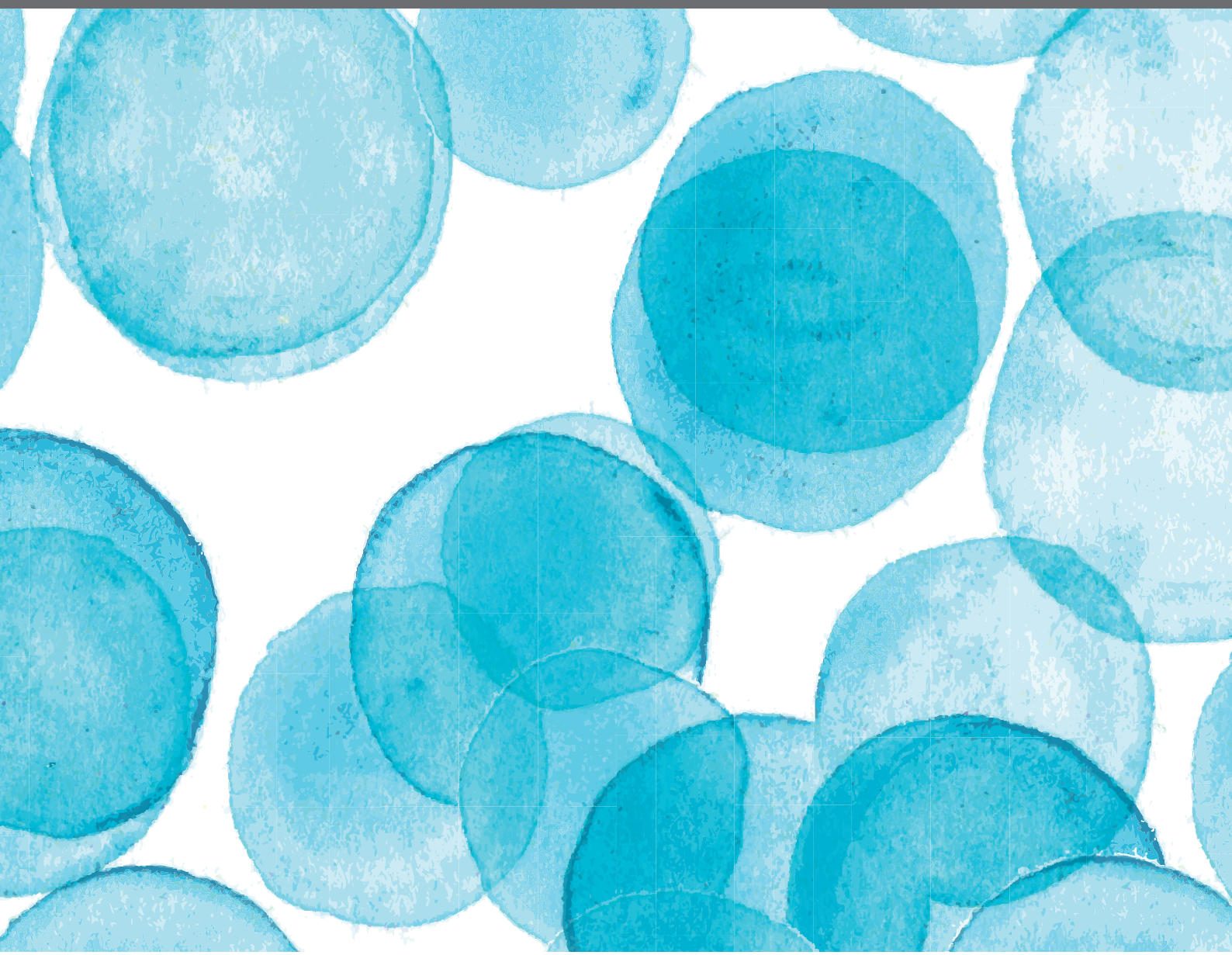
BACTERIAL CHROMOSOMES UNDER CHANGING ENVIRONMENTAL CONDITIONS

EDITED BY: Leise Riber, Monika Glinkowska and Torsten Waldminghaus
PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-729-1
DOI 10.3389/978-2-88966-729-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

BACTERIAL CHROMOSOMES UNDER CHANGING ENVIRONMENTAL CONDITIONS

Topic Editors:
Leise Riber, University of Copenhagen, Denmark
Monika Glinkowska, University of Gdansk, Poland
Torsten Waldminghaus, Darmstadt University of Technology, Germany

Citation: Riber, L., Glinkowska, M., Waldminghaus, T., eds. (2021). Bacterial Chromosomes Under Changing Environmental Conditions. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-729-1

Table of Contents

frontiers in Microbiology

EDITORIAL
published: 11 March 2021
doi: 10.3389/fmicb.2021.633466

Editorial: Bacterial Chromosomes Under Changing Environmental Conditions

Monika Glinkowska[1]*, *Torsten Waldminghaus*[2,3]* and *Leise Riber*[4]*

[1] Department of Bacterial Molecular Genetics, University of Gdansk, Gdańsk, Poland, [2] Department of Biology, Technische Universität Darmstadt, Darmstadt, Germany, [3] Centre for Synthetic Biology, Technische Universität Darmstadt, Darmstadt, Germany, [4] Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

Keywords: bacterial chromosomes, nucleoid architecture, DNA replication, chromosome segregation, DNA repair, stress conditions, environmental conditions, replication-transcription conflicts

Editorial on the Research Topic

Bacterial Chromosomes Under Changing Environmental Conditions

OPEN ACCESS

Edited by:
Brian Palenik,
University of California, San Diego, United States

Reviewed by:
Jolanta Zakrzewska-Czerwinska, University of Wrocław, Poland

****Correspondence:***
Monika Glinkowska
monika.glinkowska@ug.edu.pl
Torsten Waldminghaus
Torsten.Waldminghaus@TU-Darmstadt.de
Leise Riber
lriber@plen.ku.dk

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *25 November 2020*
Accepted: *18 February 2021*
Published: *11 March 2021*

Citation:
Glinkowska M, Waldminghaus T and Riber L (2021) Editorial: Bacterial Chromosomes Under Changing Environmental Conditions. Front. Microbiol. 12:633466. doi: 10.3389/fmicb.2021.633466

The bacterial cell cycle comprises chromosome replication and segregation of newly replicated chromosomes into daughter cells prior to cell division. Unlike in eukaryotic organisms, DNA replication, chromosome segregation, and transcription occur simultaneously in bacteria. Several molecular mechanisms act in concert to allow chromosome replication initiation once-and-only-once per cell cycle (Skarstad et al., 1986; Boye et al., 2000). Other mechanisms ensure that replication is coordinated with cell growth (Murray, 2016) and linked to chromosome segregation in a tightly coordinated manner (Blow and Tanaka, 2005; Reyes-Lamothe et al., 2012). Considering that the chromosome is a massively compact structure, organization of the bacterial nucleoid adds an extra level to cell cycle coordination. In particular, a balance has to be reached between the requirement of significant compaction and an unobstructed accessibility to molecular processes underlying essential cellular functions, such as replication, transcription, DNA repair and homologous recombination (Badrinarayanan et al., 2015; Magnan and Bates, 2015).

As single cell organisms, bacteria constantly adapt to ever changing environmental conditions. Typical examples of adverse environmental conditions include exposure to antibiotics, nutrient limitation, changes in physical parameters (pH, temperature, osmotic pressure), exposure to ultraviolet (UV) radiation or oxidative stress. As such conditions might influence control mechanisms governing chromosome dynamics (Golovlev, 2003; Tymeca-Mulik et al., 2017; Remesh et al., 2020), bacterial cells have developed highly advanced survival strategies to overcome the negative impact of the encountered stress. Those strategies enable them to adapt to unfavorable conditions, thereby allowing colonization of various ecological niches including host organisms (Boor, 2006; Hauryliuk et al., 2015). Previously, the acquisition of foreign DNA was suggested to impact on nucleoid structure as well (Krogh et al., 2018). However, the molecular mechanisms of how chromosome maintenance systems cope with such challenges and potentially help to overcome significant threats to genome stability, cell cycle regulation and viability, remain poorly understood.

