound to the high affinity sites is a

temporal and functional analog to the Origin Recognition Complex (ORC) described in yeast (Stillman 2005) and has been termed the bacterial ORC (Nievera et al. 2006). Like the yeast ORC, the bacterial ORC recruits additional proteins to the origin and serves as the framework on which the pre-Replicative Complex (pre-RC), is built. Electron microscopy studies have shown supercoiled *oriC* template reproducibly forms specific structural conformations in the presence of DnaA with up to 200 base pairs of DNA wrapped around the assembled complex (Crooke et al. 1993, Funnell et al. 1987).

Based on hi-resolution DNA footprints, the order of DnaA binding to *oriC* has been elucidated (Figure 4). DnaA bound to the high affinity site R4 nucleates loading of the proximal low affinity site C1, after which, additional DnaA is recruited to sequentially fill the right-side array (sites I3, C2, and C3). If the cell is growing rapidly, the accessory protein Fis will occupy its recognition sites between R2 and C3 as a component of the ORC (Gille et al. 1991). The presence of Fis blocks the progress of pre-RC assembly due to its negative effect on IHF binding in the left half of *oriC*, probably by producing an incompatible bend in the DNA (Ryan et al. 2004). Fis is ultimately displaced from *oriC* by the accumulation of DnaA in the right half of *oriC* (Ryan et al. 2004). The mechanism for this has yet to be determined. IHF then binds *oriC* producing a sharp bend in the DNA promoting nucleating of DnaA binding to the low affinity sites in the left half of *oriC* (Cassler et al. 1995, Ryan et al. 2004). Extension of the DnaA oligomers along the left side of *oriC* drives localized strand separation in the DUE (Baker and

Bell 1998, Speck and Messer 2001). The completed pre-RC then assists loading of DNA helicase (DnaB) (Sutton et al. 1998, Abe et al. 2007, Mott et al. 2008, Ozaki and Katayama 2012, Ozaki et al. 2012a).

During rapid growth, *E. coli* initiates new rounds of replication before synthesis of previously initiated rounds have completed, resulting in copies of chromosomal *oriC* that exceed two (Cooper and Helmstetter, 1968, Helmstetter, C. E. & Cooper 1986, Helmstetter 1996, Michelsen et al. 2003). It is worth noting that mutant strains of *E. coli* lacking either Fis or IHF are not able to initiate multiple copies of chromosome *oriC* synchronously. When cells undergo asynchronous initiation, the daughters produced may contain the wrong number of chromosomes and undergo replicative stress (Skarstad et al.1986, Løbner-Olesen et al. 1994). Use of the Fis/IHF switch mechanism during pre-RC assembly must narrow the time window for all initiation steps during the cell cycle, and the amount of DnaA required to displace Fis would be expected to exceed the amount required to complete the remaining stages of pre-RC assembly to ensure synchronous initiations. The IHF-dependent DNA bend that produces the cross-strand interaction of DnaA molecules at R1 and R5M would also be expected to impart an instability into pre-RC assembly that might be advantageous for the rapid resetting of the ORC following each round of chromosome replication.

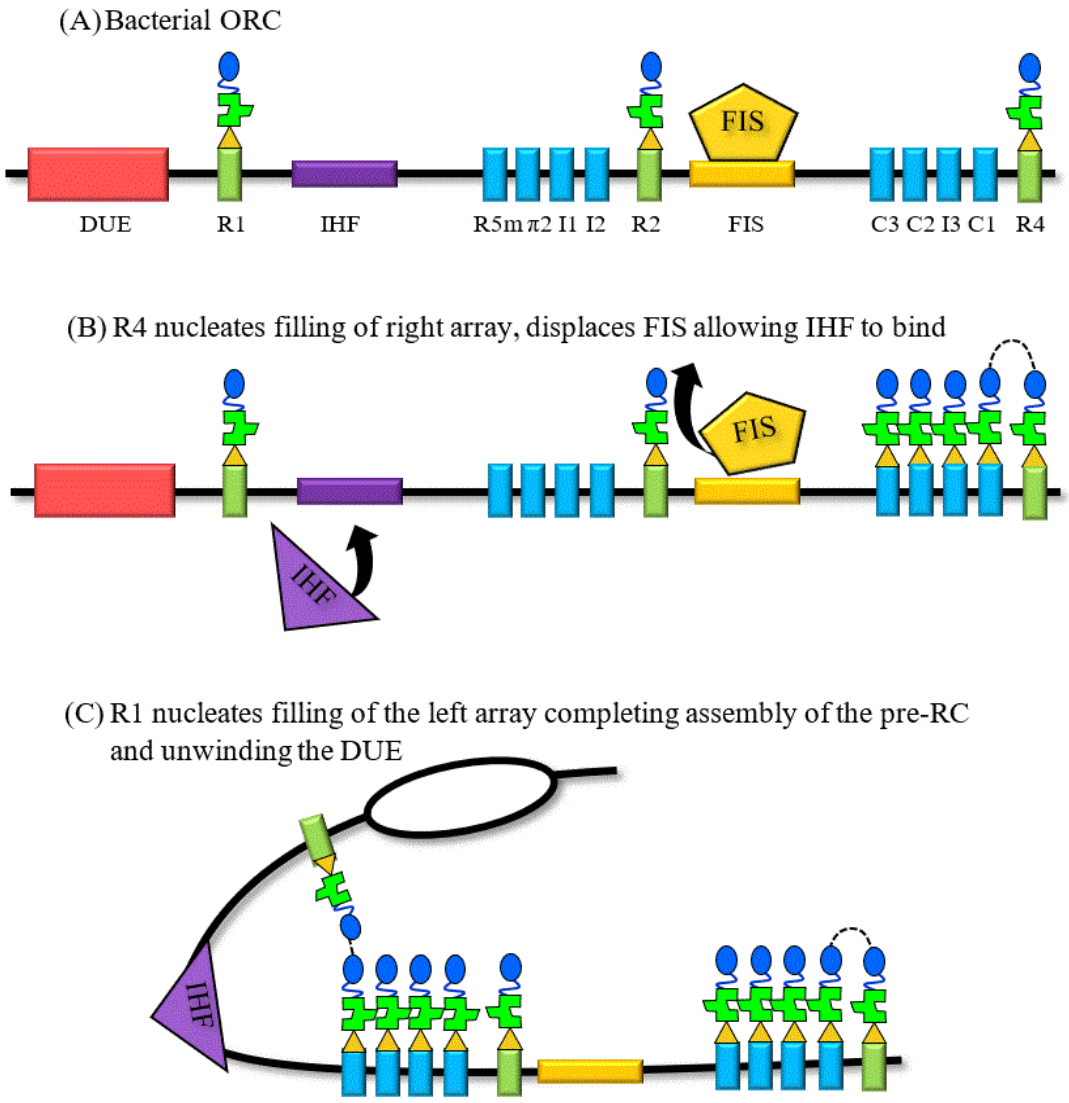

Figure 4. Stages of *E. coli* pre-RC assembly. A) The bacterial ORC consists of DnaA bound to the three high affinity sites and, when growing rapidly, Fis bound to the right half of *oriC*. B) DnaA occupying R4 nucleates elongation of a DnaA oligomer along the right array. This oligomer displaces Fis allowing IHF to bind the left half of *oriC*. C) IHF bends the origin moving DnaA bound at R1 closer to the left array nucleating elongation of the second DnaA oligomer. Filling of all low affinity DnaA recognition sites in *oriC* drives localized unwinding in the DUE. High affinity binding sites are denoted in green, low affinity sites in blue. The Fis binding site in yellow and IHF in purple. The closed DUE is denoted in red. Illustration is not to scale.

DIVERSITY AMONG BACTERIAL *ORIC*S AND THE ENIGMA OF DNAA RECOGNITION SITE ARRANGEMENTS

The ordered assembly of the *E. coli* pre-RC clearly utilizes a specific arrangement of high and low affinity DnaA recognition sites, placed in specific positions within *oriC*, and it might be expected that this arrangement would be commonly found among all bacteria, especially due to the highly conserved DnaA initiator. However, this *oriC* geography is found only in *E. coli* and its closest relatives (Shaheen et al. 2009). It is now possible to examine and compare the *oriC* nucleotide sequence of over a 1000 bacterial types (D*ORIC* database Gao et al. 2013), and although clusters of DnaA binding sites are found in all eubacterial replication origins, the number and arrangement of sites vary dramatically across species (Marczynski and Shapiro 1992, Zawilak-pawlik et al. 2005, Gao and Zhang 2007, Leonard and Mechali, 2013). Close relatives of *E. coli* have consensus DnaA boxes at the boundaries of the replication origin with one or more central high affinity sites separated by gap regions (Figure 5A). These gap regions vary in length and sequence from a few dozen to hundreds base pairs. However, within the class Gammaproteobacteria, total origin length varies by as much as fivefold (compare *Pastuerella multocida* to *Legionella pneumorphila*). In addition to a greater number of DnaA recognition sites, more distant relatives of *E. coli* have bi- and even tri-partite replication origins, with clusters of DnaA boxes separated by entire genes (Donczew et al. 2014, Moriya et al. 1988). For example, in *Helicobacter pylori*, *Staphylococcus aureus* and *Mycobacterium tuberculosis*, the

replication origin consists of two sub regions flanking the *dnaA* gene (see Figure 5B and Wolanski et al. 2015).

One of the key goals for the experiments presented in this dissertation is to advance our understanding of bacterial replication origin diversity, and provide insights into the reason that, despite the high conservation of DnaA among all bacteria, there are many different ways that DnaA can assemble into a functional pre-RC.

AN EXPERIMENTAL APPROACH TOWARDS ANALYSIS OF RECOGNITION SITE GEOGRAPHY IN *E. COLI ORIC*

In the studies reported here, two specific questions about *E. coli oriC* geography were addressed: 1) are all three high affinity sites required for *oriC* function, and 2) is wild-type nucleotide spacing between recognition sites a rigid or flexible feature of this origin. With regard to the first question, previous studies of cloned *E. coli oriC*, suggested that the high affinity recognition site R2 could be inactivated, but R1 and R4 were required for origin function. Later studies on mutated chromosomal *oriC*s revealed that cloned copies had more stringent requirements and R4 could be eliminated without completely destroying origin activity (Weigel et al. 2001). However, R1 remained absolutely required for origin function under both conditions (Weigel et al. 2001). This discrepancy between the study of plasmid and chromosomally located *oriC* is no longer a concern because it has now become easier to introduce mutations directly into specific *E. coli*

chromosomal sites by the use of recombineering methods (Court et al., 2002). By using strains that carry inducible RED recombinase genes of lambda phage, DNA fragments containing mutant versions of *oriC* can easily replace (via recombination) the wild-type version at the correct chromosomal position, and without significant genetic manipulation used in the past (Court et al. 2002, Kaur et al. 2014). For this dissertation, recombineering was used to re-examine the requirements for high affinity recognition sites in *oriC* function by scrambling each site individually and in combination with other high affinity sites, and a novel functional assay was also developed for these studies (Kaur et al. 2014).

Previous studies were also performed on *E. coli oriC* where spacing between high affinity DnaA recognition sites was altered (Weigel et al. 2001, Woelker and Messer, 1993, Crooke et al. 1993, Messer et al. 1992). Most studies were performed using cloned *oriC* copies, and it was possible to add or delete some bases between sites without disturbing origin function, but some changes inactivated *oriC*. Specifically, spacing changes to the left half of *oriC*, between R1 and R2, inactivated *oriC* (Messer et al. 1992) while insertions or deletions in the right half of *oriC*, between R2 and R4, were tolerated only when the change addition or removal of one helical turn of the DNA (Weigel et al. 2001, Messer et al. 1992, Crooke et al. 1993, Woelker and Messer, 1993).

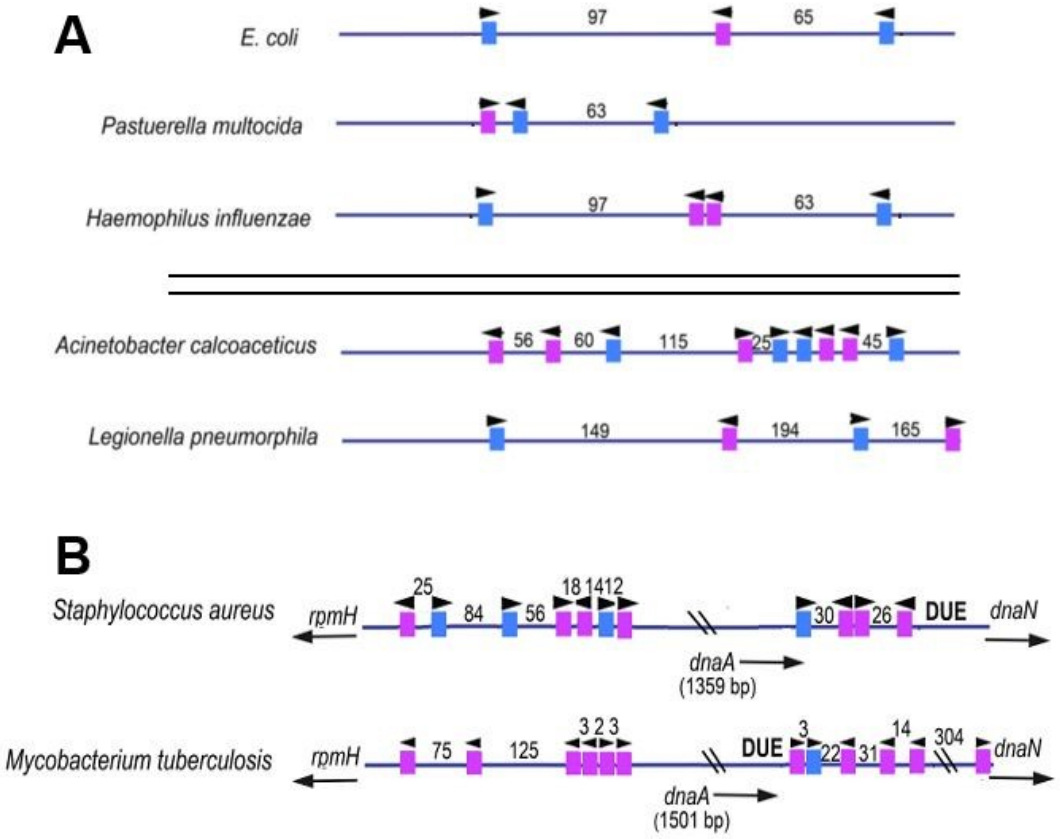

Figure 5. Comparison of replication origin geography from different bacterial types. DnaA recognition sites with the sequence 5`-TTATCCACA-3` are indicated by blue rectangles. Sites that deviate by one or two base pairs are indicated by pink rectangles. Arrowheads above each site indicate relative binding orientation. Numbers in the gap regions denote number of nucleotides between recognition sites. A) Examples of origins from the class Gammaproteobacteria. Those below the double line are notably larger than those above. B) Examples of bacteria with bi-partite replication origins.

However, these studies were performed before the arrayed low affinity recognition sites were identified in *oriC*, and it is unclear whether the results obtained were due to the change in spacing or due to alteration of an arrayed recognition site. With our current knowledge of recognition site positions, DNA fragments containing different spacing and arrangements of high and low affinity recognition sites were evaluated by electrophoretic mobility shift assays (EMSA) to measure the direct effect on DnaA binding. This data was compared to previous studies and used to generate new rules for the assembly of *E. coli* pre-RC.

The results obtained from the studies reported here support a model for the stringent positioning of proximal DnaA recognition sites due to the requirement for DnaA-DnaA interactions at both early and late stages of pre-RC assembly to properly regulate replication timing. We find, for the first time, the arrangement of high affinity DnaA recognition sites topologically constrain *oriC* in an origin recognition complex and that the placement of low affinity DnaA arrays allow DnaA to convert the constrained complex into one that imparts sufficient torsional stress to unwind *oriC*. The later complex requires that the spacing between low affinity recognition sites does not exceed three base pairs. The use of specifically placed initiator protein binding sites to direct staged assembly of a replication complex that regulates origin topology is likely a feature utilized by all bacteria.

METHODS AND MATERIALS

CHEMICALS, ENZYMES, PRIMERS AND PROTEINS

Reagent grade chemicals were purchased from IBI®, Fisher Scientific®, or Sigma® and culturing medias from Difco™. AcrylaGel™, SequaGel® and Ecoscint™ were purchased from National Diagnostics®. Enzymes were from New England Biolabs®, Bioline or Affymetrix®. Primers were ordered from Integrated DNA Technologies (IDT) or Life Technologies®.

BACTERIAL STRAINS AND PLASMIDS

Table 1 is the complete list of cell strains used for this study, along with relevant genotype information. Table 2 lists all plasmids used for this study. pOC170 (3852 bp) carries both the pBR322 replication origin and a copy of the *E. coli* wild type *oriC*, as well as the gene (*bla*) for ampicillin resistance (Langer et al., 1996). pZL411 (7273 bp) carries a poly His tagged DnaA gene regulated by the T7 promoter (Li and Crooke 1999)

Table 1. Bacterial strains used in this study.

Name of Strain	**Genotype**	**Supplier and (Reference)**
XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10 (Tetr)].	Stratagene®
BL21(DE3) pLysS	F- ompT hsdSB (rB-mB-) gal dcm (DE3) pLysS (CamR)	Life Technologies®, Li and Crooke 1999

HME6	W3110 Δ(argF-lac)U169 galKtyr145UAG [λcI857 Δ(cro-bioA)]	Ellis et al., 2001
ALC400	(HME6, *oriC*::pKN1562 (clockwise), asnB::Tn10)	Kaur et al., 2014
EH3827	CM1565 *zia*::pKN500 Δ*dnaA* mad-1	Hansen *et al.* 1986
ACL401	EH3827 ΔdnaA [λcI857 Δ(cro-bioA)] mutS::CAT $galK_{tyr145UAG}$	Kaur et al., 2014
JEG22	MG1655 asnA::Km, asnB::Tn10	Kaur et al., 2014

Table 2. Plasmids used in this study.

Plasmids	**Relevant Mutations**	**Reference**
pOC170	Wild type *oriC*	Langer et al., 1996
pZL411	Wild type DnaA	Li and Crooke 1999
pCC59	Messer R1 scramble	This study
pMV14	Messer R2 scramble	Personal Communication (Mansi Vora)
pMV13	Messer R4 scramble	Personal Communication (Mansi Vora)
pCC37	Messer R1&R4 scrambles	This study
pCC36	Messer R2&R4 scrambles	This study
pCC35	Messer R1&R2 scrambles	This study
pAL110	Wild type DnaA and *copA*	Kaur et al., 2014
pAL111	*copA* only	Kaur et al., 2014

PREPARATION OF CALCIUM CHLORIDE COMPETENT CELLS

E. coli does not naturally take up exogenous DNA, but can be made competent to do so by treatment with calcium chloride. To make *E. coli* competent for genetic transformation, cultures were plated on Luria-Bertani (LB-thy) agar plates, pH 7.0 (10 mg/ml tryptone, 5 mg/ml yeast extract, 10 mg/ml NaCl, 15 mg/ml agar) supplemented with 10 µg/ml thymine and incubated at 37°C overnight. Single colonies were used to inoculate 10 ml of LB-thy media that was incubated at 37°C with shaking overnight. 40 ml of fresh LB-thy media was inoculated with 0.4 ml of the overnight culture and grown to an optical density of 0.3-0.4 A_{600}. The culture was transferred to a sterile 50 ml Falcon® tube and rapidly chilled by incubating in an ice-water slurry for 20 minutes. Cells were harvested by centrifugation in an Eppendorf® 5810 (A462) at 3,000 RCF for 10 minutes at 4°C. Culture media was decanted, and the cell pellet gently re-suspended in 30 ml of sterile ice cold 0.1 M $CaCl_2$ and then incubated in ice-water for 30 minutes. Cells were again harvested by centrifugation as before and re-suspended in 1 ml of sterile ice cold 0.1 M $CaCl_2$ containing 15% glycerol. Competent cells were divided into aliquots of 25 or 50 µl and stored at -70°C.

TRANSFORMATION OF CALCIUM CHLORIDE COMPETENT CELLS WITH PLASMID DNA

Purified plasmid DNA (150 to 250 ng) was mixed with competent cells that had been thawed on ice. Cells were heat pulsed at 42°C for 45 seconds then

immediately placed in ice for 2 minutes. One ml of Luria-Bertani (LB-thy) media, pH 7.0, was added and samples were incubated at 37°C while shaking at 300 rpm for 1 hour in an Eppendorf® Thermomixer® R. Uptake and expression of the template DNA was determined by plating known aliquoits on selective media, typically LB-thy agar plates (15 mg/ml agar) containing 100 µg/ml ampicillin. Plates were incubating at 37°C overnight.