This Research Topic represents a comprehensive collection of articles focusing on the influence of changing environmental conditions on bacterial chromosome dynamics, such as chromosome organization, replication, segregation and DNA repair.

On the topological level, dynamic organization of the nucleoid involves a tight balance between efficient compaction within the cell and concomitant accessibility to replication, transcription and DNA repair processes. Original research by Krogh et al. demonstrates the importance of chromosome architecture on transcriptional regulation in bacteria.

Co-expression of genes was found to correlate positively with increased spatial proximity. This suggests that nucleoid structure, through the ability to bring genes together, strongly influences the amount of transcriptional spilling into neighboring genes.

Changing environmental conditions are known to induce profound topological alterations of the chromosome structure. NAPs (nucleoid associated proteins) have been shown to play a major role in adjusting bacterial chromosome architecture and affecting gene transcription regulation in response to stress, as reviewed by Holowka and Zakrzewska-Czerwinska.

Besides nucleoid organization, replication, transcription and DNA repair processes are by themselves sensitive to changing environmental conditions. In this context, Sinha et al. review the current understanding of molecular mechanisms underlying the negative impact of the stringent stress response regulator, (p)ppGpp, on chromosome replication initiation in *Escherichia coli,* and on chromosome replication elongation in *Bacillus subtilis*, respectively. In addition, they propose a model where (p)ppGpp-mediated survival during DNA damage is linked to the ability of (p)ppGpp to inhibit replication initiation, which minimizes the frequency of replication-transcription collisions, hence allowing backtracking of RNA polymerase to repair genotoxic DNA lesions. An additional facet of stringent response-mediated regulation of DNA replication initiation in bacteria is the polyphosphate-induced proteolysis and degradation of essential replication proteins, as reviewed by Ropelewska et al.

In most bacteria, DnaA-*oriC* dependent replication initiation is considered an essential mechanism. However, Ohbayashi et al. demonstrate variations in cyanobacterial chromosome replication mechanisms, manifested by regular/irregular GC skew profiles. Here, the genomes of certain free-living species are found not to encode *dnaA*, and instead chromosome replication in those organisms is initiated from multiple origins in a DnaA-independent manner. This replication mode produces irregular GC skew profiles, indicating that loss of DnaA-*oriC* dependencies might play a crucial role in cyanobacterial evolution.

In the context of elucidating novel DNA replication processes, Oliveira Paiva et al. revealed that a pathogen *Clostridioides difficile* utilizes a bipartite origin of replication, possibly conserved among *Clostridioides* species. Within this origin DnaA-dependent unwinding occurs at *oriC2,* in the *dnaA-dnaN* intergenic region.

During normal growth conditions, chromosome replication progression is tightly coordinated with simultaneous gene transcription. If, however, additional origins are engineered into different ectopic genomic locations, native replichore arrangements are disturbed and genome trafficking events, such as replication-transcription conflicts, might arise. Syeda et al. review current models of how such replication-transcription conflicts contribute to shaping of the distinct architecture of bacterial chromosomes.

Replication-transcription collisions also induce multiple repair pathways required to restart arrested replication forks. A novel insight into DNA repair mechanisms is brought by Sheng et al. who report that two *recA* variants are induced by UV in *Myxococcus xanthus* cells, each playing a different role in cell growth and UV-radiation resistance. Most bacteria, including *E. coli*, possess a single *recA* gene, and duplicate *recA* genes have been investigated only in *Bacillus megaterium* and *Myxococcus xanthus*. The findings of this research article, therefore, add a valuable insight onto the functional divergence among duplicated *recA* genes in bacteria.

Finally, chromosome segregation constitutes an essential stage of cell cycle progression, and ParA and ParB are known as main players in cellular positioning of the replication origin prior to cell division in most bacterial species. However, ParA and ParB have been shown to interact with proteins involved in cell division or cell elongation. Based on this, Pioro and Jakimowicz review evidence on the regulatory role of segregation proteins in cell cycle progression and cover the current understanding of its coordination with environmental conditions.

In conclusion, this Research Topic highlights a selection of original research- and review articles representing the current progress within understanding molecular survival strategies adopted by bacteria to preserve cell cycle regulation and genome integrity in response to changing environmental conditions. We kindly thank all contributors, authors as well as reviewers, for their valuable time, thoughts and input for this Research Topic published in the section of Evolutionary and Genomic Microbiology in Frontiers in Microbiology. We hope that the readers will enjoy their work as much as we have.

AUTHOR CONTRIBUTIONS

MG, TW, and LR were joint co-editors of this Research Topic and co-wrote the editorial. All authors contributed equally to this article and approved the submitted version.

FUNDING

This work was supported by a grant from the National Science Center (Poland) (no. UMO-2017/27/B/NZ2/00747) as well as by a grant from the Lundbeck Foundation (Denmark) (no. R324-2019-2001).

ACKNOWLEDGMENTS

We kindly thank all contributors, authors as well as reviewers, for their valuable time, thoughts and input for this Research Topic published in the section of Evolutionary and Genomic Microbiology in Frontiers in Microbiology. We hope that the readers will enjoy their work as much as we have.

REFERENCES

Badrinarayanan, A., Le, T. B. K., and Laub, M. T. (2015). Bacterial chromosome organization and segregation. *Annu. Rev. Cell Dev. Biol.* 31, 171–199. doi: 10.1146/annurev-cellbio-100814-125211

Blow, J. J., and Tanaka, T. U. (2005). The chromosome cycle: coordinating replication and segregation. *EMBO Rep.* 6, 1028–1034. doi: 10.1038/sj.embor.7400557

Boor, K. J. (2006). Bacterial stress responses: what doesn't kill them can make them stronger. *PLoS Biol.* 4:e0040023. doi: 10.1371/journal.pbio.0040023

Boye, E., Løbner-Olesen, A., and Skarstad, K. (2000). Limiting DNA replication to once and only once. *EMBO Rep.* 1, 479–483. doi: 10.1093/embo-reports/kvd116

Golovlev, E. L. (2003). Bacterial cold shock response at the level of DNA transcription, translation, and chromosome dynamics. *Microbiology* 72, 1–7. doi: 10.1023/A:1022276318653

Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* 13, 298–309. doi: 10.1038/nrmicro3448

Krogh, T. J., Møller-Jensen, J., and Kaleta, C. (2018). Impact of chromosomal architecture on the function and evolution of bacterial genomes. *Front. Microbiol.* 9:2019. doi: 10.3389/fmicb.2018.02019

Magnan, D., and Bates, D. (2015). Regulation of DNA replication initiation by chromosome structure. *J. Bacteriol.* 197, 3370–3377. doi: 10.1128/JB.00446-15

Murray, H. (2016). Connecting chromosome replication with cell growth in bacteria. *Curr. Opin. Microbiol.* 34, 13–17. doi: 10.1016/j.mib.2016.07.013

Remesh, S. G., Verma, S. C., Chen, J.-H., Ekman, A. A., Larabell, C. A., Adhya, S., et al. (2020). Nucleoid remodeling during environmental adaptation is regulated by HU-dependent DNA bundling. *Nat. Comm.* 11:2905. doi: 10.1038/s41467-020-16724-5

Reyes-Lamothe, R., Nicolas, E., and Sherratt, D. J. (2012). Chromosome replication and segregation in bacteria. *Annu. Rev. Genet.* 46, 121–143. doi: 10.1146/annurev-genet-110711-155421

Skarstad, K., Boye, E., and Steen, H. B. (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *EMBO J.* 5, 1711–1717. doi: 10.1002/j.1460-2075.1986.tb04415.x

Tymeca-Mulik, J., Boss, L., Maciag-Dorszynska, M., Rodrigues, J. F. M., Gaffke, L., Wosinski, A., et al. (2017). Suppression of the *Escherichia coli dnaA46* mutation by changes in the activities of the pyruvate-acetate node links DNA replication regulation to central carbon metabolism. *PLoS ONE* 12:e0176050. doi: 10.1371/journal.pone.0176050

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 12 February 2020
doi: 10.3389/fmicb.2020.00140

Functional Division Between the RecA1 and RecA2 Proteins in *Myxococcus xanthus*

Duo-Hong Sheng, Yi-Xue Wang, Miao Qiu, Jin-Yi Zhao, Xin-Jing Yue and Yue-Zhong Li†*

State Key Laboratory of Microbial Technology, Institute of Microbial Technology, Shandong University, Qingdao, China

Myxococcus xanthus DK1622 has two RecA genes, *recA1* (MXAN_1441) and *recA2* (MXAN_1388), with unknown functional differentiation. Herein, we showed that both *recA* genes were induced by ultraviolet (UV) irradiation but that the induction of *recA1* was more delayed than that of *recA2*. Deletion of *recA1* did not affect the growth but significantly decreased the UV-radiation survival, homologous recombination (HR) ability, and induction of LexA-dependent SOS genes. In contrast, the deletion of *recA2* markedly prolonged the lag phase of bacterial growth and increased the sensitivity to DNA damage caused by hydrogen peroxide but did not change the UV-radiation resistance or SOS gene inducibility. Protein activity analysis demonstrated that RecA1, but not RecA2, catalyzed DNA strand exchange (DSE) and LexA autocleavage *in vitro*. Transcriptomic analysis indicated that RecA2 has evolved mainly to regulate gene expression for cellular transportation and antioxidation. This is the first report of functional divergence of duplicated bacterial *recA* genes. The results highlight the evolutionary strategy of *M. xanthus* cells for DNA HR and genome sophistication.

Keywords: RecA, duplicate genes, *Myxococcus xanthus*, DNA recombination, antioxidation, functional divergence, SOS response

OPEN ACCESS

Edited by:
Torsten Waldminghaus, University of Marburg, Germany

Reviewed by:
Juan Carlos Alonso, Centro Nacional de Biotecnología (CNB), Spain
Andrei Kuzminov, University of Illinois at Urbana–Champaign, United States

****Correspondence:***
Yue-Zhong Li
lilab@sdu.edu.cn

†ORCID:
Yue-Zhong Li
orcid.org/0000-0001-8336-6638

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *16 October 2019*
Accepted: *21 January 2020*
Published: *12 February 2020*

Citation:
Sheng D-H, Wang Y-X, Qiu M, Zhao J-Y, Yue X-J and Li Y-Z (2020) Functional Division Between the RecA1 and RecA2 Proteins in Myxococcus xanthus. Front. Microbiol. 11:140. doi: 10.3389/fmicb.2020.00140

INTRODUCTION

RecA, an ATP-dependent recombinase, is the core enzyme for DNA homologous recombination (HR), as well as being a promotion agent for LexA autolysis in bacteria (Lusetti and Cox, 2002). RecA also contributes to the repair of stalled and collapsed DNA replication forks by postreplication repair pathways (translesion DNA synthesis or template switching), playing an important role in DNA lesion tolerance pathways (Bichara et al., 2011; Quinet et al., 2017; Prado, 2018; Jaszczur et al., 2019). In addition, RecA participates in horizontal gene transfer between different strains (Lawrence and Retchless, 2009; Herrero-Fresno et al., 2015; García-Solache et al., 2016; He et al., 2016), which also causes genetic diversity. Thus, HR delicately balances genomic stability and diversity (Carr and Lambert, 2013; Greene, 2016). After binding to ssDNA, the RecA/ssDNA filament complex may serve as a signal of DNA damage, resulting in the self-cleavage of LexA, which activates the SOS response, increasing the expression of LexA-repressed genes. In the best characterized *Escherichia coli* SOS response, LexA autolysis derepresses the expression of more than 40 genes involved in DNA repair, mutagenesis, and many other cellular processes (Cox, 2003, 2007; Maslowska et al., 2019). Because of its pros and cons in genomic stability and variability, the functions of RecA are strictly regulated; for example, the function of RecA in *E. coli* is regulated at the gene transcription and protein activity levels. In the gene transcription induced by the

SOS response, particularly, there is a 10–20 times difference in gene expression before and after induction (Cox, 1999, 2007).

Most bacteria, including *E. coli*, have a single *recA* gene, while some bacteria possess duplicate *recA* genes; however, duplicate *recA* genes have been investigated only in *Bacillus megaterium* and *Myxococcus xanthus*. In *B. megaterium*, duplicate *recA* genes were found to both be damage-inducible and similarly showed some DNA repair ability in *E. coli* (Nahrstedt et al., 2005). In the model strain of myxobacteria, *M. xanthus* DK1622, both RecA1 (MXAN_1441) and RecA2 (MXAN_1388) can partly restore the UV resistance of the *E. coli recA* mutant, and *recA2*, but not *recA1*, was found to be inducible by mitomycin or nalidixic acid (Norioka et al., 1995; Campoy et al., 2003). It is unclear how the duplicate RecA proteins play divergent functions in the DNA recombination and SOS induction in this organism.

In this study, we genetically and biochemically investigated the functions of RecA1 and RecA2 in *M. xanthus*. We found that both *recA* genes were inducible by UV irradiation but in different periods. The *recA1* deletion had no significant effects on cellular growth but reduced the UV-radiation resistance and induction ability of the SOS gene. In contrast, the absence of *recA2* did not affect irradiation resistance but significantly reduced bacterial growth and resistance to oxidative damage. *In vitro* protein activity analysis indicated that RecA1, but not RecA2, had the homologous strand exchange activity and was able to promote LexA autolysis. Transcriptomic analysis indicated that the *recA2* gene was crucial for intracellular substance transport and antioxidant capacity. We discuss the molecular mechanisms for the functional divergence of the RecA1 and RecA2 proteins.

MATERIALS AND METHODS

Strains, Media, and DNA Substrates

The bacterial strains and plasmids used in this study are described in **Supplementary Table S1**. The *E. coli* strains were routinely grown on Luria-Bertani (LB) agar or in LB liquid broth at 37°C. The *M. xanthus* strains were grown in CYE liquid medium with shaking at 200 rpm or grown on agar plates with 1.5% agar at 30°C (Bretscher and Kaiser, 1978). When required, a final concentration of 40 μg/ml kanamycin (Kan) or 100 μg/ml ampicillin (Amp) was added to the solid or liquid medium.

Single-stranded viral DNA was isolated from M13mp18, and its 3 kb linear dsDNA was amplified by PCR and purified by a DNA purification kit (Tiangen, Beijing, China). A 60-nt oligomer from the M13 genome, 5′-CTG TCA ATG CTG GCG GCG GCT CTG GTG GTG GTT CTG GTG GCG GCT CTG AGG GTG GTG GCT-3′, was obtained from Tsingke Biotech (Qingdao). The 60-nt oligomer was ^{32}P-labelled using a polynucleotide kinase (Ausubel et al., 1995) and stored in TE buffer (10 mM Tris–HCl, pH 7.0, and 0.5 mM EDTA).