CLONING AND ISOLATION OF PLASMID DNA

XL1-Blue *E. coli* (50 µl), previously washed with calcium chloride, were transformed per the protocol above with plasmid DNA and plated on selective media. A single colony was used to inoculate Luria-Bertani (LB-thy) broth, pH 7.0 (10 mg/ml tryptone, 5 mg/ml yeast extract, 10 mg/ml NaCl) supplemented with 10 µg/ml thymine and 100 µg/ml ampicillin and incubated at 37°C while shaking at 300 rpm overnight. Plasmid DNA was harvested and purified using the Qiaprep® Spin Miniprep Kit (Qiagen®). DNA concentrations were determined by measuring absorbance at OD_{260} using a Thermoscientific® Nanodrop 2000c spectrophotometer.

ISOLATION OF HIS_{10} TAGGED DNAA

Over-expression and isolation of his-tagged DnaA protein was performed as described by Li and Crooke (1999). The DnaA over-expression plasmid pZL411 was transformed into competent BL21(DE3) pLysS (Life Technologies®) cells and plated on Luria-Bertani (LB-thy) plates supplemented with 100 µg/ml ampicillin

and 34 μg/ml chloramphenicol. A single colony was used to inoculate 5 ml of identical media and the culture was incubated at 37°C while shaking at 300 rpm overnight. This culture was used to inoculate 500 ml of fresh LB-broth containing 100 μg/ml ampicillin, 34 μg/ml chloramphenicol and incubation was continued at 37ºC with shaking until a density of 0.4 - 0.6 A_{600} was reached. Expression of DnaA was then induced by adding 500 μl of 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubating for 90 minutes. After transferring the culture to pre-chilled centrifuge bottles, cells were harvested by centrifugation in a Sorvall™ RC2B (SA-600 rotor) at 3,800 RCF for 5 minutes at 4°C. All subsequent steps were performed at 4°C. The media was decanted, and the pellet re-suspended in 10 ml of ice cold binding buffer (20 mM sodium phosphate, pH 7.8, 500 mM sodium chloride, 5 mM imidazole). Samples were flash frozen in liquid nitrogen, and stored at -70°C overnight. To purify DnaA protein, cells were thawed on ice then 100 μl of 10 mg/ml lysozyme was added. This was incubated on ice for 15 minutes, tubes were exposed to three liquid nitrogen freeze-thaw cycles. After the final thawing, 10 μl of 10 mg/ml *Dnase* I and 100 μl of 1 M $MgCl_2$ were added and the lysate was incubated on ice for 15 minutes. Samples were centrifuged in a Beckman/Coulter® Optima (LE-80K) ultracentrifuge in a Type 90 Ti rotor at 100,000 RCF for 30 minutes at 4°C. Clear lysates were mixed with 2.5 ml of prepared Ni-NTA Agarose resin (Life Technologies®) in a 50 ml Falcon® tube and incubated for 30 minutes at 4°C with gentle rotation. Ni-NTA resin was prepared by washing with 10 ml of molecular biology grade water, then 8 ml of binding

buffer. After mixing with the lysate, the resin was allowed to settle for 5 minutes, and the supernatant was pulled off by pipetting. The resin was then washed with 2.5 ml of wash buffer (20 mM sodium phosphate pH 7.8, 500 mM sodium chloride, 100 mM imidazole) for 10 minutes with gentle rotation. After settling, the supernatant was removed and the wash step repeated two more times. The resin was then rinsed with 20 ml of urea/binding buffer (20 mM sodium phosphate, pH 7.8, 500 mM sodium chloride, 5 mM imidazole, 7 M Urea) for 30 minutes with gentle rotation. After settling, the supernatant was removed and the resin was washed with 7.5 ml of Buffer HD-1 (50 mM PIPES-KOH, pH 6.8, 10 mM magnesium acetate, 200 mM ammonium sulfate, and 20% (w/v) sucrose) with gentle rotation for 6 minutes. The wash step was repeated twice. The His-DnaA protein was eluted from the resin by adding 2.5 ml of buffer HD-2 (50 mM PIPES-KOH, pH 6.8, 10 mM magnesium acetate, 200 mM ammonium sulfate, 20% sucrose, 1 M imidazole) and gently mixed for six minutes. After settling, the supernatant was removed and kept on ice. The elution step was repeated one time to remove as much protein as possible. Eluted samples were pooled and transferred into cellulose ester dialysis membrane MWCO 20,000D (Spectra/Por®) and dialyzed against 750 ml HD-3 Buffer (50 mM PIPES-KOH, pH 6.8, 10 mM magnesium acetate, 200 mM ammonium sulfate, 20% sucrose, 0.1 mM EDTA, 2 mM dithiothreitol) for 2 hours. The dialysis buffer was refreshed and the protein dialyzed overnight. The following day the dialyzed protein was centrifuged in the Beckman/Coulter® Optima (LE-80K) ultracentrifuge (Type 90 Ti) at 100,000 RCF

for 15 minutes at 4°C. The purified protein was then divided into single-use aliquots, flash frozen and stored at -70°C. Protein binding activity was determined using high resolution DMS foot-printing as described below.

PURIFICATION OF DOUBLE STRANDED OLIGONUCLEOTIDES

Single strand oligonucleotides were purchased from IDT or Life Technology® and arrived as desalted, lyophilized pellets. Oligos were re-suspended in Qiagen® EB buffer (10 mM Tris-Cl, pH, 8.5) to a concentration of 10 µg/µl. Complementary oligos were mixed together in equal parts and annealed by incubating at 90°C for 10 minutes then at 50°C for 10 minutes and cooled to room temperature on the bench top. 6X loading dye (30% Glycerol, 60 mM Tris-Cl, pH, 7.5, 0.1% Bromophenol Blue) was added. Samples were separated by electrophoresis through a 10% polyacrylamide gel (16.7 ml 30% acrylamide, 8.3 ml 2% bisacrylamide 5 ml 5 X TBE, 5 ml 25% glycerol, 15 ml molecular biology grade water, 400 µl 10% Ammonium per-sulfate and 40 µl TEMED). Gels were run at 12.3 V/cm (13 cm x 17.5 cm x 0.1 cm) in 44.5 mM Tris-Borate, 1 mM EDTA pH 8 buffer for 1 hour at room temperature before loading of samples. Samples were then electrophoresed at 12.3 V/cm for 2 hours. The gel was transferred to a silica-coated thin layer chromatography plate. Double stranded DNA complexes were visualized using short-wave (254nm) UV shadowing. DNA bands were excised, placed in 125 µl of 0.3 M sodium acetate pH4.5, and incubated at 37ºC overnight. After a quick spin in a table-top centrifuge, the liquid was

transferred to a clean 1.7 ml microfuge tube and an equal volume (typically 100 µl) of phenol:chloroform:isoamyl alcohol (25:24:1) was added followed by vigorous vortexing. Samples were then centrifuged in a table-top centrifuge at 16,800 RCF for 2 minutes to separate the organic and aqueous phases. The aqueous (upper) phase was transferred to a clean microfuge tube and an equal volume of choloform:isoamyl alcohol (24:1) was added. Samples were vortexed and centrifuged as before and the upper phase transfer to a clean microfuge tube. The double-stranded DNA was precipitated by adding 0.4 volume of 3 M sodium acetate pH 5.2 and two volumes 100% absolute ethanol and incubated at -70ºC for 1 hour or at -20ºC overnight. DNA was pelleted by centrifugation at 22,000 RCF for 12 minutes at 4ºC. The supernatant was decanted, the pellet washed with 70% ethanol and centrifuged at 22,000 RCF for 2 minutes at 4ºC. The DNA pellet was air dried for 20 minutes then re-suspended in 30 µl of molecular biology grade water. Samples were applied to, pre-washed, spin columns packed with 0.75 ml of Bio-gel® P-6 matrix (Bio-Rad®) and centrifuged at 410 RCF for 4 minutes. Spin columns were prepared by centrifuging at 410 RCF for 2 minutes to remove 1X SSC (0.15M NaCl, 15mM NaCitrate pH 7 + 0.02% NaAzide) storage buffer, followed by 3 washes with 800 µl of molecular biology grade water. A final 1 minute centrifugation at 410 RCF was done to removes residual water. The concentration of pure double stranded DNA was determine by absorbance at OD_{260} using a Thermoscientific® Nanodrop 2000c spectrophotometer. Oligos were

diluted to a final concentration of 10 pmol/µl in molecular biology grade water and stored at -20°C.

ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)

DNA-protein interactions were detected by DNA mobility shift assays (complexes slow the mobility of DNA fragments during electrophoresis). Gel purified, double strand oligos were radiolabeled with ^{32}P (5' end) using the USB Affymetrix® optikinase protocol; 2.5 pmol of oligo was incubated with 1.5 µl of $\gamma^{32}P$ ATP (6000 µci/mmol) (Perkin Elmer®) and 5 Units Optikinase in a total volume of 12.5 µl of 1X optikinase reaction buffer at 37ºC for 30 minutes followed by 65ºC for 10 minutes. Molecular biology grade water was added to bring the final volume to 25 µl. Samples were applied to pre-washed spin columns packed with 0.75 ml of Bio-gel® P-6 matrix (Bio-Rad®) and spun in the centrifuge at 410 RCF for 4 minutes followed by 3 washes with 800 µl of molecular biology grade water. After each spin at 410 RCF for 2 minutes, a final 1 minute centrifugation was performed at 410 RCF to remove residual water. An aliquoit (0.8 µl) of radiolabeled sample was mixed with 5 ml of Ecoscint™ scintillation cocktail and radioactivity was measured using a Beckman LS 6500 Scintillation Counter.

For EMSA, 4 nM of radiolabeled oligo was mixed with different concentrations of purified DnaA on ice in 20 µl buffer (final volume) consisting of 20 mM HEPES•KOH pH 8, 1 mM EDTA, 200 ng poly(dI-dC), 2.5 mM Magnesium acetate, 4 mM DTT, 0.1 mM ATP, 5 mg/ml BSA, 0.2% Triton X-100,

and 5% Glycerol. Reactions were incubated at 37ºC for 8 minutes and 15 µl immediately loaded on a 7% polyacrylamide gel (11.7 ml 30% acrylamide, 5.8 ml 2% bisacrylamide 5 ml 5 X TBE, 2.5 ml 50% glycerol, 25 ml molecular biology grade water, 400 µl 10% Ammonium per-sulfate and 40 µl TEMED). Gels were run at 12.3 V/cm (13 cm x 17.5 cm x 0.2 cm) in 44.5 mM Tris-Borate, 1 mM EDTA pH 8 buffer for 1 hour at room temperature prior to sample loading. Samples were electrophoresed for 2 hours 15 minutes at 12.3 V/cm at room temperature. After transferring to 3 MM Whatman® Paper, gels were dried on a Bio-Rad® vacuum gel dryer (Model 583) at 90ºC for 10 minutes then at 80ºC for 35 minutes. Gels were exposed on a Bio-Rad® phosphor-imaging screen for 1-2 days and scanned using a Bio-Rad® Phosphorimager (Model 170-940). Data capture and analyses was done using Quantity One (Bio-Rad®) software v. 4.6.9.

Table 3. EMSA oligonucleotides used in this study (only top strand shown).

Description	**Sequence (5` to 3`)**
R5m-2bp-R4	CAGTCATTGGTCATTCACAGATTATCCACAGTAG ATCGCA
R5m-2bp-Scramble	CAGTCATTGGTCATTCACAGAACACATATTGTAG ATCGCA
Scramble-2bp-R4	CAGTCATTGGACACATATTGATTATCCACAGTAG AT CGC
R5m--C1inv---R4	CAGTCATTGGTCATTCACAGCTGTCAGGAAGAGT TATCCACAGTAGATCGCA
R5minv--C1---R4	CAGTCATTGGTGTGAATGAGCTTCCTGACAGAGT TATCCACAGTAGATCGCA
R5m--C1---R4inv	CAGTCATTGGTTATACACAGCTGTGGATAAGAGT GTGGATAAGTAGATCGCA
R5m--C1---R4	CAGTCATTGGTCATTCACAGCTTCCTGACAGAGTT ATCCACAGTAGATCGCA

R5m-C1--R4	CAGTCATTGGTCATTCACAGTTCCTGACAGGTTATCCACAGTAGATCGCA
R5m--C1--R4	CAGTCATTGGTCATTCACAGCTTCCTGACAGGTTATCCACAGTAGATCGCA
R5m---C1--R4	CAGTCATTGGTCATTCACAGACTTCCTGACAGGTTATCCACAGTAGATCGCA
R5m----C1--R4	CAGTCATTGGTCATTCACAGACGTTCCTGACAGGTTATCCACAGTAGATCGCA
R5m-0bp-R4	CAGTCATTGGTCATTCACATTATCCACAGTAGATCGCA
R5m-1bp-R4	CAGTCATTGGTCATTCACAGTTATCCACAGTAGATCGCA
R5m-2bp-R4	CAGTCATTGGTCATTCACAGATTATCCACAGTAGATCGCA
R5m-3bp-R4	CAGTCATTGGTCATTCACAGACTTATCCACAGTAGATCGCA
R5m-4bp-R4	CAGTCATTGGTCATTCACAGACGTTATCCACAGTAGATCGCA
R5m-5bp-R4	CAGTCATTGGTCATTCACAGACGCTTATCCACAGTAGATCGCA
R5m-7bp-R4	CAGTCATTGGTCATTCACAACACATATTATCCACAGTAGATCGCA
R5m-10bp-R4	CAGTCATTGGTCATTCACAACACATATTATTATCCACAGTAGATCGCA
R5m-13bp-R4	CAGTCATTGGTCATTCACAACACATATTACACTTATCCACAGTAGATCGCA
R5m-14bp-R4	CAGTCATTGGTCATTCACAACACATATTACACATTATCCACAGTAGATCGCA
R5m-15bp-R4	CAGTCATTGGTCATTCACAACACATATTACACATTTATCCACAGTAGATCGCA

SITE DIRECTED MUTAGENESIS

Mutations were introduced to *E. coli oriC* carried on the chimeric plasmid pOC170 (Langer et al., 1996) using a Quick-Change II Site-directed Mutagenesis kit (Stratagene®). Primers for bidirectional Polymerase Chain Reaction (PCR) were designed following the guidelines provided by the mutagenesis kit. For each insertion or deletion of a few bases, primers were designed with the desired

mutation flanked on the 5` and 3` ends by 12 to 15 bases of sequences complementary to the template DNA. For each additional base change 3 to 5 bases of complementary sequence were added to each end of the primers. PCR reactions (50 µl) contained 50 ng of purified pOC170 template DNA, 200 µM dNTP's, 125 ng each of forward and reverse primer, 2.5 Units of Accuzyme™ DNA polymerase (Bioline) in 1 X Taq polymerase buffer, and molecular biology grade water. The thermocycler was programed as follows: Step 1; 95°C for 30 seconds, Step 2; 95°C for 30 seconds, 55°C for 20 seconds, 72°C for 4 minutes (repeated for 15 cycles), Step 3; 72°C for 8 minutes, Step 4; 10°C hold. To remove the pOC170 template, each PCR tube was treated with DpnI (1 µl of 20 Units/µl DpnI) and incubated at 37°C for 10 minutes followed by a second round of digestion. 5 µl of each reaction were mixed with 1µl of 6X loading dye (30% Glycerol, 60 mM Tris-Cl, pH, 7.5, 0.1% Bromophenol Blue) and electrophoresed on a 1% agarose gel until the dye front had migrated 2/3 the length of the gel. Gels were stained with Ethidium Bromide (100 µg/mL) and visualized using short-wave (254nm) UV light. PCR products (5 µl) were transformed into competent XL1-Blue *E. coli* as described above. Mutations were verified by DMS modification and alkaline primer extension as described below. Table 2 contains the mutant plasmids used in this study.

Table 4. Complementary primers used for site directed mutagenesis.

Description	**Sequence**
R1 Scramble Top	GGATCGCACTGCCCAACTATATCCAAGGATCCGGC
R1 Scramble Bottom	GCCGGATCCTTGGATATAGTTGGGCAGTGCGATCC
R2 Scramble Top	GATCAGAATGAGGGGGATATAGTTACTCAA AAACTGAAC
R2 Scramble Bottom	GTTCAGTTTTTGAGTAACTATATCCCCCTCAT TCTGATC
R4 Scramble Top	CAAGCTTCCTGACAGAGGATATAGTTGTAGA TCGCACGATC
R4 Scramble Bottom	GATCGTGCGATCTACAACTATATCCTCTGTC AGGAAGCTTG

IN VITRO DIMETHYL SULFATE (DMS) MODIFICATION AND PIPERIDINE TREATMENT

DNA-protein interactions were detected by DMS footprinting assay. 750 ng of purified supercoiled plasmid DNA was incubated with different concentrations of DnaA in a reaction mix consisting of 40 mM Hepes• KOH pH7.6, 5 mM ATP or ADP, 8 mM $MgCl_2$, 0.32 mg/ml BSA and 30% Glycerol in a total volume of 50 µl prepared on ice. Samples were incubated at 38°C for 7 minutes, 6 µl of 1.4% dimethyl sulfate (diluted in molecular biology grad water) was added and incubated at 38°C for 5 minutes. Reactions were quenched by adding 200µl of DMS stop solution (3 M ammonium acetate, 1 M 2-mercaptoethanol, 250 µg/ml tRNA, 20 mM EDTA). DNA was precipitated by adding 600 µl of 95% ethanol and incubated at -70°C for 1 hour or at -20ºC overnight. DNA was pelleted by centrifuging at 22,000 RCF for 12 minutes at 4ºC. The supernatant was decanted,

the pellet washed with 70% ethanol then centrifuged at 22,000 RCF for 2 minutes at 4ºC. The DNA pellet was air dried for 20 minutes before being re-suspended in 100 µl of 10% piperidine solution (diluted in molecular biology grad water) and incubated at 90°C for 30 minutes. After cooling to room temperature samples were applied to Bio-gel® P-6 matrix (Bio-Rad®) packed with 0.75 ml and centrifuged at 410 RCF for 4 minutes. Spin columns were prepared by centrifugation at 410 RCF for 2 minutes to remove the 1X SSC storage buffer (0.15 M NaCl, 15 mM Na-Citrate pH 7, 0.02% Na-Azide) and a second 1 minute spin. Samples were stored at -20°C ready for alkaline primer extension.

ALKALINE PRIMER EXTENSION

The primer RS4 (5'-GGATCATTAACTGTGAATG-3') which hybridized to bases 272–290 or SR4 (5'-GTATACAGATCGTGCGATC-3') bases 124–142 of *oriC* were used to analyze the DMS modification of the top or bottom strand of *oriC* respectively. Primers were 5 prime radiolabeled with ^{32}P as follows; 20 pmol of primer DNA was incubated with 6.25 µl of [γ-32P] ATP, 6000 µci/mmol (Perkin Elmer®), and 10 Units of Polynucleotide Kinase (NEB®) in 1 X polynucleotide kinase buffer in a total volume of 50 µl. Reactions were incubated at 37°C for 1 hour followed by 10 minutes at 70°C. After cooling to room temperature reactions were applied to prepared spin columns packed with 0.75 ml of Bio-gel® P-6 matrix (Bio-Rad®) and centrifuged at 410 RCF for 4 minutes. Spin columns were prepared by centrifugation at 410 RCF for 2 minutes to remove the 1X SSC storage

buffer (0.15 M NaCl, 15 mM Na-Citrate pH 7, 0.02% Na-Azide) and a second 1 minute spin. 1 µl of radiolabeled sample was mixed with 5 ml of Ecoscint™ fluid and quantified using a Beckman LS 6500 Scintillation Counter.