Growth and Resistance Analysis

Myxococcus xanthus strains were grown in CYE medium with shaking at 200 rpm at 30°C to an optical density of 0.5 at 600 nm (OD_{600}). Cells were then collected by centrifugation at 8000 rpm for 10 min, washed with 10 mM phosphate buffer (pH 7.0), and diluted to 1 OD_{600} in the same buffer.

For the radiation damage assay, cells in 10 mM phosphate buffer (pH 7.0) were irradiated at room temperature with a gradient dose from 0 to 200 J/m^2 using a UV Crosslinker (Fisher Scientific). Then, the cells were resuspended in fresh CYE medium and incubated at 30°C for 4 h. After incubation, cells were harvested by centrifugation and either used for a further assay or stored at −80°C.

For the oxidative damage assay, cells were suspended in phosphate buffer (pH 7.0) with a concentration of 1 OD, and hydrogen peroxide (H_2O_2) was added to a final concentration from 1 to 5 mM. The bacterial suspension was incubated for 20 min at room temperature with gentle shaking. After treatment, the suspension was immediately 10-fold diluted in the same phosphate buffer to end the oxidative damage reaction. Then, cells in the suspension were collected for further assay.

The growth assay was determined by growing cells in liquid medium at 30°C. Strains were inoculated at 0.02 OD_{600} and grown in CYE media for 84 h with shaking at 200 rpm. The OD_{600} was read every 12 h.

The survival rate was determined by a soft agar colony formation assay. Briefly, to determine the cell survival rate, *M. xanthus* cells were grown to the early exponential growth stage ($OD \approx 0.5$). The cells were treated with UV or H_2O_2 as described above and were then diluted with fresh medium and mixed at a 1:2 ratio with melted 0.6% soft agar (50°C). The mixture was then spread on CYE plates. After a few minutes for medium solidification, the cultures were incubated at 30°C until clone formation. The survival percentage was calculated as the number of colony-forming units (CFUs) (damaged) divided by the total number of CFUs (undamaged).

Homologous Recombination Assay

According to a previously reported method (Sheng et al., 2005), the recombination rate in *M. xanthus* was determined by measuring the probability of a resistance gene inserted into the genome through HR. The selected insertion site was located in the noncoding sequence between the MXAN_4466 and MXAN_4467 genes. Then, 500-bp fragments upstream and downstream of the insertion site were amplified with primers (UpF 5′-cacgggctacacgcaggtgcgggg-3′/UpR 5′-ttaagctttcgtttcagcggggactgcctgg-3′ and DownF 5′-caaagcttccaggcagtccccgctgaaacga-3′/DownR 5′-ggcatcgtccctggcggcggcgtgg-3′). The Kan resistance gene (*kanR*) with its promoter was simultaneously amplified from plasmid pZJY41 with primers 5′-gctgaagcttgtgctgaccccgggtgaat-3′/5′-agaagcttccagagtcccgctcagaagaac-3′. Then, the three DNA segments were linked by the *Hind*III site according to the arrangement of the upstream segment, resistance gene, and downstream segment. The linked DNA fragment was amplified with primers (UpF 5′-cacgggctacacgcaggtgcgggg-3′ and DownR 5′-ggcatcgtccctggcggcggcgtgg-3′) and quantitatively introduced into *M. xanthus* via electroporation (1.25 kV, 300 W, 50 mF, and 1 mm cuvette gap). A serial dilution was spread on CYE plates with or without Kan and incubated at 32°C for 72 h to count CFUs. The recombination ability was calculated by the following

formula: recombination efficiency (%) = (CFUs with Cam/CFUs without Cam) × 100.

Genetic Manipulations

Escherichia coli plasmids were isolated by the alkaline lysis method, and the chromosomal DNA of *E. coli* or *M. xanthus* was extracted using a bacterial genome DNA extraction kit (Tiangen, Beijing, China). Cloning of the genes *recA1*, *recA2*, and *lexA* from *M. xanthus* was performed according to the general steps (Ausubel et al., 1995). The genes were amplified by PCR and ligated into the pET15b expression plasmid. The primers used here are listed in **Supplementary Table S2**.

Mutant construction was performed using the markerless mutation in *M. xanthus* DK1622, with the pBJ113 plasmid using the Kan-resistant cassette for the first round of screening and the *galK* gene for the negative screening (Ueki et al., 1996). Briefly, the up- and downstream homologous arms were amplified with primers (listed in **Supplementary Table S2**) and ligated at the *Bam*HI site. The ligated fragment was inserted into the *Eco*RI/*Hind*III site of pBJ113. The resulting plasmid was introduced into *M. xanthus* via electroporation (1.25 kV, 300 W, 50 mF, and 1 mm cuvette gap). The second round of screening was performed on CYE plates containing 1% galactose (Sigma). The *recA1* (named RA1) and *recA2* (named RA2) mutants were identified and verified by PCR amplification and sequencing.

We attempted to construct the *recA1*/*recA2* double mutant from the single deletion mutant (RA1 or RA2) using the same procedure as described above, but all failed.

RNA Extraction, RT-PCR, and RNA-Seq Assay

Total RNA of *M. xanthus* cells was extracted using RNAiso Plus reagent following the manufacturer's protocol (Takara, Beijing, China). cDNA synthesis was performed using the PrimeScript RT Reagent Kit with random primers. The synthesized cDNA samples were diluted five times prior to RT-PCR. The primers were designed for *lexA*, *recA*1, and *recA*2 (**Supplementary Table S2**). RT-PCR was accomplished using the SYBR Premix Ex Taq Kit (Takara, China) on an ABI StepOnePlus Real-Time PCR System (ThermoFisher Scientific, United States). Gene expression was normalized to the *gapA* expression and calculated using the equation: change (*x*-fold) = $2^{-\Delta\Delta Ct}$ (Schefe et al., 2006).

RNA sequencing was conducted by Vazyme (Beijing, China). Three independent repeats are set for each sample. All the up- and downregulated genes were obtained by comparing the expression of the genes with that of the control, and their gene functions were annotated using the NR, GO, and KEGG databases.

Protein Expression, Purification, and Characterization

The constructed expression plasmids with *recA1*, *recA2*, or *lexA* were introduced into *E. coli* BL21(DE3) competent cells. Protein expression was induced with 1 mM IPTG and purified with Ni-NTA agarose according to the manual of the Ni-NTA purification system (Invitrogen). After overnight dialysis with storage buffer [20 mM Tris–HCl (pH 7.2), 150 mM NaCl, 0.1 mM DTT, 0.1 mM EDTA, and 50% glycerol], the purified proteins were quantified and stored at −80°C.

The ATPase activity of RecA protein was determined in the presence or absence of DNA according to the methods described previously (Sheng et al., 2005). The final reaction mixture in a 2-ml volume contained: 20 mM Tris–HCl (pH 7.4), 10 mM NaCl, 5 mM $MgCl_2$, 2 mM KCl, 3 mM ATP (Sigma), 1 mM $CaCl_2$, 1 mM DTT, and 2% glycerol. The mixture was preheated to 32°C before the addition of RecA and DNA. ATPase activity was determined by measuring the free phosphate ion (Pi) released from ATP using an ultramicro ATPase activity detection kit (Nanjing Jiancheng Bioengineering, Nanjing, China).

In vitro LexA cleavage analysis was performed as described previously (Sheng et al., 2005).

D-loop assays for strand assimilation were performed according to the previously described methods (Cloud et al., 2012; Huang et al., 2017) with some modifications. Briefly, 0.2 μM RecA and 10 nM ^{32}P-labelled ssDNA was combined in 9 μl of reaction buffer containing 25 mM Tris–HCl (pH 7.5), 75 mM NaCl, 5 mM $MgCl_2$, 3 mM ATP, 1 mM DTT, and 1 mM $CaCl_2$ and incubated at 37°C for 5 min. Then, 1 μl of RF M13 plasmid was added to a final concentration of 1 μM, and the incubation at 37°C was continued for 20 min. The reaction was stopped by adding sodium dodecyl sulfate to 0.5% and proteinase K to 1 mg/ml. The deproteinated reaction products were run on a 0.9% agarose 1× TAE gel and visualized using autoradiography with phosphor screen.

In vitro DNA strand-exchange reactions were performed as described previously (Sheng et al., 2005).

RESULTS

Duplicate *recA* Genes in *M. xanthus* Are Both Induced by UV Irradiation

The two RecA proteins of *M. xanthus* DK1622 are highly conserved and are both homologous to the RecA protein of *E. coli* K12 (EcRecA). The two RecA coding genes have high G+C contents (66 and 65%, respectively); the amino acid identity of RecA1 and RecA2 is 64.6%, and they are 59.36 and 62.04% to EcRecA, respectively. Similar to EcRecA (Story et al., 1992; Lee and Wang, 2009), RecA1 and RecA2 consist of three structural domains, a small N-terminal domain (NTD), a core ATPase domain (CAD), and a large C-terminal domain (CTD). CAD contains the conserved ATPase Walker A and Walker B domains and L1 and L2 DNA-binding domains (**Figure 1A**). The CAD of RecA1 and RecA2 are highly conserved, while the NTD and CTD are varied. Compared with EcRecA, the two RecA proteins of *M. xanthus* have more basic amino acids, and the theoretical isoelectric points [pI, calculated by online software (ExPASy – Compute pI/Mw tool)] of RecA1 and RecA2 are 7.04 and 6.5, respectively; EcRecA is more acidic, with a theoretical pI of 5.09 (**Figure 1B**). Differences in the amino acid composition suggested that the RecA1 and RecA2 proteins might vary in their functions.

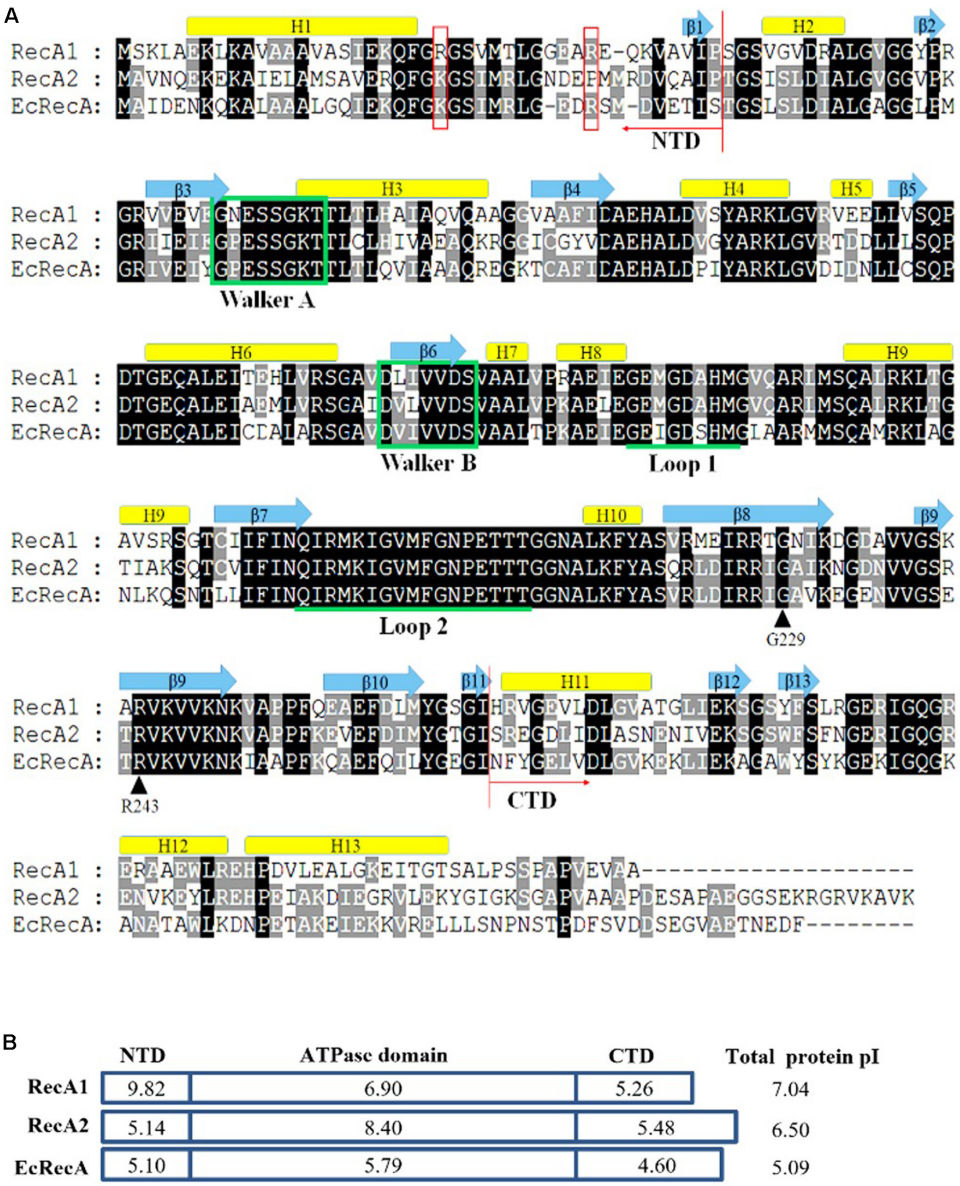


FIGURE 1 | Amino acid sequence comparison of RecA proteins. **(A)** Alignment of *M. xanthus* RecA1 and RecA2 and *E. coli* RecA (EcRecA, b2699). Positions of the N-terminus (NTD) and the C-terminus (CTD) domains are indicated with red arrows, respectively. Their secondary structures all contain 13 alpha-helixes and 13 beta-sheets, which are indicated above their corresponding amino acid sequences. The ATP binding Walker A and B motifs are marked in green frames, and the putative DNA binding sites Loop L1 and L2 are indicated by underlines of the corresponding amino acid sequences. Two reported LexA binding sites (G229 and R243) are indicated by black arrows. K23 and R33 in the N-terminal region of EcRecA are labeled with red boxes. **(B)** The pI features of the domains of the three RecA proteins. The theoretical pI values were computed using ExPASy online tools (Compute pI/Mw).

The SOS response of *M. xanthus* cells to DNA damage can be divided into LexA-dependent and -independent types (Campoy et al., 2003). The LexA-dependent SOS genes, e.g., *lexA*, typically possess a LexA-box sequence in their promoters. Each of the two *recA* genes of *M. xanthus* has its own promoter and is not a part of an operon. A typical LexA-box sequence was found in the promoter of *recA2* but not in the *recA1* promoter (**Figure 2A**). Previous studies reported that *recA2* was obviously induced by nalidixic acid and mitomycin C but that *recA1* was not induced by mitomycin C (Norioka et al., 1995; Campoy et al., 2003). We treated *M. xanthus* cells with 15 J/m^2 UV irradiation, which is also a normal induction agent for investigating the bacterial SOS response (Courcelle et al., 2001; Rastogi et al., 2010; Richa et al., 2015). RT-PCR revealed that *lexA* and *recA2* were upregulated by 8.3 times and 10.7 times, respectively, 4 h after UV irradiation at 15 J/m^2 (**Figure 2B**). Interestingly, the *recA1* gene was also UV-induced by 6.4 times. The basal expression level of *recA1* was very low and was less than one-tenth that of *recA2*. The low