Primer extensions were performed by adding 4 µl of 100 mM NaOH to 35 µl piperidine treated DNA. Reactions were set up on ice. 1 µl of radio labelled RS4 or SR4 was added and samples incubated at 80ºC for 2 minutes then placed on ice of 5 minutes. 5 µl of 10X TMD (0.5 M Tris-HCl, pH 7.2, 0.1 M magnesium sulfate, 2 mM dithiothreitol) was added and incubated at 55ºC for 3 minutes then places on ice for 5 minutes. 5 µl of 5 mM dNTP and 1 µl of 1 Units/µl Klenow fragment DNA polymerase I (NEB®) (5 Units/µl Klenow stock diluted in 1X NEB Buffer 2 (50 mM sodium chloride, 10 mM Tris-HCl, 10 mM magnesium chloride, 1 mM dithiothreitol, pH 7.9) were added and incubated at 50ºC for 10 minutes. Primer extension reactions were stopped by adding 17 µl of quencher (10 mM Tris-HCl, pH 7.5, 5 mM magnesium chloride, 7.5 mM dithiothreitol). DNA was precipitated by adding 190 µl of 95% absolute ethanol and incubated at -20ºC overnight. DNA was pelleted by centrifugation at 20,800 RCF for 12 minutes. The supernatant was decanted, the pellet washed with 70% ethanol then centrifuged at 20,800 RCF for 2 minutes. Samples were air dried for 20 minutes before being re-suspended in 6 µl of footprint tracking dye (90% formamide, 20 mM EDTA, 0.5% (w/v) bromophenol blue, 0.5% (w/v) xylene cyanol). Samples were incubated at 42ºC for 2 minutes than gently mixed. After a quick spin in a table-top micro-centrifuge, samples were incubated at 80ºC for 5 minutes. Samples were then

loaded on 6% denaturing polyacrylamide sequencing gel (66 ml of UreaGel System Diluent, 24 ml of UreaGel concentrate, 10 ml of 10X TBE, 800 µl 10% Ammonium per-sulfate and 40 µl TEMED) that had been pre-ran at 1.39 Watts/cm (34 cm x 39.5 cm x 0.044 cm) in 90 mM Tris-Borate, 2 mM EDTA pH 8 buffer for 1 hour at room temperature. Samples were electrophoresed at 1.27 watts/cm, at room temperature, until 20 minutes after the bromophenol blue dye ran off the bottom of gel (typically 2 hours 25 minutes). Gels were transferring to 3 MM Whatman® Paper and dried on a Bio-Rad® vacuum gel dryer (Model 583) at 90ºC for 10 minutes then at 80ºC for 35 minutes. Gels were exposed on a Bio-Rad® phosphor-imaging screen for 2-3 days and scanned using a Bio-Rad® Phosphorimager (Model 170-940). Data capture and analyses was done using Quantity One (Bio-Rad®) software v. 4.6.9.

MUNG BEAN NUCLEASE DIGESTION

To detect *oriC* DNA unwinding on supercoiled templates, mung bean nuclease assays were performed. 500 ng of purified supercoiled plasmids DNA was incubated with a range of DnaA concentrations in reaction mix (10 mM Tris, 1 mM EDTA, pH 7.0 and molecular biology grade water to a final volume of 18 µl) prepared on ice. DNA-protein complexes were incubated at 37°C for 7 minutes before 2 µl of 0.05 units/µl (1 µl of 10 units/µl diluted in 200 µl of 1X Mung Bean Nuclease Reaction Buffer: 30 mM NaCl, 50 mM sodium acetate, 1 mM $ZnSO_4$, pH 5) Mung Bean Nuclease was added. Samples were further incubated at 37°C for 15

minutes. Reactions were stopped by adding 100 µl quench (3 M ammonium acetate, 250 µg ml−1 tRNA) and DNA was precipitated by adding 300 µl cold of 95% absolute ethanol and incubating at -20ºC overnight. DNA was pelleted by centrifugation at 22,000 RCF for 12 minutes at 4ºC. The supernatant was decanted, the pellet washed with 70% ethanol then centrifuged at 22,000 RCF for 2 minutes at 4ºC. The DNA pellet was air dried for 20 minutes before being re-suspended in 50 µl EB (10 mM Tris-Cl, pH 8.5). Samples were stored at -20ºC ready for alkaline primer extended using the primer SR4 as described above.

RECOMBINEERING

Mutant versions of *oriC* were placed into the chromosome at the correct position by recombineering following the protocol of Court et al., 2002. Briefly, single stranded (Table 5) or double stranded oligonucleotides (see below) were introduced (by electroporation) into *E. coli* strains which carried a copy of the RED recombination genes from the bacteriophage lambda which are under the control of the lambda cI857 promoter allowing induced expression only at 42ºC. Following induction, these strains become able to recombine linear DNA into homologous regions of the chromosome. To test *oriC* replacements for function, it is important to note that the RED strains ACL400 and 401 used in these experiments cannot initiate DNA synthesis from their chromosomal *oriC* copies and must use a chromosomally-integrated copy of a plasmid replication origin (*oriR* derived from the R factor R1) that does not require DnaA. *OriR* can be inactivated in the

presence of the plasmid p*copA* that expresses the *oriR* repressor normally used for copy control. By co-transforming with p*copA* it is possible to repress the growth of ACL400 unless a functional *oriC* was electroporated into the chromosome (see below). The *oriC* replacements were then verified for chromosomal position and presence of the mutation by PCR and DNA sequence analysis (Genewiz Corp.).

Mutant *oriC*s were moved from pOC170 to the chromosome by PCR. Primers OR84 (5'-CTCAACTTTGTCGGCTTGAG-3') and OR10 (5'-ATCCCATACTTTTCCACAGG-3') amplify a 761-base pair segment consisting of the mutant *oriC* flanked by additional sequences complementary to the adjacent genes on the chromosome. PCR reactions contained 5 ng of purified plasmid template DNA, 800 μM dNTP's, 0.4 μM each of forward and reverse primer, 2.5 units of Accuzyme™ DNA polymerase (Bioline®) in 1X polymerase enzyme buffer, and molecular biology grade water to a final volume of 50 μl. The thermocycler was programed as follows: Step 1 - 95°C for 3 minutes, Step 2 - 95°C for 15 seconds, 52°C for 15 seconds, 72°C for 1 minutes, repeated for 30 cycles, Step 3 - 72°C for 3 minutes, Step 4 - 10°C indefinite hold. Double stranded PCR products were purified using the Qiagen® PCR purification kit and eluted in molecular biology grade water.

Cultures of RED inducible strains ACL400 or ACL401 (Table 1) were prepared by inoculating 45ml fresh Luria-Bertani (LB-thy) (10-5-5) broth, pH 7.0 (10 mg/ml tryptone, 5 mg/ml yeast extract, 5 mg/ml NaCl) supplemented with 10

μg/ml thymine with 450 μl of overnight culture and incubated at 32°C while shaking until the culture reached a density of 0.4-0.5 A_{600}. 20 ml of culture was transferred to a 250 ml baffled Erlenmeyer flask and incubated at 42ºC for precisely 15 minutes to induce expression of the λ RED recombination genes. The culture was immediately put in ice/water and incubated for 10 minutes. Samples were transfer to sterile pre-chilled 30 ml Oak Ridge Centrifuge tubes and centrifuged at 6500 x RCF for 10 minutes at 4°C. The media was decanted, and the cell pellet was gently re-suspended, by pipetting, in 30 ml ice cold sterile distilled water. Samples were centrifuged and decanted as above then gently re-suspended in 1ml ice cold sterile distilled water. Cells were then transferred to pre-chilled, sterile, 1.7 ml microfuge tubes and centrifuged in a table-top micro-centrifuge at 22,000 RCF for 30 seconds at 4°C. The supernatant was removed, and the pellet re-suspended in 200 μl ice cold sterile distilled water. 50 μl of cells was transferred to a pre-chilled, sterile, electroporation cuvette (50 mm x 1 mm gap) and 100 ng of purified double or single-stranded DNA along with 100 ng of p*copA* (Table 2) when appropriate.

Samples were electroporated using a MicroPulser™ (Bio-Rad®) set at 1.8 kVolts followed by immediate recovery in 1 ml LB-Thy (10-5-5) media and incubated at 32°C for 2 hours. Samples were centrifuged in a table-top micro-centrifuge at 16,800 RCF for 30 seconds at room temperature, decanted and re-suspended in 500 μl LB-thy (10-5-5) media before being spread on LB-thy agar plates containing 50 ug/ml ampicillin. Plates were incubated at 32°C for 22 to 24

hours. Resulting colonies were screened for sensitivity to kanamycin and resistance to ampicillin (p*copA*) by patch plating on LB-thy agar plates containing 30 µg/ml kanamycin or 50 µg/ml ampicillin by incubating at 32°C overnight.

Table 5. Recombineering primers used in this study.

Description	**Sequence (5` to 3`)**
R1 Scramble Bottom	GGTTGTTGATCTTAAAAGCCGGATCCTTGGA TATAGTTGGGCAGTGCGATCCTAATATGTGA TCACAA
R2 Scramble Bottom	GGTAGTTATCCAAAGAACAACTGTTGTTCAG TTTTTGAGTAACTATATCCCCTCATTCTGATC CCAGCTTATACGGTCCAGGATCACC
R4 Scramble Bottom	GGGTTAATTTACTCAAATAAGTATACAGATC GTGCGATCTACAACTATATCCTCTGTCAGGA AGCTTGGATCAACCGGTAGTTATCCAAAG AACAAC

COLONY PCR

Detection primers (Table 6) were designed so that their 3`-end carried either the mutant or wild type sequence. These primers were paired with either a forward primer OR9 (5'- CACGGCCACCGCTGTAATTAT- 3') or a reverse primer OR10 (5' - ATCCCATACTTTTCCACAGG- 3') as was appropriate. Reactions were set up on ice as follows: 12.5 µl of NEB 2X Taq Mix, 1 µl of diluted cells (prepared from a small amount of a fresh colony (suspended in 4µl of molecular biology grade by pipetting up and down), 2.5 µM of forward or reverse primers, 2.5 µM of detection primer and molecular biology grade water to a final volume of 25 µl. PCR cycles parameters were as follows: Step 1; 95°C for 5 minutes, Step 2; 95°C for 15 seconds, *66°C to 59°C for 15 seconds (* Annealing temperature was

decreased by 0.5°C per cycle), 68°C for 1 minute, repeated 14 cycles, Step 3; 95°C for 15 seconds 59 °C for 15 seconds, 68°C for 1 minute, repeated 20 cycles, Step 4; 68°C for 5 minutes, Step 5; 10°C indefinite hold. 5 µl PCR product was mix with 1µl 6X loading dye (30% Glycerol, 60 mM Tris-Cl, pH, 7.5, 0.1% Bromophenol Blue) and electrophoresed on a 1% agarose gels until the dye front had migrated 2/3 the length of the gel. Gels were stained with Ethidium Bromide 100µg/mL and visualized using short-wave (254nm) UV light.

Table 6. Mutation detection primers.

Description	**Sequence**
Wild-type R1 Reverse primer	AGCCGGATCCTTGTTATCC
R1 Scramble Reverse primer	AGCCGGATCCAACGATATA
Wild-type R2 Reverse primer	GTTCAGTTTTTGAGTTGTG
R2 Scramble Reverse primer	GTTCAGTTTTTGAGTAACT
Wild-type R4 Reverse primer	GATCGTGCGATCTACTGTGG
R4 Scramble Reverse primer	GATCGTGCGATCTACAACTA
gidA Universal forward primer	CACGGCCACCGCTGTAATTAT
mioC Universal reverse primer	ATCCCATACTTTTCCACAGG

PHAGE PREPARATION AND TRANSDUCTION

Mutant origins on the chromosome of ACL400 (Table 1) were moved to JEG22 (Table 1) by P1 phage transduction so that origin function could be studied in cells able to grow at 37°C and normal growth conditions. 50µl of overnight mutant ACL400 culture grown in LB-thy was used to inoculate 5 ml of fresh LB-

thy broth supplemented with 0.2% glucose and 5 mM CaCl2. The culture was incubated at 32°C while shaking for 30 minutes after which 100 µl P1 phage stock [5 x 10^8 phage/ml] was added and the incubation continued until the culture lysed, typically 2 to 3 hours. The reaction was transferred to a sterile 15 ml Falcon® tube, 100 µl of chloroform was added and vigorously vortexed. P1 phage was separated from unlysed cells by centrifugation in an Eppendorf 5810R (A-4-62 rotor) at 3200 RCF for 10 minutes. The supernatant was transferred to a clean 15 ml Falcon® tube, 100 µl of chloroform was added and the sample was vigorously vortexed. The mutant P1 phage was stored at 4°C.

For P1 phage transductions, 5 ml overnight cultures of JEG22 (Table 1) were centrifuged at 1500 RCF for 10 minutes to harvest the cells. The media was decanted, and the cells re-suspended in 2.5 ml of 10 mM $MgSO_4$ and 5 mM $CaCl_2$. The cells were divided into 100 µl aliquots and placed in sterile 1.7 ml micro-centrifuge tubes. Different volumes (0, 10 µl, 50 µl and 100 µl) of mutant phage were added to each tube followed by incubation at 30°C for 30 minutes without shaking. 100 µl of 1 M sodium citrate and 1 ml of LB-thy broth were added and the samples were incubated at 37°C for 1hr. Infected cells were then centrifuged at 15,000 RCF for 15 seconds. The supernatant was then decanted and the cells re-suspended in 1 ml minimal salts medium without MgSO4 (2 mg/ml NH_4Cl, 6 mg/ml Na_2HPO_4, 3 mg/ml KH_2PO_4, 3 mg/ml NaCl). Cells were washed once more and then plated on Glucose minimal salts agar (minimal salts medium, 0.1% Dextrose, 1 mM MgSO4) containing tetracycline (10 ug/ml). The plates were

incubated at 37°C overnight to selecting for cells able to grow in the absence of asparagine. The presence of the transduced *oriC* was confirmed by nucleotide sequence analysis (Genewiz Corp.)

FLOW CYTOMETRY

To evaluate chromosomal replication timing of transduced JEG22 during the cell cycle, run-out experiments were performed and DNA content was measured by flow cytometry. Cultures were grown in minimal salts media (2 mg/ml NH_4Cl, 6 mg/ml Na_2HPO_4, 3 mg/ml KH_2PO_4, 3 mg/ml NaCl) supplemented with 10% dextrose, 10% casamino acids, 0.1% thiamine and 25 mg/ml $MgSO_4$ and incubated at 37°C while shaking overnight. The overnight culture was used to inoculate fresh minimal media, supplemented as above, at 1/100 volume and growth was continued to a density of 0.2 A_{600}. 300 µg/ml rifampicin and 15 µg/ml cephalexin were added to block new rounds of chromosome replication as well as cell division, and samples were incubated for an additional 3 to 4 hours. Cells in 1 ml of culture were fixed by adding 9 ml 70% ethanol and the sample was stored at 4°C. Prior to flow cytometric analysis, samples were centrifuged at 14,000 RCF for 3 minutes. Supernatant was decanted and the pellet was washed with 1ml Tris-Buffered Saline (TBS) (200 mM Tris, 500 mM NaCl, pH 7.5). The cells were re-suspended in 100 µl TBS containing Vybrant Green (0.5µl/ml TBS). Stained cells (about 3000 cells/ml) were analyzed using an Accuri™ C6 personal flowcytometer. Data from 10,000 cells was collected and analyzed using CFlow software.

RESULTS

DNAA BINDS TO TWO OPPOSITELY ORIENTED HELICALLY-PHASED ARRAYS OF LOW AFFINITY SITES IN *ORIC*

Previous work demonstrated that high affinity DnaA recognition sites are required for further DnaA interactions with lower affinity sites in *oriC* (Schaper and Messer, 1995; Miller et al., 2009). Studies with mutant *oriC* containing a single high affinity recognition site at either R1, R2, or R4 revealed that DnaA occupation was localized to low affinity sites that were immediately adjacent to the strong site, with binding to additional low affinity sites decreasing as a function of distance from the strong site (Miller et al., 2009). Since not all low affinity sites in *oriC* are positioned next to a high affinity R box (Figure 6A), and, since weak sites cannot bind DnaA independently (Schaper and Messer, 1995; Miller et al., 2009), there must be additional parameters that allow cooperative DnaA binding to extend over larger distances to internal weak sites within the DNA "gap" regions between R1 and R2 (97 bp) and between R2 and R4 (65 bp).

To define these parameters, the pattern of DnaA contacts was examined on a supercoiled *oriC* plasmid using high resolution *in vitro* dimethyl sulfate (DMS) footprinting. DnaA contacts are revealed by the distinctive and reproducible

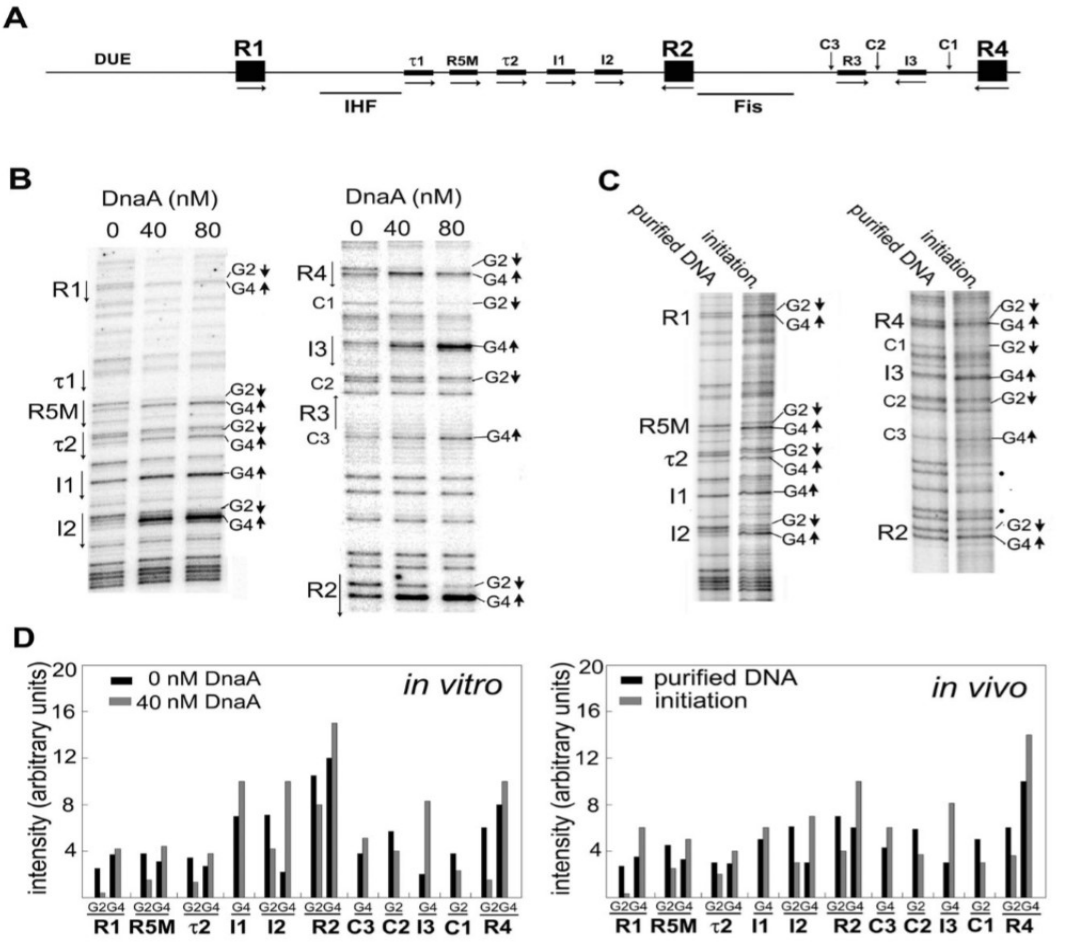

Figure 6. Arrays of DnaA contacts exist in each half of *oriC*. A) Map of *oriC* with positions of DnaA, IHF, and Fis binding sites, as well as the DNA unwinding element (DUE), marked. The three high affinity sites are designated by large black squares, and the low affinity sites are marked by small black rectangles. Horizontal arrows indicate orientation of sites. C1, C2, and C3, and vertical arrows mark regions of unmapped DnaA contacts between R4 and R2. B) *In vitro* DMS modification patterns of *oriC*, after incubating *oriC* plasmids with the indicated concentrations of DnaA prior to treatment with DMS. Two different primers were used to analyze modifications between R1 and R2 (left panel) and between R2 and R4 (right panel). Binding site positions are marked, with their orientations indicated by vertical arrow. Guanosine residues in position 2 or 4 are labeled, with up or down arrows indicating enhanced or diminished DMS sensitivity, respectively. C) *In vivo* DMS modification patterns of *oriC* on purified chromosomal DNA, or DNA isolated from cells aligned at the stage of initiation just prior to helicase loading, measured by LMPCR (Nievera et al., 2006). Guanosine residues in position 2 or 4 are labeled, with up or down arrows indicating enhanced or diminished DMS sensitivity, respectively. Filled circles indicate changes in the modification pattern in the Fis binding region. D) Relative intensities of DMS modification at guanosine residues within DnaA binding sites.

changes to the DMS modification pattern produced when DnaA binds to its 9mer consensus sequence 5'- TGTGNATAA-3' (or variations of this recognition sequence) (Grimwade et al., 2000; McGarry et al., 2004; Nievera et al., 2006). Specifically, the guanosine residue in the fourth position of the 9-mer becomes hypersensitive, and modification of the guanosine in the second position is repressed (Figure 6B, D), causing darker and lighter bands, respectively, on footprinting gels. This well-defined DnaA "footprint" not only allows resolution of individual, closely spaced DnaA recognition sites, but also predicts the nucleotide sequence of putative sites, since any repressed G signal should be in position 2, and any enhanced signal should be in position 4 (McGarry et al., 2004). At DnaA-ATP levels high enough to unwind the DUE of supercoiled *oriC* (80 nM, corresponding to 20–25 molecules per *oriC*) (Bramhill and Kornberg, 1988; Grimwade et al., 2000), a distinctive pattern of DMS modifications, indicative of four regularly-spaced low affinity DnaA contact regions, can be seen in the left gap region between R1 and R2 (Figure 6B left panel), as well as another array of 4 contact sites in the right gap region, between R2 and R4 (Figure 6B right panel). The contacts in the left region were mapped to the previously identified sites I2, I1, τ2, and R5M (McGarry et al., 2004; Kawakami et al., 2005) which are all in the same orientation, with each site separated by 2bp.

In the right half of *oriC*, only two weak DnaA recognition sites (I3 and R3) are reported, and in the *in vitro* DMS footprint, only I3 can be mapped to a contact

site (Figure 6B). Of the three remaining footprints detected in the right gap region, one (C1) is positioned between R4 and I3, and the other two (C2 and C3) partially overlap the R3 sequence. Using the DMS modification pattern, we predicted specific 9mer binding sites for the three contact sites (shown in Figure 7, with correctly positioned suppressed or enhanced G residues marked). A previous report also suggested the existence of a low affinity binding site between R4 and I3, based on weak similarity to the R box consensus (Hansen et al., 2007). Each 9mer in the right *oriC* array is separated by 2 bp, and oriented in the same direction as R4 and I3. Thus, the spacing and number of low affinity contact sites between R2 and R4 is identical to the left array between R2 and R1, but the two arrays are oriented in opposite directions. These data suggest that pre-RC formation in *oriC* proceeds by filling of two arrays of helically-phased, oppositely oriented, low affinity DnaA contacts.

We previously reported that, *in vivo*, the pre-RC is formed when low affinity sites become occupied just prior to the time of initiation of chromosome replication (Cassler et al., 1995; Nievera et al., 2006). To verify that all contact sites in the two arrays are filled *in vivo* when the pre-RC is completely assembled, we performed Ligation Mediated PCR (LMPCR) on chromosomal DNA isolated from DMS-treated cells halted at the stage of helicase loading, compared to chromosomal DNA isolated non-aligned cells (Figure 6C). These *in vivo* footprints show that DnaA contacts the same arrayed sites that were detected by

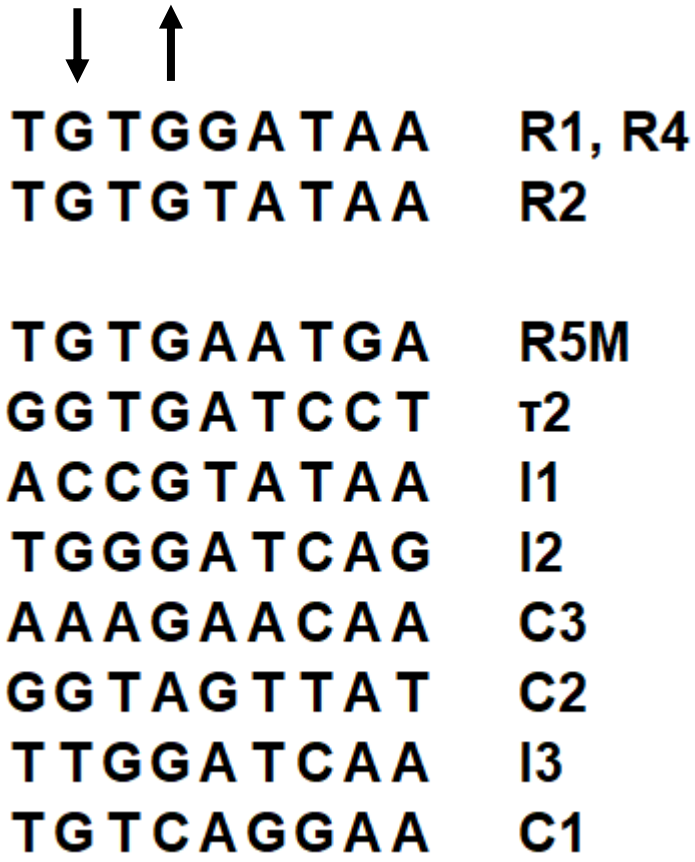

Figure 7. Sequence of DnaA contact sites found in *E. coli oriC*, based on placing G residues hypersensitive to DMS in position 4, and suppressed Gs in position 2.

in vitro footprinting, and the contact pattern is verified by densitometric scans quantifying band intensities (Figure 6D). The bands in the *in vivo* chromosomal footprint for the left half of *oriC* are separated better than those in the right half, since constraints inherent to the LMPCR procedure necessitated the use of a probe for the right half that hybridizes to a region relatively distant from *oriC.* We also note that there are subtle changes in the DMS modification pattern in the Fis region (marked by closed circles in Figure 6C). We do not currently know what protein is causing these changes. We previously reported that Fis is displaced from its primary binding site in *oriC* at the time of initiation (Cassler et al., 1995), but it remains possible that Fis is bound to an overlapping secondary site (Hengen et al., 2003), or that some other, unidentified factor, is binding in this region.

E. COLI ORIC REMAINS FUNCTIONAL AFTER THE LOSS OF INDIVIDUAL DNAA HIGH-AFFINITY RECOGNITION SITES

Pre-RC assembly in bacteria must begin with DnaA binding to high-affinity sites in *oriC* (Schaper and Messer, 1995; Rozgaja et al., 2011). Although *E. coli oriC* contains three high-affinity sites (R1, R2 and R4; see Figure 6A), a previous chromosomal *oriC* mutagenesis study suggested an essential role for only R1 (Weigel et al., 2001). Because this determination was based largely on the inability to recover a strain carrying a non-binding mutation in R1 (Weigel et al., 2001) the possibility remained that viable R1 mutants could be made using alternative methodology. In order to test this idea, we attempted to increase the efficiency of

mutating chromosomal *oriC* by developing methods based on the recombineering (homologous recombination-mediated genetic engineering) technique (Court et al., 2002), shown schematically in Figure 8. In recombineering, cells expressing the lambda RED system are transformed with either a single-stranded or double-stranded DNA fragment containing the desired mutation, which is inserted into the chromosome by homologous recombination (Court et al., 2002).

There are several obstacles facing any method for replacing *oriC*. One is the lack of a selectable phenotype for recombinants. Other groups have solved the selection problem by placing a drug-resistance determinant on the donor DNA (Weigel et al., 2001), but for *oriC* replacement this solution is not ideal, since transcription near the *oriC* region might affect *oriC* function (Baker and Kornberg, 1988; Skarstad et al., 1990). Another challenge for *oriC* mutagenesis is that, because *oriC* function is essential for cell viability, one must be able to distinguish between a failure to recover a mutation because it knocks out function, or because the mutation efficiency is too low.

To overcome these challenges, we have developed two recombineering methods to generate both functional and non-functional *oriC* mutations on the chromosome. For functional *oriC* versions, we generated a recombineering strain (ACL400) in which chromosomal *oriC* was inactivated by insertion of a plasmid R1 derivative that confers resistance to kanamycin (Koppes and Nordström, 1986),

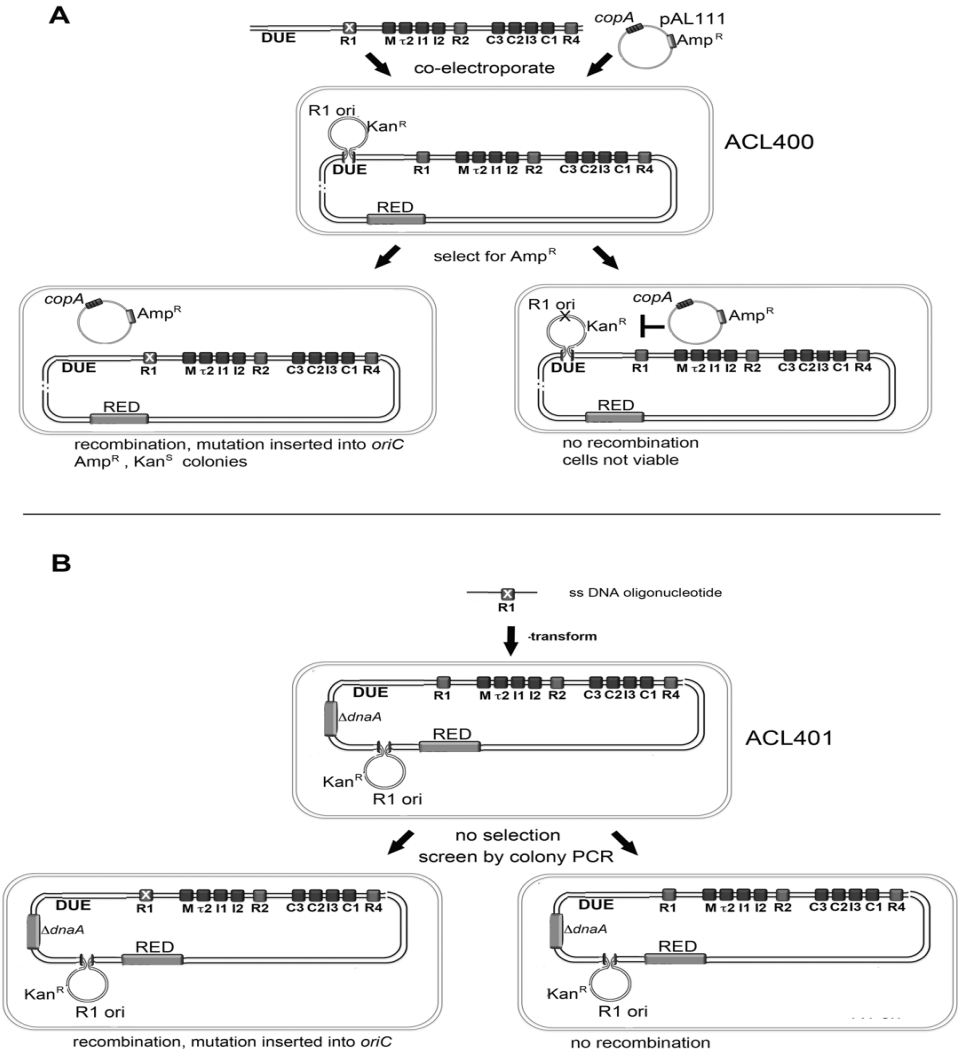

Figure 8. Scheme of the *oriC*-specific recombineering method. A) A PCR fragment carrying the desired mutation in *oriC* (marked by white X in R1) was co-electroporated with pAL111 to transform ACL400 cells in which the RED system was induced. Recombination results in replacement of the chromosomal origin (containing the plasmid R1 origin inserted into the DUE) with the mutated *oriC*. If the mutated *oriC* is functional, loss of the plasmid R1 origin allows recombineered cells to survive in the presence of *copA*, and makes them sensitive to kanamycin. pAL111 confers ampicillin resistance. B) A single-stranded DNA oligonucleotide carrying the desired mutation was electroporated into ACL401 cells in which the RED system was induced. All cells from the electroporation are plated on non-selective media and screening for the mutation is done using colony PCR. Cells are viable regardless of *oriC* function because this strain lacks DnaA and chromosome replication can initiate only from the integrated plasmid origin.

into the DUE (Figure 8A). In these cells, chromosomal DNA replication is initiated by the plasmid origin. Transformation of ACL400 with a plasmid (pAL111) expressing *copA*, an antisense RNA that inhibits synthesis of the plasmid initiator protein RepA, prevents initiation of chromosome replication (Molin and Nordström, 1980). Thus, co-transformation with a DNA fragment carrying an *oriC* mutation and a plasmid expressing *copA* and ampicillin resistance will give ampicillin-resistant colonies only when recombination resurrects a functional *oriC* copy (Figure 8A). Loss of the plasmid R1 origin also confers sensitivity to kanamycin.

Replacement of *oriC* in ACL400 is done in a single step, with a strong selection for recombinants. However, because lethal mutations cannot be recovered by this method, we also developed an alternative recombineering strain (ACL401) in which a *dnaA* deletion renders the wild type copy of *oriC* inactive, with initiation of chromosome replication driven by a plasmid R1 origin integrated into the chromosome outside of the *oriC* region (Figure 8B). Any *oriC* mutation, regardless of activity, can be made in ACL401, but there is no simple selection for mutants. Fortunately, recombination efficiencies are high enough (1–5%) to enable identification of mutants without selection, by screening colonies by PCR, using mutation-specific and wild-type specific primers. After outgrowth and verification of the mutation (see below), the function of mutated origins in ACL401 can be evaluated by transforming cells with a plasmid (pAL110) expressing both DnaA (to

activate *oriC*) and *copA* (to inactivate the plasmid R1 origin). Only cells harboring a functional *oriC* will yield colonies after this transformation.

To generate non-binding mutations in R1, R2, or R4, base pair alterations were first made on an *oriC* plasmid that also carries the pBR322 origin, which allows replication even when a mutation knocks out *oriC* function (Weigel et al., 1997). Each site was converted to 5′-GATATAGTT-3′, the same non-binding mutation tested in previous studies by the Messer lab (Langer et al., 1996; Weigel et al., 1997; 2001). The mutations on the plasmids were verified by DMS sequence analysis.

To confirm that the mutation knocked out both high and low-affinity binding, end-labelled double-stranded oligonucleotides containing the mutated site adjacent to the R4 box were incubated with purified DnaA, and the complexes separated by gel electrophoresis (Figure 9A). Only one complex (formed by DnaA binding to the R4 site) was seen. In contrast, DnaA formed two complexes on probes containing the low-affinity R5M site and R4 (Figure 9B). Probes containing R5M and the mutated site did not bind any DnaA (Figure 9C), consistent with previous studies which demonstrated that cooperative binding is required for DnaA occupation of low-affinity sites (Schaper and Messer, 1995).

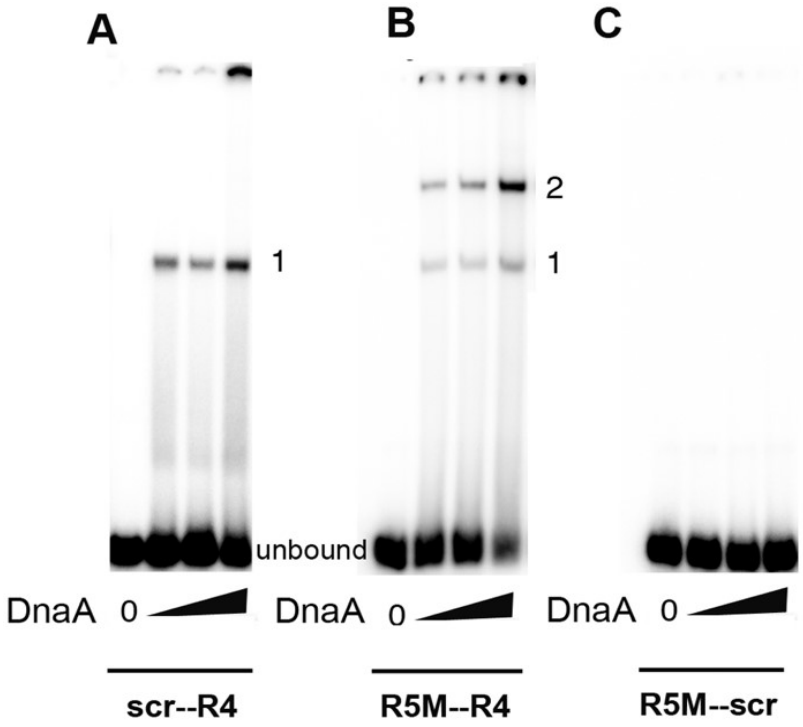

Figure 9. The mutated R-box does not bind DnaA. Double-stranded DNA oligonucleotides containing R4 and the non-binding GATATAGTT site (A), R4 and the low-affinity binding site R5M (B), or R5M adjacent to the non-binding site (C), were incubated with DnaA-ATP at DnaA/DNA molar ratios of 0:1, 2:1, 10:1 and 20:1, and the resulting complexes were resolved on polyacrylamide gels by electrophoresis. All DnaA sites were in the same binding orientation. The position of the unbound probe, and complexes resulting from 1 or 2 molecules of DnaA bound to the probe are marked.

For recombineering, PCR fragments carrying the *oriC* region from *gidA* to *mioC* were generated from the mutant plasmids, as well as from the wild-type *oriC* plasmid, and co-transformed with pAL111 into ACL400. ACL400 was used for the first attempts to generate non-binding mutations in R1 because it allows simple selection (ampicillin resistance) for recombinants.

While the transformation efficiency varied somewhat among experiments, *oriC* mutants with defective R1, R2 or R4 replaced the integrated plasmid R1 origin with similar frequencies as wild-type *oriC* (Table 7). Seventy percent to 80% of the primary colonies from all transformations were kanamycin sensitive, indicating loss of the integrated plasmid origin. While we did not examine the kanamycin-resistant cells in detail, we speculate that these colonies may have retained the integrated plasmid origin but had insufficient *copA* expression to render it inactive. Typically, 50–100 of the kanamycin-sensitive colonies were tested by PCR analysis, using mutant-specific and wild-type-specific primers paired with a primer in *gidA*, and approximately 70% of these contained the mutated site on the correct sized fragment (712 bp for R1, 801 bp for R2, and 872 bp for R4). The remaining 30% gave no signal with the mutant-specific primer, and strong signal with the wild-type primers, suggesting that although recombineering removed the integrated plasmid R1 origin, the recombination site was located between the DUE and the position of the desired mutation. Colonies that tested positive for the mutation were passaged and re-tested to ensure that the

origin region could be amplified only with mutant-specific primers, with no fragment generated when the wild-type primer was used. Primers flanking *oriC* (in *gidA* and *mioC*) were also used to amplify 1432 bp encompassing the *oriC* region from five different isolates, and the entire region was sequenced.

All five isolates carried the correct mutation. The *oriC* mutations were then successfully transduced into an *asnAB* MG1655 derivative (JEG22) to place them into cells with a clean genetic background. The presence of a single *oriC* copy carrying the desired mutation in the transductants was verified by: 1) PCR analysis, using mutant-specific and wild-type-specific primers paired with a primer in *gidA*; 2) PCR amplification of the region from *gidA* to *mioC*, followed by sequence analysis; and 3) qPCR, using primers within *oriC* or ter (Johnsen et al., 2011) to verify that the *oriC*/ter ratio was 1:1 in cells from stationary cultures. No evidence of extra copies of *oriC* was detected in any of the transductants, nor did we find any cells containing an *oriC* with both a wild-type and mutated site.

Table 7. Screening for the replacement of R1, R2 and R4 with the non-binding sequence: 5′-GATATAGTT-3′ (Langer et al., 1996; Weigel et al., 1997; 2001) in the chromosomal copy of *oriC*. Correct fragment size was determined by colony PCR using a forward primer in *gidA* and mutation specific reverse primers at R1, R2 and R4. Double DnaA binding site mutations were screened for using a forward primer in *gidA* and reverse primer in *mioC*.

Site Mutated	**Primary amp^R Colonies**	**Percent kan^S Colonies**	**Correct size PCR Fragment**	**Correct *oriC* Sequence**
Wild-Type	230	82%	100%	N/D
R1	140	75%	70%	5/5
R2	222	80%	72%	5/5
R4	153	82%	72%	5/5
R1/ R2	67	50%	15%	0/5
R1/ R4	66	79%	35%	0/5
R2/ R4	120	79%	30%	0/5
copA only	41	0	N/D	N/D

These results strongly suggest that no single high-affinity DnaA binding site is essential for *oriC* function. Because this result is inconsistent with previous findings (Weigel et al., 2001), and with models in which R1 has an essential role in unwinding and in helicase loading (Speck and Messer, 2001; Soultanas, 2012), we also used short (75–85 bp) single-stranded DNA oligomers to generate the R1, R2 and R4 mutations in the ACL401 recombineering strain with the inactive *oriC*. In this case, the *oriC* independent replication removed any selective pressure for *oriC* function, and the use of shorter, single-stranded oligonucleotides eliminated any chance of recombining multiple copies of the entire origin into the cells. Approximately 200 hundred colonies from recombineered cells were screened using mutant-specific and wild-type-specific primers as described above, and colonies testing positive for the presence of the mutation were detected with similar frequencies (2–5%) for all three mutations. After passaging and rescreening the positive colonies for the presence of the mutated site and absence of the wild-type site, the origin regions (*gidA-mioC*) of the mutated cells were sequenced. All of the sequenced origins verified that cells carrying each of the three mutations were successfully isolated. Transforming the mutant cells with pAL110 (expressing *dnaA* and *copA*, as described above) tested function of the mutated *oriC*. For all three (R1, R2 and R4) mutations, colonies were obtained in this transformation, with an efficiency similar to that seen with cells containing a wild-type *oriC* (approximately 2.5×10^4 colonies per μg of DNA). A plasmid expressing DnaA but not *copA* also transformed ACL401 cells with similar efficiencies in wild-type

and mutant strains, and transformation with a plasmid that expressed only *copA* gave no colonies in any ACL401 derivative.