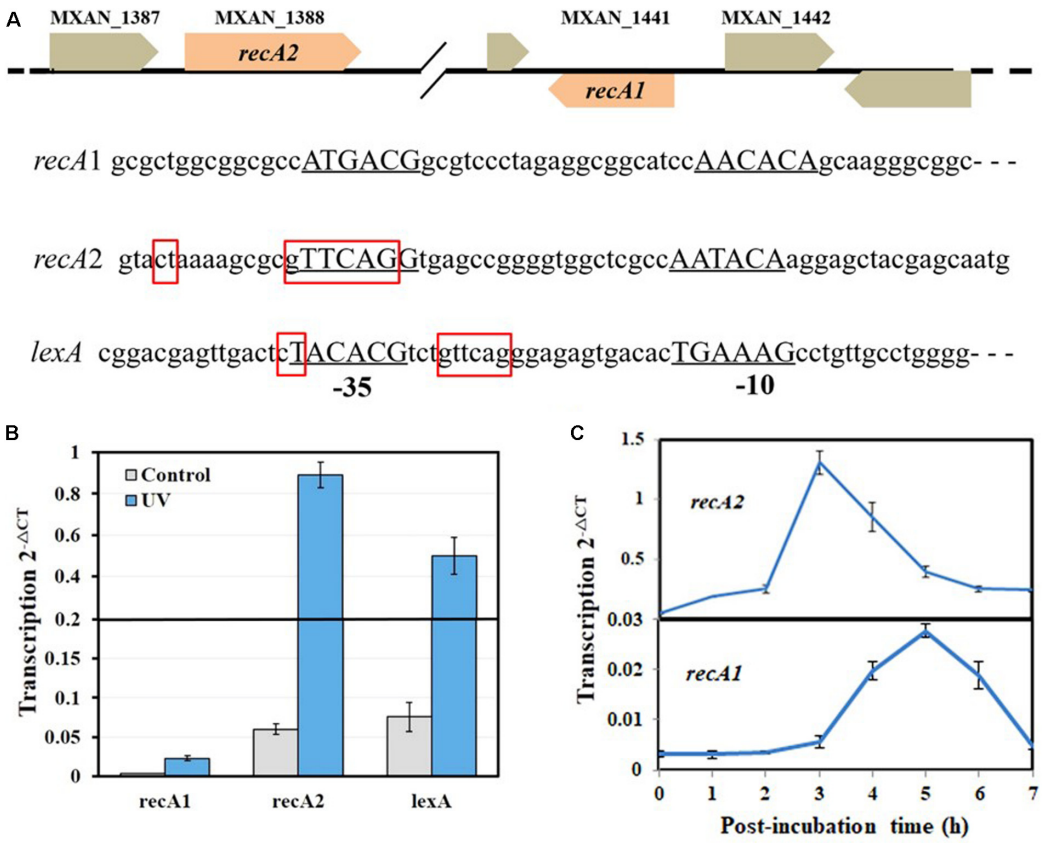


FIGURE 2 | Organization and UV inducibility of the *recA1* and *recA2* genes of *M. xanthus* DK1622. **(A)** Schematic gene location and promoter alignment of *M. xanthus recA1* and *recA2*. RNA polymerase binding sites (–10 and –35 regions) are underlined, and the corresponding nucleotide sequences are in capitals. The SOS box regions are framed in red squares, and the sequence in the promoter of the *lexA* gene (*MXAN_4446*) was used as a control. **(B)** UV inducibility of *recA1* and *recA2*. The strains were incubated for 4 h after UV irradiation treatment with a dose of 15 J/m^2 in a UV crosslinking machine and used to detect transcription of *recA1* and *recA2* by RT-PCR. *lexA* was used as a control. **(C)** The induction time points of *recA1* and *recA2*. After being exposed to UV irradiation at a dose of 15 J/m^2, the cell cultures were post-incubated at 30°C, sampled at certain intervals to extract the total RNA for RT-PCR. The error bars in panels **B** and **C** represent means ± SEM ($n = 3$, $p < 0.05$ versus inner reference).

expression level of *recA1* might be the reason why the expression of *recA1* was not detected by Northern blotting (Norioka et al., 1995). The generation time of *M. xanthus* cells is about 3–4 h in the exponential growth stage. We found that the induction of *recA2* peaked at approximately 3 h after UV treatment, whereas the induction time of *recA1* was delayed and peaked 5 h after the treatment (**Figure 2C**). The different expression levels and induction time points implied that the two RecA proteins participate in the repair of different types of DNA damage caused by UV irradiation.

Inactivation of *recA2* Compromises the Growth of *M. xanthus* Cells

In previous studies, *recA2* deletion mutants were not obtained in either *M. xanthus* or *B. megaterium* (Norioka et al., 1995; Campoy et al., 2003; Nahrstedt et al., 2005). However, in this study, we successfully obtained the deletion mutant of both *recA1* and *recA2* in *M. xanthus*, named RA1 and RA2, respectively (**Figure 3A**). According to the two-step screening method employed, the acquisition probability from reverse screening was $\sim 10^{-6}$ for the deletion of *recA1* and $\sim 3.3 \times 10^{-10}$ for the deletion of *recA2*, and this may be the reason why it is difficult to make a *recA2* mutant. At present, there is no evidence of a suppressor mutation, which may be required to achieve the deletion of *recA2*. However, although more evidence is needed, we speculate that the difficulty in the screening of RA2 is probably related to the function of RecA2 in growth. We also tried to construct the double knockout mutant of *recA1* and *recA2* but failed, and this might be because the double mutation had a serious impact on cell survival and was synthetically lethal. *recA1* deletion had no significant effects on cellular growth, but deletion of *recA2* caused the mutant to have a long lag phase. After the lag phase, growth of the RA2 mutant did not slow down significantly in the logarithmic phase, and the mutant culture reached a similar density as wild-type DK1622 (**Figures 3B,C**).

When treated with 15 J/m^2 UV irradiation, compared with those without UV treatment, the growth abilities were delayed in DK1622, RA1, and RA2 cells, and the growth delay was more notable in RA2 (**Figure 3B**). When treated with 3 mM