MUTATIONS DISRUPTING THE BACTERIAL ORC PERTURB INITIATION TIMING

Previous studies showed that, although viable, cells carrying non-binding mutation in R4 dramatically under-initiate chromosome replication, while the same non-binding mutation in the R2 site caused less perturbation of initiation timing (Weigel et al., 2001; Riber et al., 2009). To verify this, and to examine cells lacking a functional R1 site, we used flow cytometry to evaluate initiation timing in wild-type and mutant cells. Under rapid growth conditions, the *E. coli* generation time is less than the time required to complete chromosome replication (Cooper and Helmstetter, 1968), so exponentially growing cells normally contain more than one copy of *oriC*, which synchronously trigger chromosome replication once per cell cycle on partially duplicated chromosomes (Skarstad et al., 1986). Because all origins in the cell fire simultaneously, the origin number doubles at the time of initiation. To examine the number of replication origins, cultures are treated with rifampicin to inhibit new rounds of chromosome replication and cephalexin, which prevents cell division (Skarstad et al., 1986). After adding the two drugs and incubating long enough to allow completion of ongoing rounds of replication, cellular DNA is stained using a fluorescent dye, and the number of chromosome

equivalents in the cell, detected by flow cytometry, reflects the number of *oriC* copies present at the time of drug addition.

JEG22 cells (wild-type *oriC*), growing exponentially in minimal media supplemented with glucose and casamino acids, contain 2 (pre-initiation) or 4 (post initiation) copies of *oriC* (Figure 10A). Since there are only 2n chromosome equivalents, both copies of *oriC* must have fired synchronously in all cells, a key feature of normal initiation timing. In contrast, both R1 and R4 mutants, despite growing with the same doubling time as wild-type cells (τ = 31 minutes), had perturbed initiation timing, including asynchronous initiations (Figure 10B and D), with cells containing 1, 2, 3 and 4 origins. The average number of origins per cell, and origins per cell mass, were also decreased (

Table 8) indicating that the mutant cells had under-initiated. The results for the R4 mutant are essentially the same as were previously reported (Weigel et al., 2001; Riber et al., 2009). Initiation timing in the R1 mutant was slightly less perturbed than in the R4 mutant. Several separate transductants of the R1 mutant, with origins generated by the two methods described above, were tested, and each showed a similar pattern in the flow cytometry experiments (not shown). The results support a model in which both R1 and R4 sites play key roles in pre-RC assembly, which is consistent with their acting as primary nucleation sites for DnaA oligomers (Rozgaja et al., 2011).

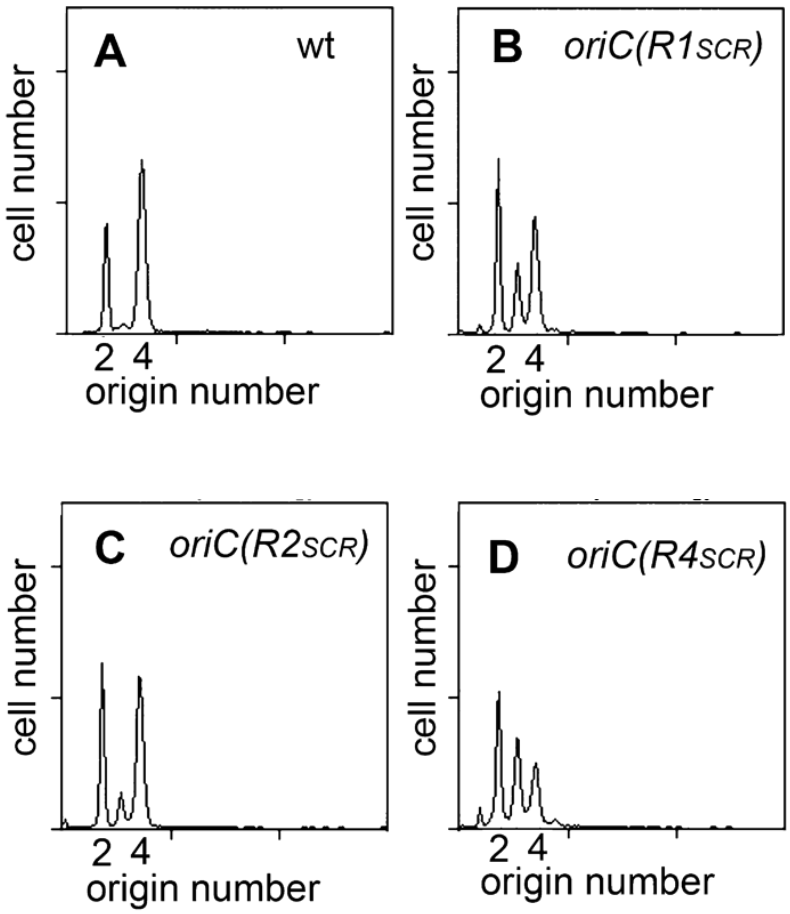

Figure 10. DNA histograms of JEG22 cells with R1 (B), R2 (C) and R4 (D) DnaA binding sites changed to the non-binding sequence: 5′-GATATAGTT-3′ (Langer et al., 1996; Weigel et al., 1997; 2001). Cells were treated with cephalexin and rifampicin and analyzed by flow cytometry to determine the number of origins per cell.

FUNCTIONAL ORCS IN RAPIDLY GROWING CELLS MUST CONTAIN AT LEAST TWO HIGH-AFFINITY SITES

Since each high-affinity DnaA binding site in *oriC* is capable of nucleating a DnaA oligomer associated with weak arrayed recognition sites (Miller *et al*., 2009), it seemed possible that *oriC* might function with only one strong site, particularly if only one oligomer was required to form a functional pre-RC. However, if DnaA oligomers are unstable without a high-affinity anchor site, or if their length is limited, then it is likely that two high-affinity sites would be required for *oriC* function. To evaluate these two possible scenarios, the recombineering procedures described above were used in an attempt to introduce *oriC*s carrying mutations in two high-affinity DnaA binding sites (R1/R2, R1/R4 and R2/R4) into the chromosome.

When PCR fragments (from *gidA* to *mioC*) carrying double high affinity site mutations were co-transformed with pAL111 (*copA* plasmid) into ACL400 cells, the number of ampicillin-resistant colonies recovered was twofold to threefold lower compared to wild-type *oriC* (Table 7). While most colonies were sensitive to kanamycin, screening using mutant-specific or wildtype specific primers yielded mixed signals. Sequence analysis showed all cells had wildtype DUE but only one or none of the desired high affinity site mutations in *oriC*. These results suggest an *oriC* deficient in binding at two high affinity sites is nonfunctional under these growth conditions.

Table 8. Doubling times for JEG22 cells with non-binding mutations at R1, R2 and R4 were determined by measured change in optical density (OD600) over time. The number of origins per cell and origins per cell mass were measured by flow cytometry. Relative cell mass was calculated by measuring light scatter of mutant strains compared to the wild type strain.

Site Mutated	Doubling Time (min.)	Number of Origins $Cell^{-1}$	Relative Cell Mass	Origins $Cell\ Mass^{-1}$
Wild-Type	31	3.68	1.0	3.68
R1	30	3.24	1.1	2.95
R2	31	3.56	1.0	3.56
R4	31	3.32	1.2	2.76

To verify this result, and to ensure we did not miss a rare event, the double mutant PCR fragments were transformed into ACL401, which allowed for recombineering independent of chromosomal *oriC* function. Using this method, cells carrying R1/R4, R1/R2 and R2/R4 mutations on the chromosome were recovered with roughly a 2% efficiency. *OriC* function was evaluated by transforming each mutant strain with pAL110 (expressing *dnaA* and *copA*). No colonies were observed. The transformation efficiency was < 0.1% compared to cells carrying a wild-type *oriC*. Taken together, the inability to make double high affinity site mutants in ACL400 and the ACL401 results, these data clearly show an origin with a single functional high affinity site is unable to form a viable pre-RC under fast growth conditions. Under slow growth conditions the rules for pre-RC formation are likely different as a previous study found the entire right half of *oriC* could be deleted, leaving only R1 and the left array of low affinity sites, however, these cells were highly stressed and could not support multi-fork growth (Stepankiw et al., 2009).

HIGH-AFFINITY SITES ARE REQUIRED TO CONSTRAIN *ORIC* DNA

In vivo, the transition from the bacterial ORC to the pre-RC, is switch-like, with full assembly of the DnaA oligomers delayed until just before the time of initiation (Cassler et al., 1995; Nievera et al., 2006). The switch from ORC to pre-RC results largely from the ability of Fis to inhibit DnaA and IHF from binding to the left region of *oriC* (Ryan et al., 2004) until sufficient DnaA-ATP accumulates

to displace Fis and relieve the repression. How Fis accomplishes this inhibition is not clear, since Fis repression must operate over a distance of up to 90 bp. One possibility is that the *oriC* region is contained in a constrained topological domain, so that bending in one region would make additional bends in the domain energetically unfavorable. The decreased flexibility could affect both IHF binding and nucleation of a DnaA oligomer from R1. If this scenario is true, then the question of how the domain is constrained remains. The DnaA molecules bound to the high-affinity sites are logical candidates for the constraining function, since these are the only other sites observed to be occupied by proteins in the bacterial ORC (Nievera et al., 2006). To test this idea, we first looked for evidence that DnaA binding to high-affinity sites could constrain *oriC*, by developing an assay that exploited the inherent ability of the DUE on supercoiled *oriC* plasmids to unwind in the absence of replication proteins (Kowalski and Eddy, 1989) under particular buffer conditions. Supercoiled *oriC* plasmids were cleaved with Mung Bean Nuclease (MBN), which is specific for single-stranded DNA, and primer extension revealed that the MBN cut sites were predominantly in the 13-mer region (Figure 11). As previously reported, incubated of supercoiled plasmids with DnaA, at levels that result in occupation of high- affinity sites, prior to MBN cutting caused loss of MBN sensitivity in the 13mer region (Kowalski and Eddy, 1989)(Figure 11, first panel), indicating that DnaA repressed unwinding in the 13mer region. It is likely that the loss of unwinding was due to DnaA binding to only high-affinity sites since, under the buffer conditions optimal for MBN

cleavage, binding to low-affinity sites is not observed when checked by DMS footprinting (data not shown). As a control, a mock incubation with buffer lacking DnaA was added to supercoiled plasmid and did not change the MBN sensitivity pattern (data not shown).

To determine if high-affinity sites acted cooperatively to constrain *oriC*, DnaA was incubated with plasmids lacking a functional R1, R2, or R4 site, and then treated with MBN (Figure 11, second-fourth panels). Interestingly, DnaA binding did not inhibit MBN-induced cutting in the 13mer region in any of these plasmids. These data are consistent with the idea that cooperative DnaA binding among high-affinity sites in *oriC* topologically constrains the origin.

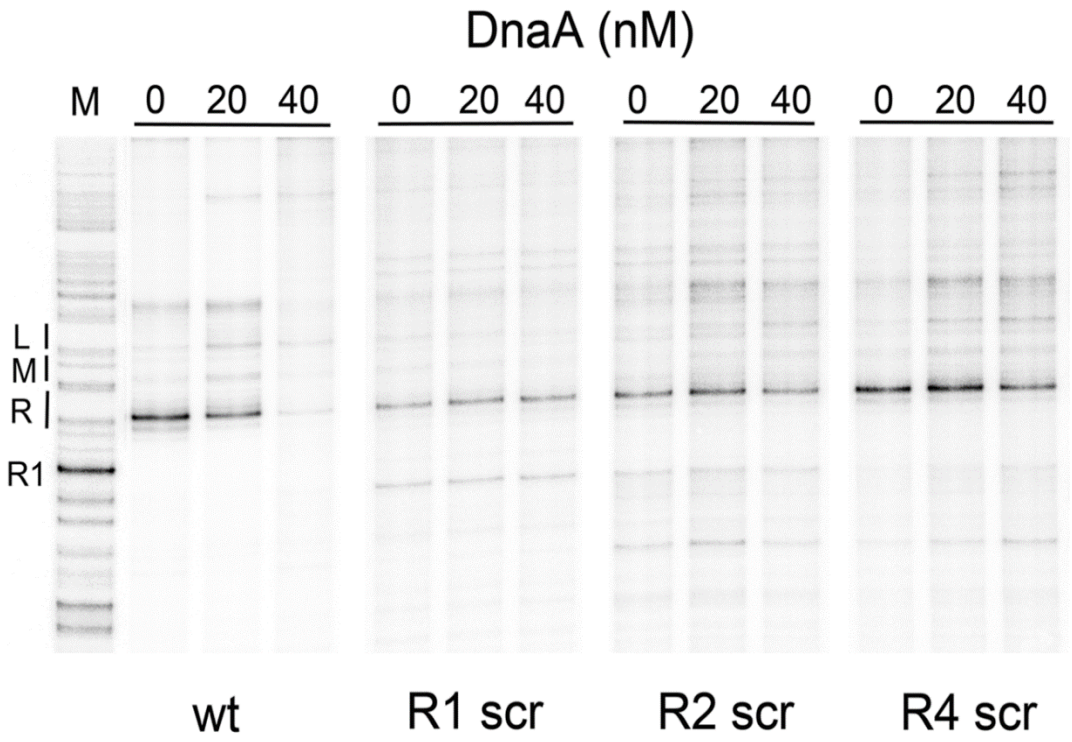

Figure 11. DnaA binding to high-affinity sites constrains *oriC* DNA and represses spontaneous unwinding in the DUE. Supercoiled wt or mutant *oriC* plasmids were incubated with the indicated concentrations of DnaA, and then treated with Mung Bean Nuclease as described in the Methods. Cut site location was determined by primer extension. Position of the right (R), middle (M) and left (L) 13-mer repeats of the DUE and the DnaA binding sites R1 are indicated next to the marker lane (M).

COOPERATIVE BINDING BETWEEN LOW AFFINITY SITES IS DEPENDENT ON SITE ORIENTATION AND SPACING

Low affinity sites in the arrays are in the same orientation and separated by 2 bp. If this arrangement is required for cooperative DnaA binding, then inverting arrayed sites, or placing sites further apart should prevent the arrays from becoming filled, and thus prevent pre-RC assembly. We first tested whether site orientation played a role in extension of DnaA oligomers within an array (Figure 12). Oligonucleotide probes, containing a peripheral donor strong site and two weak sites, in varying orientations, were incubated with increasing concentrations of DnaA, and the complexes that formed were examined using EMSA. When the two weak sites were in the same orientation, three complexes were formed, by binding to the donor site (complex 1), as well as to the donor site plus the proximal (middle) site (complex 2) or both weak sites (complex 3) (Figure 12A, D). When the two weak sites were in opposite orientations, only two complexes could form, indicating that only one weak site could be occupied in this configuration (Figure 12B, C). Given that a low affinity site must be paired with a strong site to bind DnaA, it is likely that it is the middle weak site that is bound in complex 2. Extension of DnaA from the high affinity anchor site was less dependent on site orientation, since the strong site assisted loading of the middle weak site on all probes, regardless of the orientation of the anchor (Figure 12C, D), although DnaA binding was most efficient when all three sites were in the same orientation.

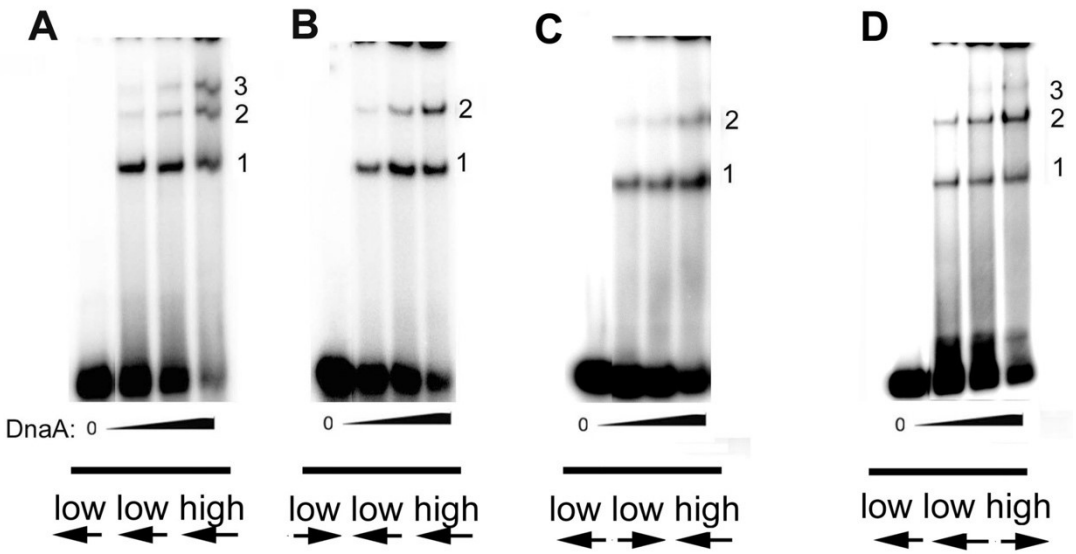

Figure 12. Cooperative binding between low affinity sites is dependent on site orientation, while strong sites in either orientation can donate DnaA. The EMSA system described in the text was used to examine DnaA-ATP binding to double-stranded DNA oligonucleotide probes containing R4 flanking two low affinity sites (C1 and R5m) in varying orientations, shown by arrows below each panel (A–D). End-labeled probes were incubated with DnaA-ATP at DnaA/DNA molar ratios of 0:1, 10:1, 20:1, and 50:1, and the resulting complexes (1, 2 and 3) were resolved on polyacrylamide gels.

To determine if the 2 bp spacing between arrayed sites is also critical for cooperative binding between low affinity sites, we designed probes for EMSA that had a donor strong site and two low affinity sites, with varying number of bases (1–4) between the weak sites (Figure 13). When DnaA was incubated with these probes, a clear dependency on site spacing was observed. Extension of DnaA from the middle to the distal weak site was optimal when sites were separated by 2 bp, shown by the formation of complex 3 in the gel shift assay shown in Figure 13B. Adding or removing an additional bp had a strong effect on site occupation (Figure 13A, D). Spacing of 3bp decreased binding to the distal weak site by approximately twofold (Figure 13C) and spacing of 1 bp or 4 bp eliminated occupation of the downstream weak site. These experiments reveal a stringent base pair spacing requirement between arrayed low affinity recognition sites for optimal occupation and raises the possibility that a specific structural interaction takes place between adjacent DnaA molecules which can be disturbed by positioning the molecules too closely or too distantly.