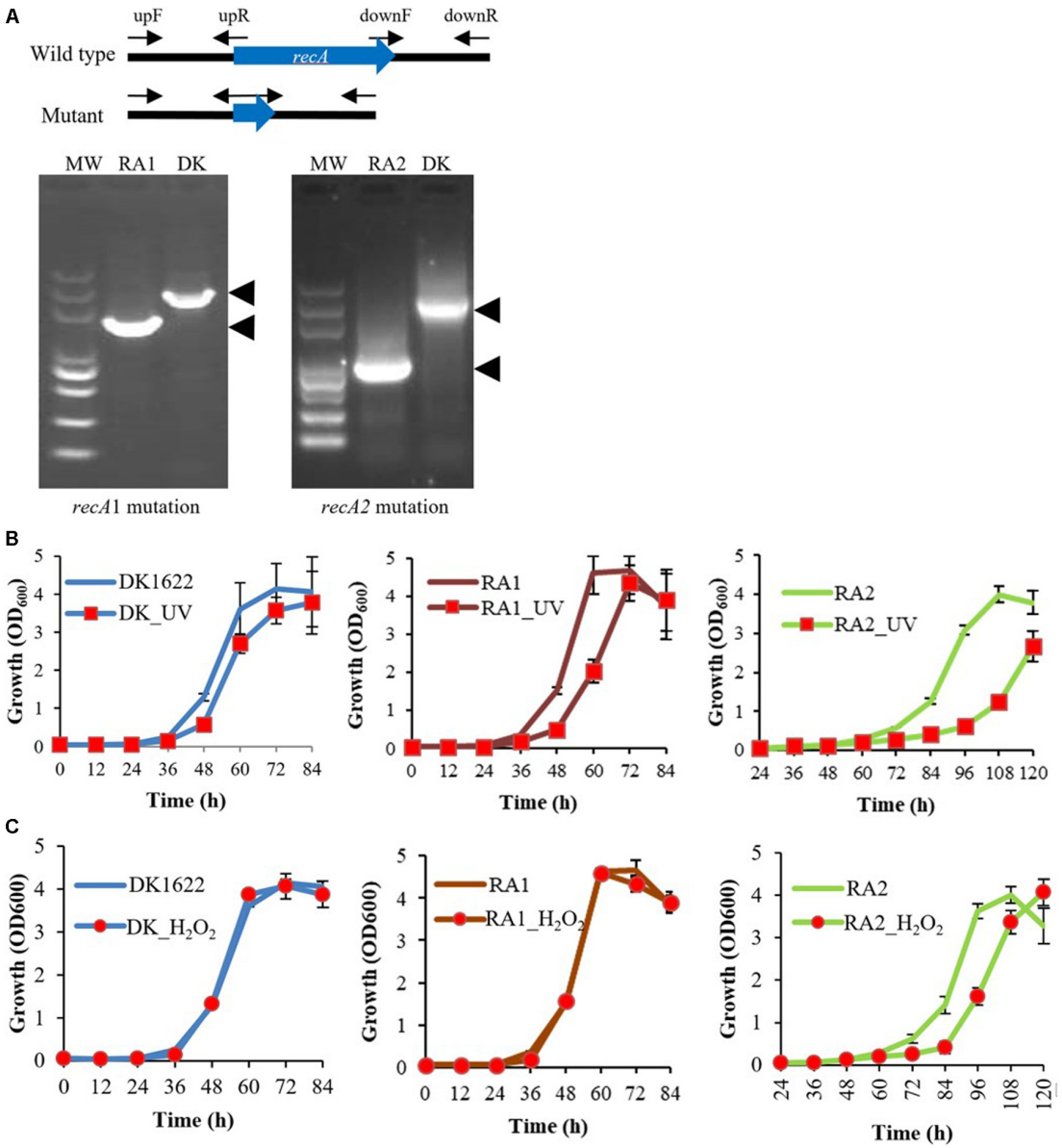


FIGURE 3 | Mutations of *recA1* and *recA2*, and their effects on the growth of *M. xanthus*. **(A)** Deletion of *recA1* or *recA2* in *M. xanthus* DK1622, using the markerless knockout plasmid pBJ113, producing the RA1 or RA2 mutants. The deletion was verified by PCR using their primer pairs (upF/downR) and sequencing. **(B)** Separate growth comparisons of DK1622, RA1, and RA2 with and without UV treatment at a dose of 15 J/m^2. **(C)** Separate growth comparisons of DK1622, RA1, and RA2 with and without the H_2O_2 treatment at a final concentration of 3 mM for 15 min. The error bars indicate the SEM for six replicates.

H_2O_2 for 15 min, DK1622 and RA1 cells showed almost the same growth curve, while the growth of RA2 cells was delayed significantly compared with that of the strains without the treatment (**Figure 3C**). The results demonstrated that *recA2*, but not *recA1*, is an important factor for cell growth after UV irradiation and oxidation damage.

recA1 and *recA2* Are Separately Crucial for UV Resistance and H_2O_2 Resistance

We measured the survival rates of the wild-type strain and the *recA* deletion mutants treated with different dosages of UV irradiation (0–25 J/m^2) and H_2O_2 (1–5 mM). All three strains had decreased survival rates with increasing UV irradiation or H_2O_2 concentration. Interestingly, the survival rate of RA1 cells decreased more significantly than that of RA2 at each UV-radiation dosage, which had a highly similar survival curve to the wild-type strain (**Figure 4A**). In addition, the survival rate of RA2 cells decreased more significantly at each H_2O_2 concentration than that of RA1 and DK1622 cells, which showed similar survival curves when treated with hydrogen peroxide (**Figure 4B**). Thus, RecA1 is needed for the survival of *M. xanthus* cells under UV irradiation, which is similar to that of EcRecA

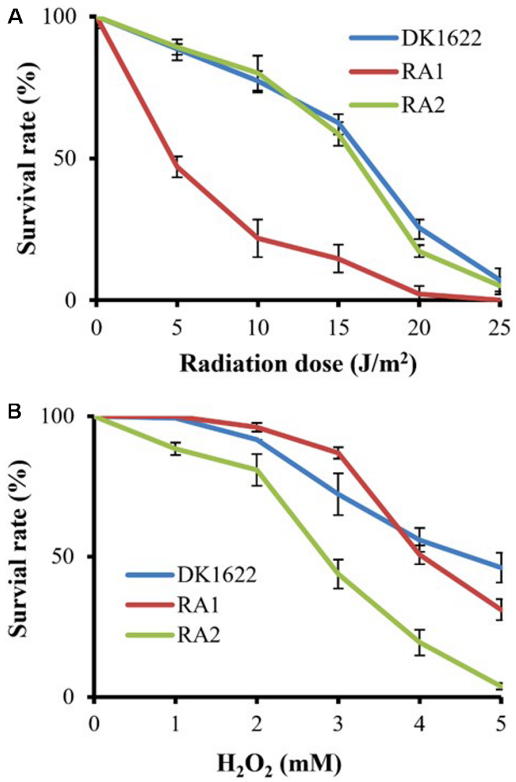


FIGURE 4 | Survival of *M. xanthus* wild-type strain DK1622 and the mutants RA1 and RA2. **(A)** Survival curves after exposure to UV irradiation at different dosages (0–25 J/m^2). **(B)** Survival curves after hydrogen peroxide treatment at different concentrations (0–5 mM). The percentage of surviving cells was calculated by comparing with the corresponding non-treated cells. The error bars indicate the SEM for six replicates.

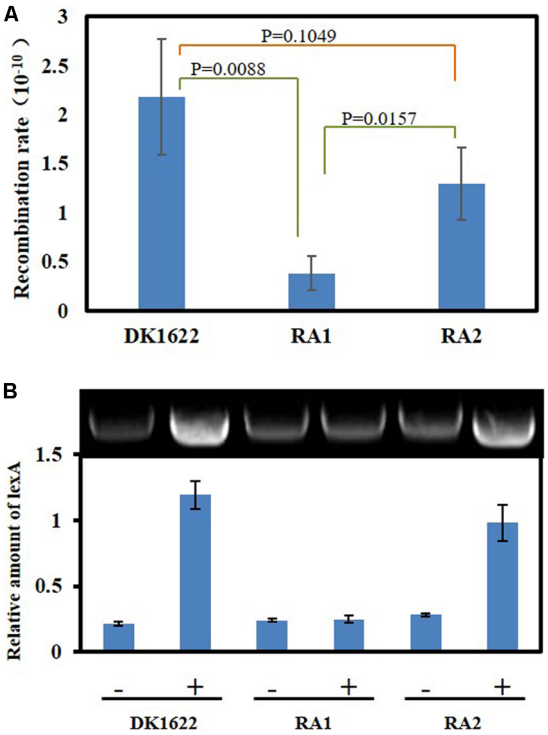


FIGURE 5 | Intracellular DNA recombination rate and induction analysis of the *lexA* gene. **(A)** The cellular recombination rate of DK1622, RA1, and RA2. The two ends of the resistance gene (*kan*) have homologous DNA sequences with the two ends of the insertion site, respectively, and the *kan* gene is transferred into the *M. xanthus* through electrical transformation. The recombination rate was calculated by measuring the proportion of kanamycin-resistant colonies. **(B)** The inducibility of the *lexA* gene. *Myxococcus lexA* is a known SOS gene induced through LexA-dependent SOS response, and herein, its UV inducibility represents the activation of LexA-dependent SOS response. The strains were irradiated with 15 J/m^2 UV irradiation (+) or mock treatment (–), and the transcription of *lexA* was determined by RT-PCR.