STRONG SITES CAN ASSIST LOADING OF A PROXIMAL WEAK SITE ONLY SEPARATED BY SEVERAL BASES

The experiments above show cooperative occupation of arrayed low affinity sites is limited to spacing of 2 to 3 bp. The arrangement of DnaA binding sites in *oriC* suggest the rules for high to low affinity site interactions may be less stringent. To determine the parameters governing these interactions,

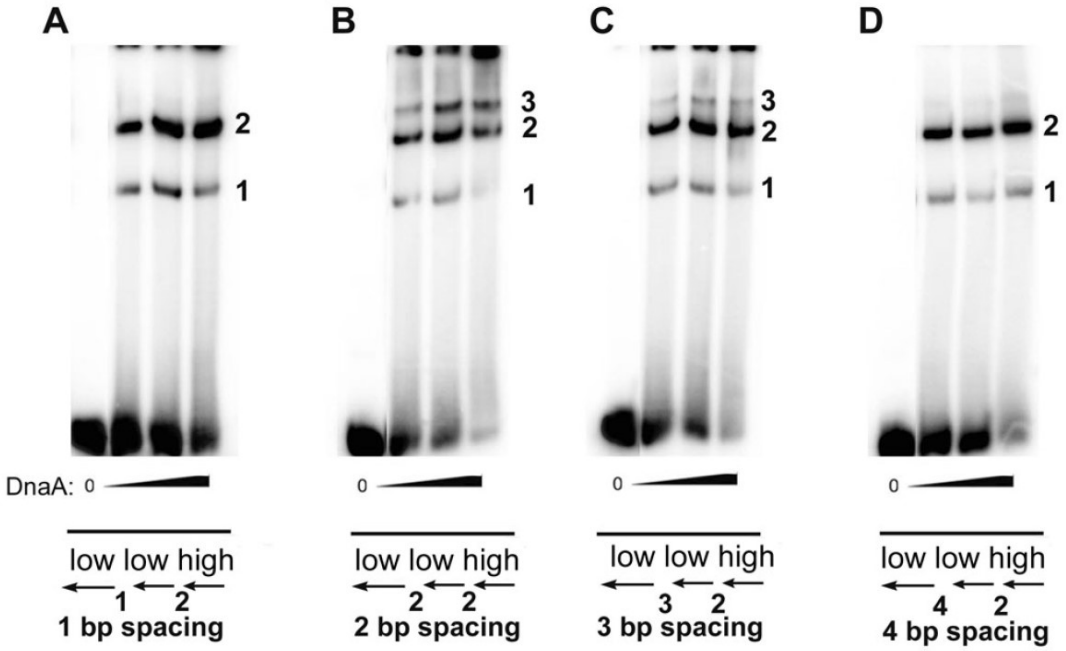

Figure 13. Extension of DnaA requires precise spacing between low affinity sites. Double-stranded oligonucleotide probes containing R4, C1, and R5m were designed so that C1 was separated from R5m by 1bp (A), 2bp (B), 3 bp (C), or 4 bp (D). All sites were in the same orientation. The configuration of the probes is indicated below each panel. End-labeled probes were incubated with DnaA-ATP at DnaA/DNA molar ratios of 0:1, 10:1, 20:1, and 50:1, and the resulting complexes (1, 2, and 3) were resolved on polyacrylamide gels.

oligonucleotide probes containing a high affinity site and single low affinity site, separated by varying numbers of base pairs (0-15), were incubated with increasing concentrations of DnaA and the complexes formed were examined using EMSA (Figure 14 & Figure 15). When DnaA was incubated with these probes, a clear dependency on site spacing was observed. Cooperative occupation of the low affinity site was best when 2 bp separated the nucleator from the weak site, evident by the greatest relative intensity of the second shift complex (Figure 14 B & Figure 15 B). As spacing increased, cooperative occupation of the weak site diminished. When 7 to 12 bp separated the sites, a second shift was not observed (Figure 14 E, data not shown). When 13 bp separated the sites, a second shift was seen, however, only at the highest protein concentrations and only faintly, suggesting the complex was unstable (Figure 15 D). Beyond 13 bp spacing, second shifts were essentially absent, with only the slightest hint of a band at the highest DnaA levels (Figure 15 E-F). One possible explanation for the recurrence in occupation of the weak site over 13 bp is that this spacing configuration places the binding sites in helical phase, puts the two DnaA monomers on the same face of the DNA.

When the low affinity site was positioned one base pair from the nucleation site, the weak site could not be bound, evident by the lack of a second shift (Figure 14 A). This result is consistent with the observation above between two low affinity sites separated by 1 bp (Figure 13 A). Likely, there is a specific structural interaction that takes place between adjacent DnaA molecules which is disturbed

when binding sites are positioned too closely. When the binding sites were adjacent (0 bp spacing), the EMSA pattern was highly unusual with two faint shifts of equivalent intensity (Figure 15A). These shifts appeared only at the highest DnaA concentrations and there was no dose-dependent change in band intensity. It is difficult to predict how two DnaA molecules might interact under these conditions to produce the super-shifted complex. Collectively, these data show the rules governing high to low affinity site interactions are less stringent that those regulating interactions between low sites.

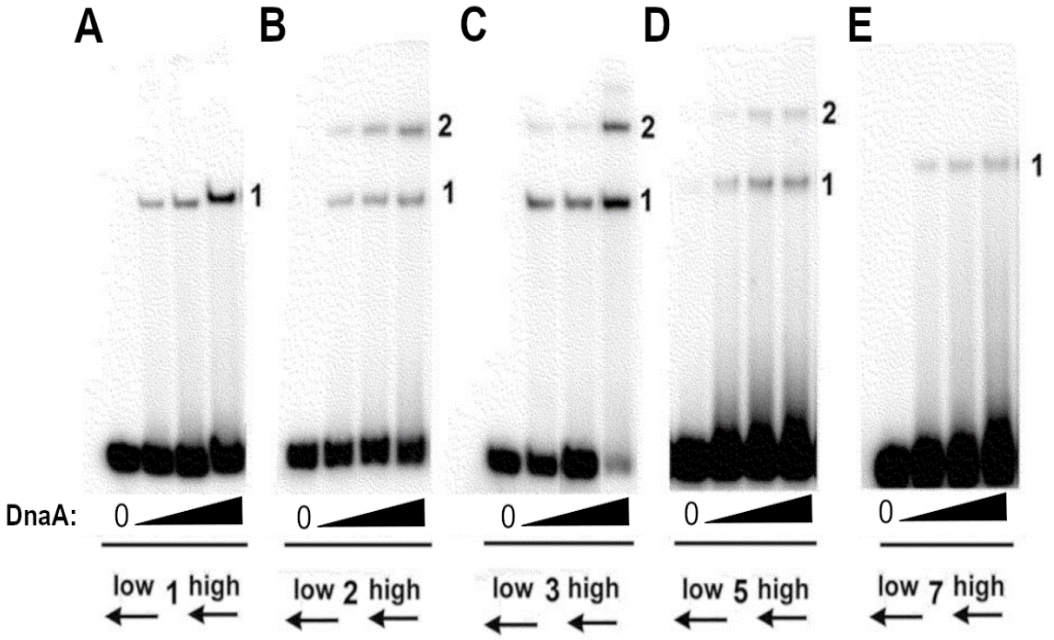

Figure 14. Extension of DnaA requires precise spacing between high and low affinity sites. Double-stranded oligonucleotide probes containing R4 and R5m were designed so that the sites were separated by 1bp (A), 2bp (B), 3 bp (C), 5 bp (D), or 7 bp (E). All sites were in the same orientation. The configuration of the probes is indicated below each panel. End-labeled probes were incubated with DnaA-ATP at DnaA/DNA molar ratios of 0:1, 10:1, 20:1, and 50:1, and the resulting complexes (1, 2, and 3) were resolved on polyacrylamide gels.

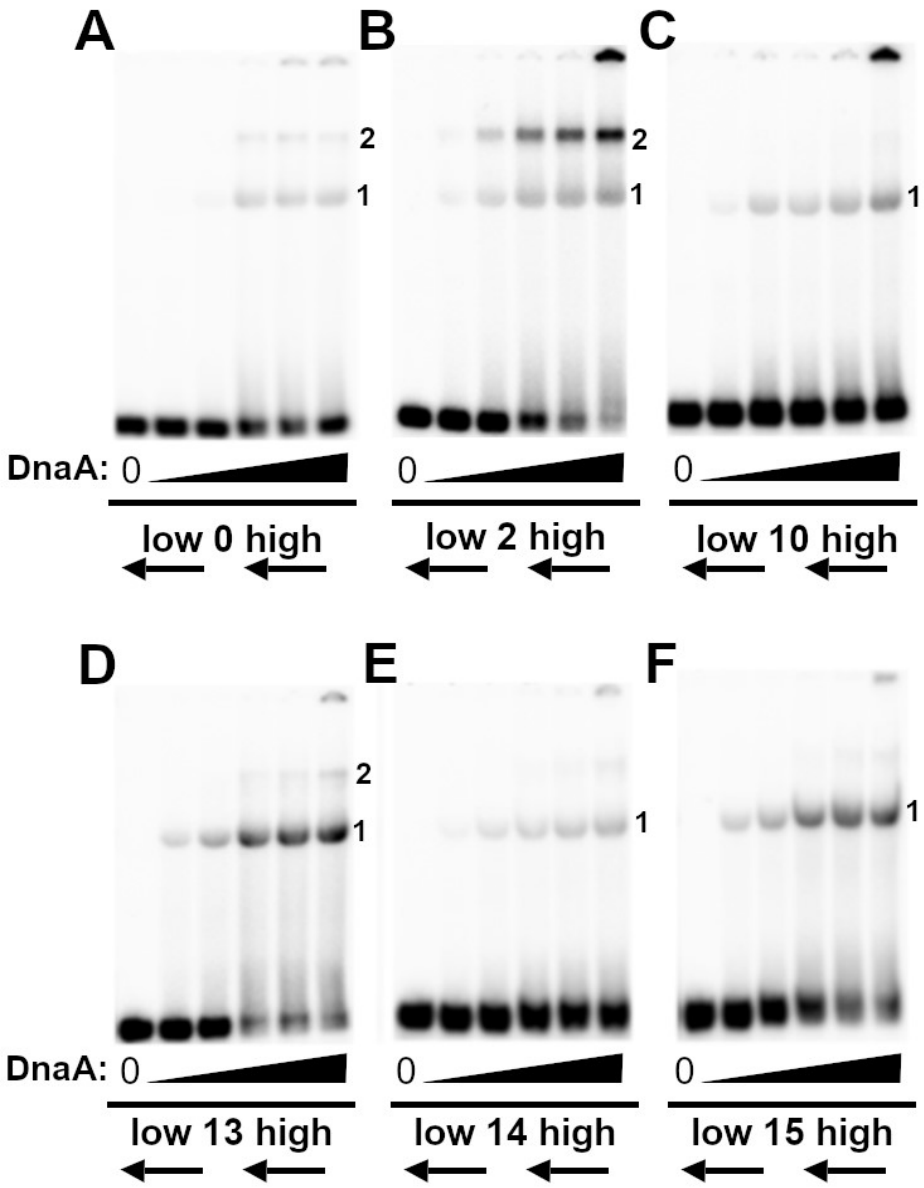

Figure 15. Extension of DnaA between high and low affinity sites is reduced by spacing less than 2 bp or greater than 7 bp. Double-stranded oligonucleotide probes containing R4 and R5M were designed so that the sites were separated by 0bp (A), 2bp (B), 10bp (C), 13bp (D), 14bp (E), or 15bp (F). All sites were in the same orientation. The configuration of the probes is indicated below each panel. End-labeled probes were incubated with DnaA-ATP at DnaA/DNA molar ratios of 0:1, 1:1, 2:1, 5:1, 10:1, and 20:1 and the resulting complexes (1 and 2) were resolved on polyacrylamide gels.

DISCUSSION

DnaA oligomerization along *oriC* DNA plays an important role in initiating new rounds of chromosomal DNA replication in bacteria (Erzberger *et al*., 2006; Rozgaja *et al*., 2011), and nucleation of these oligomers requires the high-affinity interaction of DnaA with recognition sites in *oriC* (Miller *et al*., 2009). Although DnaA is highly conserved among all bacteria, the number and arrangement of *oriC* high-affinity recognition sites for DnaA varies widely among bacterial types (Zawilak-Pawlik *et al*., 2005; Gao and Zhang, 2007), and it remains unclear why any particular configuration is preferred. Based on the results of our *oriC*-specific recombineering, we report that high-affinity DnaA–*oriC* interactions on the *E. coli* chromosome play two important roles: (1) as nucleation (start) sites for DnaA oligomerization, and, unexpectedly, (2) in maintaining *oriC* in a constrained conformation that prohibit spontaneous DUE unwinding and promotes the correct topological transition of the complex as the pre-RC is assembled. This latter attribute adds an important new function for the bacterial ORC, and suggests that not only are there instructions for ordered DnaA oligomerization encoded into *oriC* DNA, but also instructions (in the form of specially positioned recognition sites) to ensure the correct spatial orientation of *oriC* during pre-RC assembly.

No individual high-affinity DnaA recognition site was essential for *oriC* function, and for the first time, viable strains with a defective R1 recognition site were constructed. This result was unexpected, since R1 non-binding mutants were

reported to be non-viable (Weigel *et al*., 2001), and deletion mutants of *oriC*, in which R1 is the only high-affinity site retaining, have unwinding and ssDNA binding activity *in vitro* (Speck and Messer, 2001; Ozaki and Katayama, 2012), and are viable under slow growth conditions (Stepankiw *et al*., 2009). Based on these reports, R1 was proposed to be essential, and required to nucleate DnaA oligomers extending into the DUE that stabilize unwound DNA and assist in the loading of DNA helicase (Richardson et al., 2016, Hsu *et al*., 1994; Speck and Messer, 2001; Soultanas, 2012). However, the results from this study indicate that these activities can be accomplished using other paths. For example, the right half of *oriC* has been reported to play a role in helicase loading (Ozaki and Katayama, 2012; Ozaki *et al*., 2012a), and this role might increase in the absence of R1. Further, the left DnaA oligomer, which can be nucleated by R2, combined with the sharp DNA bend in the IHF region, is sufficient for unwinding and ssDNA binding (Ozaki and Katayama, 2012; M. Vora, manuscript in preparation). It remains unclear why strains unable to bind DnaA at R1 were not isolated in other labs, but the increased efficiency of the origin-specific recombineering method used for the studies reported here may have facilitated their construction. In particular, the powerful and continuous selection (the *copA* expression plasmid) against chromosome replication in non-recombinant ACL 400 cells allows a significant proportion of survivors (> 50%) to carry the desired chromosomal mutation in a minimally altered strain.

In our model of pre-RC assembly, DnaA oligomers fill the gap regions between high-affinity sites (Leonard and Grimwade, 2010). In origins lacking a functional R1 or R4 site, R2 must act as an alternative, albeit less efficient, nucleation site for the left or right oligomer respectively. The gap-filling model also predicts that one or both of the DnaA oligomers are normally anchored by DnaA bound to R2, and the mildly perturbed initiation timing observed in cells with a mutated R2 site is consistent with this idea (Weigel et al., 2001, Riber et al., 2009). However, even though any high-affinity site has the capability of nucleating a DnaA oligomer, cells containing a single high-affinity site were not viable in our growth conditions, indicating that one unanchored oligomer is either not long enough or stable enough to provide origin activity. Based on this observation in *E. coli*, extending single DnaA oligomers over long regions of DNA may be incompatible with the function of any bacterial *oriC*, and large replication origins (for example, see Zawilak-Pawlik *et al*., 2005) may require multiple high-affinity interactions to nucleate and anchor DnaA oligomers that span the origin.

While the exact nature of the configuration in the bacterial ORC remains to be determined, we propose a model in which the DnaA molecules occupying high-affinity sites interact to form a constrained loop/toroid (Figure 16A). The ability of DnaA to bend DNA (Schaper and Messer, 1995) may contribute to the DnaA/DnaA

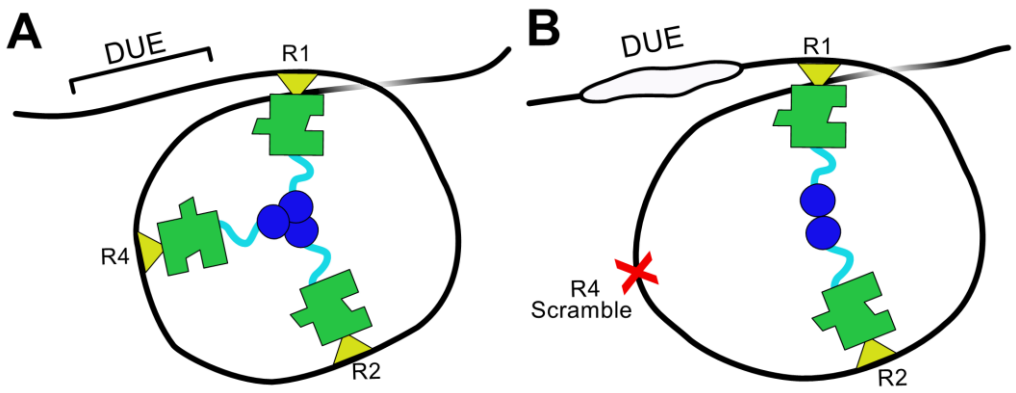

Figure 16. Model of the *E. coli* ORC. A) DnaA bound at the three high affinity sites R1, R2 and R4 suppress spontaneous unwinding of the DUE. Bound molecules are suggested to interact via their N-terminal domains (blue circles) to form a constrained loop. B) Loss of binding, at any high affinity sites, eliminates DnaA dependent constraint allowing spontaneous unwinding of the DUE. Model not to scale.

interactions, by placing the DnaA molecules in closer proximity. We speculate that the interactions are via Domain I, but this remains to be determined. These interactions would be a plausible target for the positive regulator of initiation DiaA, which has been shown to stabilize DnaA oligomers (Keyamura *et al*., 2009). This model of a loop wrapped around a core of DnaA is consistent with electron microscopy studies (Funnell *et al*., 1987; Crooke *et al*., 1993). When either R1 or R4 is missing, *oriC in vivo* activity is not lost and other *oriC* associated proteins such as Fis and IHF may help form a similar looped structure, but one which has lost constraint (Figure 16B). If this model is correct, then the structure that is formed in the bacterial ORC has similarity to that of a proto-nucleosome. In fact, a similarly phased arrangement of nucleosome positioning proteins with looped Green monkey cell α-satellite DNA was previously reported by Strauss and Varshavsky (1984). It appears that the earliest role of DnaA at *oriC* may be to generate the appropriate loop on which to assemble a nucleosome-like pre-RC (Funnell *et al*., 1987).

An appealing feature of this model is that it provides a mechanism governing the dynamic interplay among DnaA, Fis and IHF during pre-RC assembly (Ryan *et al*., 2004). Under rapid growth conditions, the constrained loop would also contain a bend induced by Fis binding. If this structure does not accommodate additional DNA bending in the region of the IHF binding site, then simultaneous interaction of Fis and IHF with *oriC* would be prohibited, as was

previously observed (Ryan *et al*., 2004). Also, because it is likely that bending in the IHF region is needed for R1 to nucleate the DnaA oligomer in the left half of *oriC* (Rozgaja *et al*., 2011), the Fis-induced bend would also repress assembly of the DnaA oligomer in the left region of *oriC* (Ryan *et al*., 2004). This model is supported by the finding that DnaA binding to high-affinity sites in wild type *oriC* caused a loss of supercoiled DNA-dependent unwinding in the DUE, but DnaA was not able to constrain supertwists when bound to origins lacking one of the high-affinity sites.

Further studies will be necessary to assess cross-strand DnaA interactions, but a regulatory system based on dynamic DNA looping or bending (Travers, 2006) would be particularly well suited to rapidly assemble pre-RC in fast growing bacteria. By transitioning through different constrained conformational states, subassemblies of the pre-RC could permit or inhibit access of *oriC* binding proteins, ensuring that each conformation was a prerequisite (and regulator) for the next assembly stage. For example, in *E. coli*, a loop (bound to Fis) initially made by interacting DnaA molecules at high-affinity sites would be remodeled when IHF-induced bends catalyzed interactions between DnaA molecules at R1 and R5M. However, the remodeling would be possible only when sufficient DnaA was available to displace Fis and fill the rest of the available low-affinity recognition sites.

If regulation by DNA looping is a common attribute of bacterial replication origins, the diversity in position and number of high-affinity sites among bacterial replication origins would then be determined not only by the number of anchored oligomers required to span the origin, but also by the need to produce different *oriC* configurations (in the form of dynamic DNA loops and bends) during pre-RC assembly that ensure correct initiation timing. It is reasonable to expect that the high affinity recognition sites in bacterial *oriC*s would play a role in both topological constraint and nucleation, as is the case for *E. coli*, but in other bacterial types these roles may be assigned separately to individual sites. This scenario may explain the reason some bacterial origins contain very high numbers of consensus DnaA boxes. Interestingly, the looped origin DNA conformation proposed here for *E. coli oriC* may also be a common feature of eukaryotic replication origins. For example, using cryo-electron microscopy, the path of origin DNA within yeast ORC was recently shown to be a DNA loop (Sun *et al*., 2012) that becomes available for cross-strand interactions promoted by Cdc6.

Historically, bacterial replication origins have been described as containing two types of nucleotide sequence information necessary for biological function: canonical DnaA recognition sites (R-boxes) to recruit initiator protein, and "spacer" regions of low sequence specificity, that were proposed to properly separate the R-box bound DnaA molecules (Oka et al., 1984; Zyskind and Smith, 1986; Woelker and Messer, 1993). The resulting origin maps, exemplified by *E. coli oriC* (Figure

5), provided few obvious clues regarding how the R-boxes and spacer regions work together to assemble pre-RC during the cell cycle. This work, combined with previously published studies (McGarry et al., 2004; Kawakami et al., 2005), demonstrates that regions with "spacer" function contain nucleotide sequence motifs that consist of four closely-spaced specific low affinity initiator recognition sites, arranged into oppositely-oriented symmetrical arrays in each half of the origin (Figure 3). The newly identified C sites, combined with the previously identified I sites (Grimwade et al., 2000), τ2 site (Kawakami et al., 2005), and R5M site (Matsui et al., 1985), yield 11 mapped DnaA contacts in *E. coli oriC*, with additional DnaA occupying the A-T rich DUE (see below). Thus, the revised map moves closer to accommodating the 15–20 DnaA monomers previously estimated by EM (Fuller et al., 1984) and *in vitro* unwinding assays to be required for complete pre-RC assembly (Ryan et al., 2004), and promotes development of a testable model for properly timed pre-RC assembly during the *E. coli* cell cycle.

Precise spacing and orientation of arrayed sites was required for proper pre-RC assembly, and work by our lab found disrupting a contact site at one end of the array perturbed further downstream contacts, indicating polarity in the sequential placement of DnaA. This polarity was determined by the arrangement of low affinity sites, rather than the orientation of the nucleating site. These data are consistent with a model of pre-RC assembly in which the arrayed sites, by precisely and closely positioning DnaA monomers, direct sequential growth of two DnaA

oligomers anchored by R1 and R4, both converging on R2. In this model, regulators of *oriC* activity would function by targeting oligomer extension. For example, the negative regulator SeqA, previously shown to prevent DnaA from re-binding to low affinity sites after initiation (Nievera et al., 2006) would prevent oligomer growth in both halves of *oriC*. Fis, which also inhibits DnaA binding to low affinity sites (Ryan et al., 2004), might block extension of the right filament (see Figure 6). The polarized formation of two oppositely oriented DnaA oligomers also suggests a mechanism for origin unwinding. Although the two oligomers appear to assemble independently, it is possible that they ultimately become joined into a contiguous filament. Since the two arrays in each *oriC* half are not in helical phase, connecting the two DnaA oligomers as they converge at R2 may require twisting of the DNA strand, which would impart torsional stress on the duplex DNA and aid strand separation in the DUE. Alternatively, preventing free rotation of the DNA helix by having it locked by multiple contacts within the DNA oligomer may be sufficient to focus topological stress within the A-T rich region.

Normal pre-RC assembly must also be coordinated with the growth rate of the cell, to ensure proper initiation timing during the cell cycle (Leonard and Grimwade, 2010). Based on spacing, one of the earliest DnaA-DnaA interactions in pre-RC assembly should be between DnaA monomers occupying R4 and the proximal C1 site, and therefore the right half of oriC might play a key role in proper timing of pre-RC assembly. Consistent with this idea, cells harboring R4

binding mutations or deletions on chromosomal *oriC* exhibit dramatic perturbations in initiation timing (Bates et al., 1995; Riber et al., 2009), and chromosomal *oriC* mutants lacking the entire R2–R4 region, while viable, show profound replication initiation and growth defects, and are particularly sensitive to rich media (Stepankiw et al., 2009). The DnaA binding pattern in the pre-RCs formed from these mutant origins is not currently known, but since initiation timing is not normal, it seems logical to suggest that pre-RC assembly is altered, and somehow uncoupled from cellular mass accumulation. Since the left half of *oriC* is reported to be required for viability (Stepankiw et al., 2009), it is possible that the oligomer assembled in this half plays a greater role in DNA unwinding and helicase loading.

The structure of the higher order DnaA complexes formed on the arrays remains to be determined. Each array contains sites that preferentially bind DnaA-ATP suggesting that at least some of the positioned DnaA monomers must be in the ATP-form. Domain III-specific interactions between adjacent DnaA-ATP molecules have been previously proposed to form a right-handed helical filament (Erzberger et al., 2006), but, as currently modeled, this structure is not capable of double-stranded DNA binding (Duderstadt et al., 2010). While the modeled DnaA-ATP filament is able to interact with single-stranded DNA in an unwound DUE, the closely spaced DnaA-DNA interactions within the arrays are likely to promote assembly of a novel, alternative DnaA filament, in which the protomers are less compact and capable of stable interaction with double-stranded *oriC* DNA.

Additionally, all the DnaA in the alternative filament structure may not be in the ATP form, since both *in vitro* and *in vivo* studies have shown that some DnaA-ADP is permitted in the pre-RC (Yung et al., 1990; Grimwade et al., 2007). We are currently evaluating the requirement for DnaA-ATP and domain III interactions in array filling.

Based on the studies presented here, the binding of DnaA to neighboring recognition sites is affected by even small changes in spacing, although the interaction between high and low affinity DnaA recognition sites are less stringent than those for low affinity site interactions. This observation suggests that a different type of DnaA-DnaA contact may be used to extend DnaA from a high affinity binding site than the contacts between low affinity binders. We previously demonstrated that DnaA defective in domain I oligomerization was capable of binding strong sites, but could not occupy low affinity sites (Miller et al., 2009). Since domain I is attached to the flexible domain II linker region (Nozaki and Ogawa, 2008), by using domain I interactions, it should be possible to extend a DnaA monomer from a nucleating site to a weak site over increased distances if the linker length is extended, and we are investigating this possibility. Accessory factors, such as DiaA (Ishida et al., 2004), which associate with Domain I and increase initiation efficiency, may also promote or stabilize these longer-range DnaA-DnaA interactions. Additionally, the effective proximity between strong and weak sites could be modulated by the DNA bending proteins Fis and IHF, which

bind in the right and left half of *oriC*, respectively (Figure 4), and are known to regulate pre-RC assembly (Ryan et al., 2004). IHF would increase interaction of R1 with R5M, which are separated by 46 bp, and Fis might affect the ability of R2 to interact with C3 (separated by 20 bp). Further studies will be necessary to determine how removal of the Fis and IHF sites, as well as altering the distance between high and low affinity sites, affects pre-RC assembly and initiation timing.

The response to positioning recognition sites at distances smaller than 2 bp was unexpected and suggests that there may be features of DnaA that limit the ability to occupy the closely spaced sites. While these remain speculative, there are several possible explanations. First, Domain I-I interactions may act as a hinge so that DnaA molecules have difficulty interacting with one another in the absence of DNA. Second, a structural feature of DnaA domain III, such as the protruding arginine finger may produce interference between DnaA molecules trying to occupy closely spliced sites. Third, the bending of DNA when DnaA occupies a recognition site may alter the nearby site so that it no longer is accessible by a nearby partner. Further studies will be required to rule out these possibilities.

Is the information on *E. coli* pre-RC assembly relevant for other bacterial replication origins? DnaA is the initiator protein for all bacteria, but DnaA binding sites in bacterial replication origins have been identified primarily by sequence similarity to the consensus DnaA R box. Further, the arrangement of mapped DnaA recognition sites in other bacteria is very different from those in *E. coli*, for

example see Zawilak-Pawlik et al. (2005). Our studies reveal that extreme deviation from the consensus recognition site is tolerated for DnaA recognition, although deviations greatly reduce binding affinity. Since it appears that clustered low affinity contacts adjacent to high affinity DnaA recognition sites play a key role in *E. coli* pre-RC assembly, it will be necessary to re-evaluate DnaA binding to other origins to determine if these features are shared by all bacteria. Nevertheless, some rules for *oriC* geography are emerging as is the possibility that there are sequences that are required for origin function that are distinct from those required for proper regulation of function. While appreciating the diversity seen among different bacterial types, every *oriC* is expected to contain the following features: 1) a set of high affinity DnaA recognition sites that define the ORC, whose placement constrains the origin under low DnaA concentrations and whose spacing is based on the requirement to form the appropriate topological structure (loops) to accomplish this task. The spacing will be related to the overall length of the origin as well as the length of the flexible linker domain II of the DnaA that determines the limit of domain I-domain I interactions. 2) additional recognition sites for post-ORC DnaA interactions that focus torsional stress to the DUE. These can be high or low affinity sites, but their role is to bend the DNA and stabilize the DNA helix so that a specific topology is reached at the end of pre-RC assembly. Closely spaced and low affinity DnaA-ATP preferential recognition sites can be used for this purpose, analogous to the case for *E. coli* to couple pre-RC assembly to DnaA-ATP availability with the formation of tight DNA loops (closely spaced bends of

50 degrees) or helical filaments. Alternatively, high affinity interactions may substitute if DnaA availability is tightly regulated. 3) DnaA recognition sites that are components of regulatory mechanisms that utilize proteins other than DnaA. For example, DNA bending proteins that place two recognition sites in proximity or blocking proteins that prohibit DnaA interactions. Targets could be at any stage of pre-RC assembly and DnaA bound at a particular recognition sites might play dual roles. For example, a bound DnaA molecule might play an important role as a component of an ORC, but also be blocked from accessing *oriC* during the cell cycle by a regulatory protein.

LITERATURE CITED

Abe Y, Jo T, Matsuda Y, Matsunaga C, Katayama T, Ueda T. 2007. Structure and function of DnaA N-terminal domains: specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on *oriC*. J. Biol. Chem. 282:17816–17827.

Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A. 1999. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. J. Bacteriol. 181,6361–6370.

Aranovich A, Braier-Marcovitz S, Ansbacher E, Granek R, Parola AH, Fishov I. 2015. N-terminal-mediated oligomerization of DnaA drives the occupancy-dependent rejuvenation of the protein on the membrane. BioSci. Rep. Aug 13;35(5). pii: e00250. doi: 10.1042/BSR20150175.

Aranovich A, Gdalevsky GY, Cohen-Luria R, Fishov I, Parola AH. 2006. Membrane-catalyzed nucleotide exchange on DnaA: effect of surface molecular crowding. J. Biol Chem 281:12526–12534.

Atlung T, Clausen ES, Hansen FG. 1985. Autoregulation of the dnaA gene of *Escherichia coli* K12. Mol Gen Genet 200:442-50.

Atlung T, Lobner-Olesen A, Hansen FG. 1987. Overproduction of DnaA protein stimulates initiation of chromosome and minichromosome replication in *Escherichia coli*. Mol Gen Genet 206:51-9.

Baker TA, Bell SP. 1998. Polymerases and the Replisome: Machines within Machines. Cell 92, 295–305.

Baker TA, Kornberg A. 1988. Transcriptional activation of initiation of replication from the *E. coli* chromosomal origin: an RNA-DNA hybrid near *oriC*. Cell. Oct 7;55(1):113-23.

Ball CA, Osuna R, Ferguson KC, Johnson RC. 1992. Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. J Bacteriol 174:8043-56.

Bates DB, Asai T, Cao Y, Chambers MW, Cadwell GW, Boye E, Kogoma T. 1995. The DnaA box R4 in the minimal *oriC* is dispensable for initiation of *Escherichia coli* chromosome replication. Nucleic Acids Res 23:3119–3125.

Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277:1453-74.

Boeneman K, Crooke E. 2005. Chromosomal replication and the cell membrane. Current Opinion in Microbiology. Apr 8(2):143-8. PMID: 15802244.

Bramhill D, Kornberg A. 1988. Duplex opening by DnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. Cell 52:743-55.

Braun RE, O'Day K, Wright A. 1985. Autoregulation of the DNA replication gene *dnaA* in *E. coli* K-12. Cell 40:159-69.

Campbell JL, Kleckner N. 1990. *E. coli oriC* and the *dnaA* gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. Cell 62:967-79.

Cassler MR, Grimwade JE, Leonard AC. 1995. Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. The EMBO Journal 14:5833–5841.

Cooper S, Helmstetter CE. 1968. Chromosome replication and the division cycle of *Escherichia coli* B/r. J Mol Biol. Feb 14;31(3):519-40.

Court DL, Sawitzke J, Thomason LC. 2002. Genetic engineering using homologous recombination. Annu. Rev. Genet. 36:361-388.

Crooke E, Castuma CE, Kornberg A. 1992. The chromosome origin of *Escherichia coli* stabilizes DnaA protein during rejuvenation by phospholipids. J Biol Chem 267:16779–16782.

Crooke E, Thresher R, Hwang DS, Griffith J, Kornberg A. 1993. Replicatively active complexes of DnaA protein and the *Escherichia coli* chromosomal origin observed in the electron microscope. J Mol Biol. 233:16–24.

Diffley JFX. 2010. The many faces of redundancy in DNA replication control. Cold Spring Harb. Symp. Quant. Biol., 75, 135–142.

Diffley JFX. 2011. Quality control in the initiation of eukaryotic DNA replication. Philos Trans R Soc Lond B Biol Sci 366:3545–3553.

Dillon SC, Dorman CJ. 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat. Rev. Microbiol. 8:185–195.

Donczew R, Mielke T, Jaworski P, Zakrzewska-Czerwi´nska J, ZawilakPawlik A. 2014. Assembly of Helicobacter pylori initiation complex is determined by sequence-specific and topology-sensitive DnaA-*oriC* interactions. J. Mol. Biol. 426:2769–2782.

Duderstadt KE, Berger JM. 2008. AAA+ ATPases in the initiation of DNA replication. Crit Rev Biochem Mol Biol 43:163-87.

Duderstadt KE, Chuang K, Berger JM. 2011. DNA stretching by bacterial initiators promotes replication origin opening. Nature 478:209–213

Duderstadt KE, Mott ML, Crisona NJ, Chuang K, Yang H, Berger JM. 2010. Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator. J. Biol. Chem. 285:28229–28239.

Ellis HM, Yu D, DiTizio T, Court DL. 2001. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. Proc. Natl. Acad. Sci. 98(12):6742-6746.

Erzberger JP, Mott ML, Berger JM. 2006. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. Nat. Struct. Mol. Biol. 13:676–683.

Erzberger JP, Pirruccello MM, Berger JM. 2002. The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. EMBO J. Sep 16;21(18):4763-73.

Felczak MM, Kaguni JM. 2004. The box VII motif of *Escherichia coli* DnaA protein is required for DnaA oligomerization at the *E. coli* replication origin. J. Biol. Chem. 279:51156–51162.

Felczak MM, Simmons LA, Kaguni JM. 2005. An essential tryptophan of *Escherichia coli* DnaA protein functions in oligomerization at the *E. coli* replication origin. J Biol Chem. Jul 1;280(26):24627-33.

Fingland N, Flåtten I, Downey CD, Fossum-Raunehaug S, Skarstad K, Crooke E. 2012. Depletion of acidic phospholipids influences chromosomal replication in *Escherichia coli*. Microbiology open 1:450–466.

Finkel SE, Johnson RC. 1992. The Fis protein: it's not just for DNA inversion anymore. Mol Microbiol 6:3257-65.

Fujikawa N, Kurumizaka H, Nureki O, Terada T, Shirouzu M, Katayama T, Yokoyama S. 2003. Structural basis of replication origin recognition by the DnaA protein. Nucleic Acids Res 31:2077-86.

Fujimitsu K, Katayama T. 2004. Reactivation of DnaA by DNA sequence-specific nucleotide exchange in vitro. Biochem Biophys Res Commun. Sep 17;322(2):411-9.

Fujimitsu K, Senriuchi T, Katayama T. 2009. Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. Genes Dev. 23:1221–1233.

Fuller RS, Funnell BE, Kornberg A. 1984. The DnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites. Cell Oct;38(3):889-900.

Fuller RS, Kornberg A. 1983. Purified DnaA protein in initiation of replication at the *Escherichia coli* chromosomal origin of replication. Proc. Natl. Acad. Sci. U.S.A. 80:5817–5821.

Funnell BE, Baker TA, Kornberg A. 1987. In vitro assembly of a prepriming complex at the origin of the *Escherichia coli* chromosome. J Biol Chem 262:10327-34.

Gamas P, Chandler MG, Prentki P, Galas DJ. 1987. *Escherichia coli* integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot spot in pBR322. J Mol Bio 2:262-272.

Gao F, Luo H, Zhang CT. 2013. D*oriC* 5.0: an updated database of *oriC* regions in both bacterial and archaeal genomes. Nucleic Acids Res. 41:D90–D93.

Gao F, Zhang CT. 2007. D*oriC*: a database of *oriC* regions in bacterial genomes. Bioinformatics 23:1866–1867.

Gardner JF, Nash HA. 1986. Role of *Escherichia coli* IHF protein in lambda site-specific recombination. A mutational analysis of binding sites. J Mol Bio 191:181-189.

Garner J, Crooke E. 1996. Membrane regulation of the chromosomal replication activity of *E. coli* DnaA requires a discrete site on the protein. EMBO J 15:3477–3485.

Gille H, Egan JB, Roth A, Messer W. 1991. The Fis protein binds and bends the origin of chromosomal DNA replication, *oriC*, of *Escherichia coli*. Nucleic Acids Res. 19:4167-72.

Grimwade JE, Ryan VT, Leonard AC. 2000. IHF redistributes bound initiator protein, DnaA, on supercoiled *oriC* of *Escherichia coli*. Mol. Microbiol. 35:835-44.

Grimwade JE, Torgue JJC, McGarry KC, Rozgaja T, Enloe ST, Leonard AC. 2007. Mutational analysis reveals *Escherichia coli oriC* interacts with both DnaA-ATP and DnaA-ADP during pre-RC assembly. Mol. Microbiol. 66:428–439.

Hansen EB, Yarmolinsky MB. 1986. Host participation in plasmid maintenance: Dependence upon dnaA of replicons derived from P1 and F. Proc. Natl. Acad. Sci. U.S.A. 83(12):4423-4427.

Hansen FG, Christensen BB, Atlung T. 2007. Sequence characteristics required for cooperative binding and efficient in vivo titration of the replication initiator protein DnaA in *E. coli*. J Mol Biol. Apr 6;367(4):942-52.

Heacock PN, Dowhan W. 1987. Construction of a lethal mutation in the synthesis of the major acidic. J Biol Chem. Sep 25;262(27):13044-9.

Helmstetter CE. 1996. Timing of synthetic activities in the cell cycle, p. 1627–1639. In F. C. Neidhardt, R. Curtis, III, J. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), Escherichia coli and Salmonella: Cellular and Molecular Biology. ASM Press, Washington, DC.

Helmstetter CE. Cooper S. 1968. DNA synthesis during the division cycle of rapidly growing *Escherichia coli* B/r. J Mol Biol 31, 507–518.

Hengen PN, Lyakhov IG, Stewart LE, Schneider TD. 2003. Molecular flip-flops formed by overlapping Fis sites. Nucleic Acids Res. Nov 15;31(22):6663-73.

Hsu J, Bramhill D, Thompson CM. 1994. Open complex formation by DnaA initiation protein at the *Escherichia coli* chromosomal origin requires the 13-mers precisely spaced relative to the 9-mers. Mol Microbiol. Mar; 11(5):903-11.

Ishida T, Akimitsu N, Kashioka T, Hatano M, Kubota T, Ogata Y, Sekimizu K, Katayama T. 2004. DiaA, a novel DnaA-binding protein, ensures the timely initiation of *Escherichia coli* chromosome replication. J Biol Chem. Oct 29;279(44):45546-55.

Iyer LM, Leipe DD, Koonin EV, Aravind L. 2004. Evolutionary history and higher order classification of AAA+ ATPases. J Struct Biol 146:11–31.

Johnsen L, Flatten I, Morigen, Dalhus B, Bjoras M, Waldminghaus T, Skarstad K. 2011. The G157C mutation in the *Escherichia coli* sliding clamp specifically affects initiation of replication. Mol Microbiol 79: 433–446.

Kaguni JM, Kornberg A. 1984. Replication initiated at the origin (*oriC*) of the *E. coli* chromosome reconstituted with purified enzymes. Cell 38:183–190.

Kang S, Han JS, Park JH, Skarstad K, Hwang DS. 2003. SeqA protein stimulates the relaxing and decatenating activities of Topoisomerase IV. J Biol Chem 278:48779–48785.

Kasho K, Fujimitsu K, Matoba T, Oshima T, Katayama T. 2014. Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production and replication initiation. Nucleic Acids Res. 42:13134–13149.

Kasho K, Katayama T. 2013. DnaA binding locus datA promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. Proc. Natl. Acad. Sci. U.S.A. 110:936–941.

Kasho K, Tanaka H, Sakai R, Katayama T. 2017. Cooperative DnaA Binding to the Negatively Supercoiled datA Locus Stimulates DnaA-ATP Hydrolysis. J Biol Chem. Jan 27;292(4):1251-1266.

Katayama T, Kasho K, Kawakami H. 2017. The DnaA Cycle in *Escherichia coli*: Activation, Function and Inactivation of the Initiator Protein. Front Microbiol. Dec 21; 8:2496.

Katayama T, Kubota T, Kurokawa K, Crooke E, Sekimizu K. 1998. The initiator function of DnaA protein is negatively regulated by the sliding clamp of the *E. coli* chromosomal replicase. Cell 94:61–71.

Katayama T, Ozaki S, Keyamura K, Fujimitsu K. 2010. Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. Nat Rev Microbiol 8:163–170.

Kato J, Katayama T. 2001. Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*. EMBO J 20:4253–4262.

Kaur G, Vora MP, Czerwonka CA, Rozgaja TA, Grimwade JE, Leonard AC. 2014. Building the bacterial orisome: high-affinity DnaA recognition plays a role in setting the conformation of *oriC* DNA. Mol Microbiol. Mar; 91(6):1148-63.

Kawakami H, Katayama T. 2010. DnaA, ORC, and Cdc6: similarity beyond the domains of life and diversity. Biochem. Cell Biol., 88:49–62.

Kawakami H, Keyamura K, Katayama T. 2005. Formation of an ATP-DnaA-specific initiation complex requires DnaA Arginine 285, a conserved motif in the AAA+ protein family. J. Biol. Chem. 280:27420–27430.

Kawakami H, Ozaki S, Suzuki S, Nakamura K, Senriuchi T, Su'etsugu M, Fujimitsu K, Katayama T. 2006. The exceptionally tight affinity of DnaA for ATP/ADP requires a unique aspartic acid residue in the AAA+ sensor 1 motif. Mol Microbiol 62:1310–1324.

Keyamura K, Abe Y, Higashi M, Ueda T, Katayama T. 2009. DiaA dynamics are coupled with changes in initial origin complexes leading to helicase loading. J Biol Chem 284:25038–25050.

Keyamura K, Fujikawa N, Ishida T, Ozaki S, Su'etsugu M, Fujimitsu K, Kagawa W, Yokoyama S, Kurumizaka H, Katayama T. 2007. The interaction of DiaA and DnaA regulates the replication cycle in E. coli by directly promoting ATP-DnaA-specific initiation complexes. Genes Dev. 21, 2083–2099.

Keyamura K, Katayama T. 2011. DnaA protein DNA-binding domain binds to Hda protein to promote inter-AAA+ domain interaction involved in regulatory inactivation of DnaA. J Biol Chem 286:29336–29346.

Kim JS, Nanfara MT, Chodavarapu S, Jin KS, Babu VMP, Ghazy MA, Chung S, Kaguni JM, Sutton MD, Cho Y. 2017. Dynamic assembly of Hda and the sliding clamp in the regulation of replication licensing. Nucleic Acids Res. Apr 20;45(7):3888-3905.

Kitagawa R, Mitsuki H, Okazaki T, Ogawa T. 1996. A novel DnaA protein-binding site at 94.7 min on the *Escherichia coli* chromosome. Mol. Microbiol. 19:1137–1147.

Kitagawa R, Ozaki T, Moriya S, Ogawa T. 1998. Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein. Genes Dev 12:3032–3043.

Koppes L, Nordström K. 1986. Insertion of an R1 plasmid into the origin of replication of the *E. coli* chromosome: random timing of replication of the hybrid chromosome. Cell. Jan 17;44(1):117-24.

Kowalski D, Eddy MJ. 1989. The DNA unwinding element: a novel, cis acting component that facilitates opening of the *Escherichia coli* replication origin. Embo J 8:4335-44.

Kurokawa K, Nishida S, Emoto A, Sekimizu K and Katayama T. 1999. Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*. The EMBO Journal 18 (23):6642–6652.

Langer U, Richter S, Roth A, Weigel C, Messer W. 1996. A comprehensive set of DnaA-box mutations in the replication origin, *oriC*, of *Escherichia coli*. Mol Microbiol 21:301-11.

Leonard AC, Grimwade JE. 2010. Regulating DnaA complex assembly: it is time to fill the gaps. Curr. Opin. Microbiol. 13(6):766–772.

Leonard AC, Grimwade JE. 2011. Regulation of DnaA assembly and activity: taking directions from the genome. Annu. Rev. Microbiol. 65, 19–35.

Leonard AC, Grimwade JE. 2015. The orisome: structure and function. Front Microbiol. Jun 2;6:545.

Leonard AC, Méchali M. 2013. DNA replication origins. Cold Spring Harb Perspect Biol. Oct 1; 5(10):a010116.

Leong JM, Nunes-Duby S, Lessner CF, Youderian P, Kusskind MM, Landy A. 1985. The φ80 and P22 attachment sites: primary structure and interaction with *Escherichia coli* integration host factor. J Biol Chem 260:4468-4477.

Lgbner-Olesen A, Skarstad K, Hansen FG, Von Meyenburg K, Boye E. 1989. The DnaA Protein Determines the Initiation Mass of *Escherichia coli* K-12. Cell. Cell Press, June 2, Vol. 57. 881-889.

Li Z, Crooke E. 1999. Functional Analysis of Affinity-Purified Polyhistidine Tagged DnaA Protein. Protein Expres Purif 17: 41-48.

Løbner-Olesen A, Hansen FG, Rasmussen KV, Martin B, Kuempel PL. 1994. The initiation cascade for chromosome replication in wild-type and Dam methyltransferase deficient *Escherichia coli* cells. EMBO J. Apr 15; 13(8):1856-62.

Løbner-Olesen A, Skarstad K, Hansen FG, Von Meyenburg K, Boye E. 1989. The DnaA protein determines the initiation mass of *Escherichia coli* K-12. Cell 57:881–889.

Lu M, Campbell JL, Boye E, Kleckner N. 1994. SeqA: a negative modulator of replication initiation in *E. coli*. Cell. May 6; 77(3):413-26.

Marczynski GT, Shapiro L. 1992. Cell-cycle control of a cloned chromosomal origin of replication from Caulobacter crescentus. J. Mol. Biol. 226:959–977.

Margulies C, Kaguni JM. 1996. Ordered and sequential binding of DnaA protein to *oriC*, the chromosomal origin of *Escherichia coli*. J. Biol. Chem. 271:17035–17040.

Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M. 2010. Eukaryotic chromosome DNA replication: where, when, and how? Annu. Rev. Biochem. 79:89–130.

Matsui M, Oka A, Takanami M, Yasuda S, Hirota Y. 1985. Sites of dnaA protein-binding in the replication origin of the *Escherichia coli* K-12 chromosome (1985) J. Mol. Biol. 184:529–533.

McGarry KC, Ryan VT, Grimwade JE, Leonard AC. 2004. Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA- ATP. Proc Natl Acad Sci USA 101:2811–2816.

McMacken R, Silver L, Georgopoulos C. 1987. DNA replication. In: *Escherichia coli and Salmonella typhimurium*: Cellular and Molecular Biology, pp. 564-612. Edited by Ingraham JL, Low KB, Magasanik B, Neidhardt FC, SchaechterM, Umbarger HE. Washington, DC: American Society for Microbiology

Messer W, Blaesing F, Majka J, Nardmann J, Schaper S, Schmidt A, Seitz H, Speck C, Tüngler D, Wegrzyn G, Weigel C, Welzeck M, Zakrzewska-Czerwinska J. 1999. Functional domains of DnaA proteins. Biochimie. Aug-Sep; 81(8-9):819-25.

Messer W, Hartmann KH, Langer U, Mahlow E, Roth A, Schaper S, Urmoneit B, Woelker B. 1992. The complex for replication initiation of *Escherichia coli*. Chromosoma. 102:S1–S6.

Michelsen O, De Mattos MJT, Jensen PR, Hansen FG. 2003. Precise determinations of c and d periods by flow cytometry in *Escherichia coli* k-12 and b/r. Microbiology. 149(4):1001–1010.

Miller DT, Grimwade JE, Betteridge T, Rozgaja T, Torgue JJ, Leonard AC. 2009. Bacterial origin recognition complexes direct assembly of higher-order DnaA oligomeric structures. Proc Natl Acad Sci USA 106:18479–18484.

Molin S, Nordström K. 1980. Control of plasmid R1 replication: functions involved in replication, copy number control, incompatibility, and switch-off of replication. J Bacteriol. Jan; 141(1):111-20.

Moriya S, Fukuoka T, Ogasawara N, Yoshikawa H. 1998. Regulation of initiation of chromosomal replication by DNA-boxes in the origin region of the Bacillus subtilis chromosome. The EMBO Journal 7(9):2911-7.

Mott ML, Berger JM. 2007. DNA replication initiation: mechanisms and regulation in bacteria. Nat Rev Microbiol 5: 343–354.

Mott ML, Erzberger JP, Coons MM, Berger JM. 2008. Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators. Cell 135:623–634.

Neuwald AF, Aravind L, Spouge JL, Koonin EV. 1999. AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res 9:27–43.

Nievera C, Torgue JJ, Grimwade JE, Leonard AC. 2006. SeqA blocking of DnaA-*oriC* interactions ensures staged assembly of the E coli pre-RC. Mol Cell 24:581–592.

Nishida S, Fujimitsu K, Sekimizu K, Ohmura T, Ueda T, Katayama T. 2002. A nucleotide switch in the *Escherichia coli* DnaA protein initiates chromosomal replication: evidence from a mutant DnaA protein defective in regulatory ATP hydrolysis in vitro and in vivo. J Biol Chem 277:14986–1499.

Nozaki S, Ogawa T. 2008. Determination of the minimum domain II size of *Escherichia coli* DnaA protein essential for cell viability. Microbiology 154:3379-84.

Nozaki S, Yamada Y, Ogawa T. 2009. Initiator titration complex formed at datA with the aid of IHF regulates replication timing in *Escherichia coli*. Genes Cells 14:329–341.

Ogawa T, Yamada Y, Kuroda T, Kishi T, Moriya S. 2002. The datA locus predominantly contributes to the initiator titration mechanism in the control of replication initiation in *Escherichia coli*. Mol. Microbiol. 44: 1367–1375.

Oka A, Sasaki H, Sugimoto K, Takanami M. 1984. Sequence organization of replication origin of the *Escherichia coli* K-12 chromosome. J Mol Biol. Jul 15; 176(4):443-58.

Ozaki S, Katayama T. 2012. Highly organized DnaA-*oriC* complexes recruit the single-stranded DNA for replication initiation. Nucleic Acids Res 40:1648–1665.

Ozaki S, Kawakami H, Nakamura K, Fujikawa N, Kagawa W, Park SY, Yokoyama S, Kurumizaka H, Katayama T. 2008. A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. J Biol Chem 283:8351– 8362.

Ozaki S, Noguchi Y, Hayashi Y, Miyazaki E, Katayama T. 2012a. Differentiation of the DnaA-*oriC*-subcomplex for DNA unwinding in a replication initiation complex. J Biol Chem 287:37458–37471.

Ozaki S, Noguchi Y, Nishimura M, Katayama T. 2012b. Stable nucleotide binding to DnaA requires a specific glutamic acid residue within the AAA+ box II motif. J Struct Biol 179:242–250.

Pan CQ, Feng JA, Finkel SE, Landgraf R, Sigman D, Johnson RC. 1994. Structure of the *Escherichia coli* Fis-DNA complex probed by protein conjugated with 1, 10-phenanthroline copper (I) complex. Proc Natl Acad Sci USA 91:1721-1725.

Parker MW, Botchan MR, Berger JM. 2017. Mechanisms and regulation of DNA replication initiation in eukaryotes. Critical Reviews in Biochemistry and Molecular Biology Apr;52(2):107-144 .

Pierucci O, Rickert M, Helmstetter CE. 1989. DnaA protein overproduction abolishes cell cycle specificity of DNA replication from *oriC* in *Escherichia coli*. J Bacteriol 171:3760-6.

Riber L, Fujimitsu K, Katayama T, Løbner-Olesen A. 2009. Loss of Hda activity stimulates replication initiation from I-box, but not R4 mutant origins in *Escherichia coli*. Mol Microbiol 71:107–122.

Rice PA, Yang S, Mizuuchi K. 1996. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell 87:1295-1306.

Richardson TT, Harran O, Murray H. 2016. The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. Nature. Jun 16; 534(7607):412-6.

Roth A, Messer W. 1995. The DNA binding domain of the initiator protein DnaA. Embo J 14:2106-11.

Roth A, Messer W. 1998. High-affinity binding sites for the initiator protein DnaA on the chromosome of *Escherichia coli*. Mol. Microbiol. 28:395–401.

Rozgaja TA, Grimwade JE, Iqbal M, Czerwonka C, Vora M, Leonard AC. 2011. Two oppositely oriented arrays of low-affinity recognition sites in *oriC* guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. Mol. Microbiol. 82:475–488.

Ryan VT, Grimwade JE, Camara JE, Crooke E, Leonard AC. 2004. *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. Mol. Microbiol. 51:1347–1359.

Samitt CE, Hansen FG, Miller JF, Schaechter M. 1989. In vivo studies of DnaA binding to the origin of replication of *Escherichia coli*. EMBO J. 8:989–993.

Saxena BR. 2013b. The initiator protein DnaA, chromosomal origin *oriC* and their interaction to generate origin recognition complex and pre-replication complex like nucleoprotein structures in *Escherichia coli*. OA Biochem. 1:13.

Saxena R, Fingland N, Patil D, Sharma AK, Crooke E. 2013a. Crosstalk between DnaA protein, the initiator of *Escherichia coli* chromosomal replication, and acidic phospholipids present in bacterial membranes. Int J Mol Sci 14:8517–8537.

Saxena R, Rozgaja T, Grimwade J, Crooke E. 2011. Remodeling of nucleoprotein complexes is independent of the nucleotide state of a mutant AAA+ protein. J. Biol. Chem. 286:33770–33777.

Schaper S, Messer W. 1995. Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. J Biol Chem 270:17622-6.

Seitz H, Weigel C, Messer W. 2000. The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. Mol Microbiol 37:1270-9.

Sekimizu K, Bramhill D, Kornberg A. 1987. ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E.coli chromosome. Cell 50:259–265.

Sekimizu K, Bramhill D, Kornberg A. 1988b. Sequential early stages in the in vitro initiation of replication at the origin of the *Escherichia coli* chromosome. J.Biol.Chem. 263:7124–7130.

Sekimizu K, Kornberg A. 1988. Cardiolipin activation of dnaA protein, the initiation protein of replication in *Escherichia coli*. J Biol Chem 263:7131–7135.

Sekimizu K, Yung BY, Kornberg A. 1988a. The DnaA protein of *Escherichia coli*. Abundance, improved purification, and membrane binding. J Biol Chem 263:7136–7140.

Shaheen SM, Ouimet MC, Marczynski GT. 2009. Comparative analysis of Caulobacter chromosome replication origins. Microbiology 155:1215–1225.

Shimizu M, Noguchi Y, Sakiyama Y, Kawakami H, Katayama T, Takada S. 2016. Near-atomic structural model for bacterial DNA replication initiation complex and its functional insights. Proc Natl Acad Sci U S A 113:E8021–E8030

Simmons LA, Felczak M, Kaguni JM. 2003. DnaA protein of *Escherichia coli*: oligomerization at the *E. coli* chromosomal origin is required for initiation and involves specific N-terminal amino acids. Mol Microbiol 49:849– 858

Skarstad K, Baker TA, Kornberg A. 1990. Strand separation required for initiation of replication at the chromosomal origin of E.coli is facilitated by a distant RNA--DNA hybrid. EMBO J. Jul; 9(7):2341-8.

Skarstad K, Boye E, Steen HB. 1986. Timing of initiation of chromosome replication in individual *Escherichia coli* cells. EMBO J Nov; 5(11):3074.

Skarstad K, Katayama T. 2013. Regulating DNA replication in bacteria. Cold Spring Harb. Perspect. Biol. 5:a012922.

Snider J, Thibault G, Houry WA. 2008. The AAA+ superfamily of functionally diverse proteins. Genome Biol. Apr 30; 9(4):216.

Soultanas P. 2012. Loading mechanisms of ring helicases at replication origins. Mol Microbiol. Apr; 84(1):6-16.

Speck C, Messer W. 2001. Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. The EMBO Journal 20:1469-1476.

Speck C, Weigel C, Messer W. 1999. ATP- and ADP-dnaA protein, a molecular switch in gene regulation. EMBO J. Nov 1; 18(21):6169-76.

Stepankiw N, Kaidow A, Boye E, Bates D. 2009. The right half of the *Escherichia coli* replication origin is not essential for viability, but facilitates multi-forked replication. Mol Microbiol 74: 467-479.

Stillman B. 2005. Origin recognition and the chromosome cycle. FEBS Lett 579:877-84.

Strauss F, Varshavsky A. 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37:889–901.

Su'etsugu M, Shimuta TR, Ishida T, Kawakami H, Katayama T. 2005. Protein associations in DnaA-ATP hydrolysis mediated by the Hda-replicase clamp complex. J. Biol. Chem. 280:6528–6536.

Su'etsugu M, Takata M, Kubota T, Matsuda Y, and Katayama T. 2004. Molecular mechanism of DNA replication-coupled inactivation of the initiator protein in *Escherichia coli*: interaction of DnaA with the sliding clamp-loaded DNA and the sliding clamp-Had complex. Genes Cells 9:509–522.

Sun J, Kawakami H, Zech J, Speck C, Stillman B, Li H. 2012. Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. Structure 20:534–544.

Sutton MD, Carr KM, Vicente M, Kaguni JM. 1998. *Escherichia coli* DnaA protein. The N-terminaldomain and loading of DnaB helicase at the *E. coli* chro-mosomal origin. J Biol Chem 273:34255–34262.

Sutton MD, Kaguni JM. 1997. The *Escherichia coli* dnaA gene: four functional domains. J Mol Biol 274:546-61.

Theisen PW, Grimwade JE, Leonard AC, Bogan JA, Helmstetter CE. 1993. Correlation of gene transcription with the time of initiation of chromosome replication in Escherichia coli. Mol Microbiol 10:575-84.

Thompson JF, Landy A. 1988. Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. Nucleic Acids Res 16(20):9687-705.

Torheim NK, Skarstad K. 1999. *Escherichia coli* SeqA protein affects DNA topology and inhibits open complex formation at *oriC*. EMBO J 18:4882–4888.

Travers A. 2006. DNA topology: dynamic DNA looping. Curr Biol 16: R838–R840.

Waldminghaus T, Skarstad. 2009. The *Escherichia coli* SeqA protein. Plasmid 61:141–150.

Weigel C, Messer W, Preiss S, Welzeck M, Morigen, Boye E. 2001. The sequence requirements for a functional *Escherichia coli* replication origin are different for the chromosome and a minichromosome. Mol Microbiol 40:498–507.

Weigel C, Schmidt A, Ruckert B, Lurz R, Messer W. 1997. DnaA protein binding to individual DnaA boxes in the *Escherichia coli* replication origin, *oriC*. Embo J 16(21):6574-83.

Weigel C, Schmidt A, Seitz H, Tüngler D, Welzeck M, Messer W. 1999. The N-terminus promotes oligomerization of the *Escherichia coli* initiator protein DnaA. Mol Microbiol. Oct;34(1):53-66.

Woelker B, Messer W. 1993. The structure of the initiation complex at the replication origin, *oriC*, of *Escherichia coli*. Nucleic Acids Res 21:5025–5033.

Wolanski M, Donczew R, Zawilak-Pawlik A, Zakrzewska-Czerwinska J. 2015. *oriC*-encoded instructions for the initiation of bacterial chromosome replication. Front. Microbiol. 5:1–14.

Yung BY, Crooke E, Kornberg A. 1990. Fate of the DnaA initiator protein in replication at the origin of the *Escherichia coli* chromosome in vitro. J Biol Chem 265:1282–1285

Yung BY, Kornberg A. 1988. Membrane attachment activates dnaA protein, the initiation protein of chromosome replication in *Escherichia coli*. Proc Natl Acad Sci USA 85:7202–7205

Zakrzewska-Czerwinska J, Jakimowicz D, Zawilak-Pawlik A, Messer W. 2007. Regulation of the initiation of chromosomal replication in bacteria. FEMS Microbiol Rev31:378–387.

Zawilak-Pawlik A, Kois A, Majka J, Jakimowicz D, Smulczyk-Krawczyszyn A, Messer W, Zakrzewska-Czerwińska J. 2005. Architecture of bacterial replication initiation complexes: orisomes from four unrelated bacteria. Biochem. J. 389:471–481.

Zawilak-Pawlik A, Nowaczyk M, Zakrzewska-Czerwińska J. 2017. The Role of the N-Terminal Domains of Bacterial Initiator DnaA in the Assembly and Regulation of the Bacterial Replication Initiation Complex. Genes (Basel). May 10;8(5):E136.

Zyskind JW, Smith DW. 1986. The bacterial origin of replication, *oriC*. Cell. Aug 15; 46(4):489-90.