(Alexseyev et al., 1996), whereas RecA2 is involved in tolerance to H_2O_2 damage in cells.

RecA1, Not RecA2, Is Responsible for HR and LexA-Dependent SOS Induction

DNA HR and SOS induction are the two main cellular functions of the RecA proteins (Cox, 2003). We analyzed the *in vivo* integration abilities of an antibiotic resistance gene into the genomes of the DK1622, RA1, and RA2 strains. Calculated from the appearance of resistant colonies, the recombination rate of RA1 cells was significantly lower than that in either DK1622 (p = 0.0088) or RA2 (p = 0.0157) cells, while the differences between the recombination rates of RA2 and DK1622 cells were not significant (p = 0.1049) (**Figure 5A**). The results showed that *recA1* is important for the recombination process in *M. xanthus*.

Previous studies indicated that the expression of *lexA* is induced by the LexA-dependent SOS response in *M. xanthus* (Norioka et al., 1995). We compared the transcription of *lexA* in the *M. xanthus* DK1622, RA1, and RA2 strains in response to the 15 J/m^2 UV irradiation treatment. The RT-PCR results showed that *lexA* could be induced by UV in both DK1622 and RA2 but not in the RA1 mutant (**Figure 5B**). Thus, the deletion of *recA1*, rather than *recA2*, affected the UV-induction of *lexA*, i.e., RecA1 is responsible for LexA-dependent SOS induction.

RecA1 and RecA2 Both Have ss- and ds-DNA Promoted ATPase Activities

We further expressed and purified the RecA1 and RecA2 proteins (**Figure 6A**) and measured their *in vitro* ATPase activities by the quantitative analysis of inorganic phosphorus released from ATP hydrolysis (**Figure 6B**). In the reaction mixture without the addition of DNA, RecA1 and RecA2 both exhibited low ATPase activities, and the ATPase activity of RecA2 was somewhat higher than that of RecA1. For example, a microgram of purified RecA2 released 0.1428 nanomole Pi in an hour, which is approximately 2.44 times the hydrolysis capacity of RecA1 on ATP (0.0586 nmol Pi/μg/h). The addition of DNA, especially ssDNA, markedly promoted the ATPase activity of both RecA1 and RecA2, and this is consistent with the functionality of classic RecA proteins (Cox, 2003; Greene, 2016). Thus, RecA1 and RecA2 are both DNA-dependent (more dependent on ssDNA) ATPase enzymes.

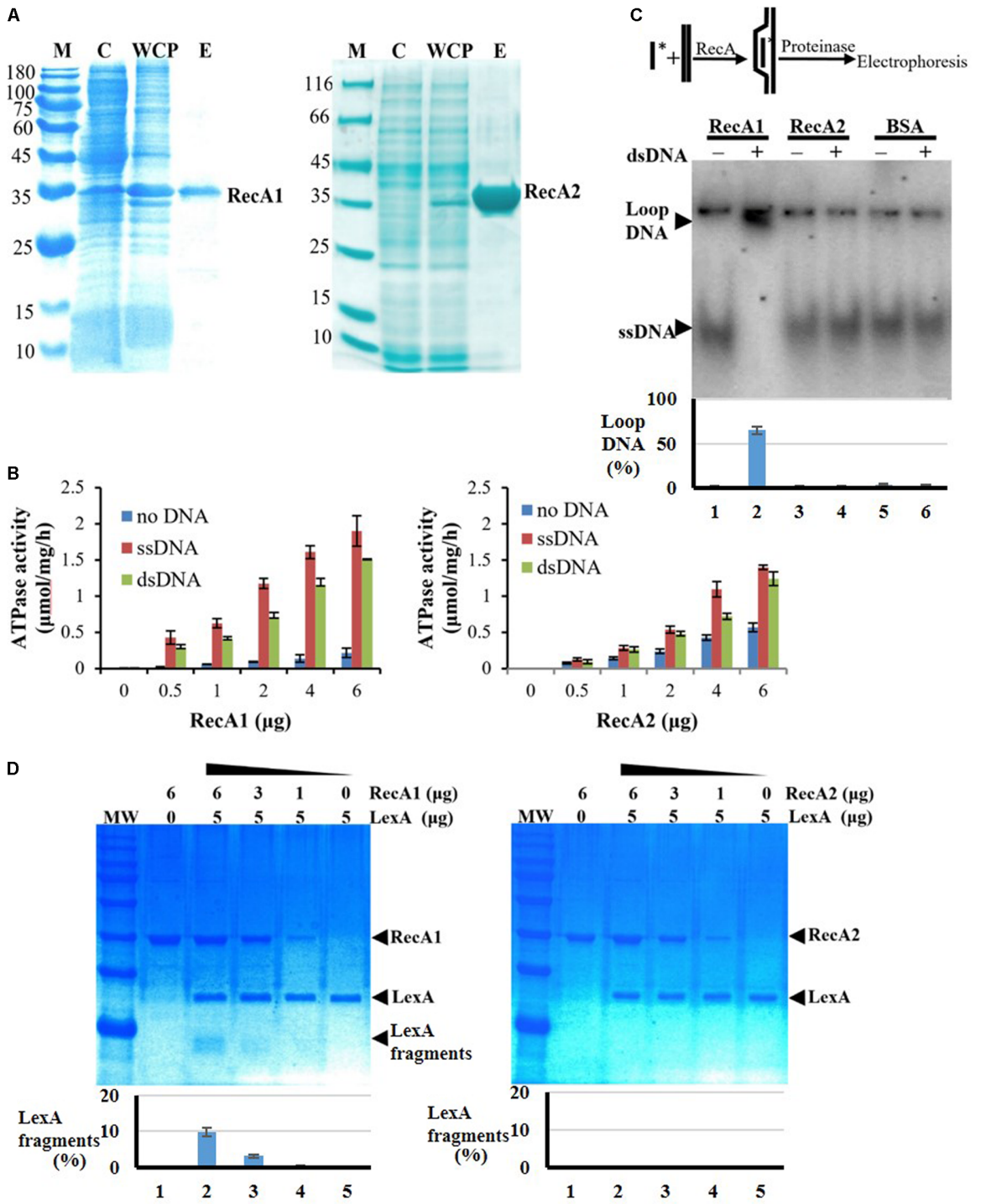


FIGURE 6 | Expression and activity analysis of RecA proteins. **(A)** Expression and purification of RecA1 and RecA2. M, marker; C, control; WCP, whole-cell protein; E, eluent of purified protein. **(B)** Assays of ATPase activities. The ATPase activity was determined by measuring free phosphate ion (Pi) released from enzymolysis of ATP. The error bar is calculated from three independent repeats. **(C)** D-loop assay. A 60-nt ^{32}P-labeled ssDNA fragment and a superhelical dsDNA (RF M13) sequence were mixed and incubated with and without the addition of purified RecA1 or RecA2 proteins. If the protein has HR activity, the homologous pairing reaction will be initiated, thus forming the ssDNA–dsDNA complex. Bovine serum albumin (BSA) was used as a control. Relative DNA-labeled intensities of the bands were quantified by a Gel-Doc image analysis system (Bio-Rad). The percentage of loop-DNA-labeled intensity in the total labeled intensity (including the labeled strength in ssDNA, loop-DNA, and residual DNA in the origin) was used to quantitate the RecA activity. **(D)** The promotion ability of RecA1 (left) or RecA2 (right) on the cleavage of LexA proteins. The MxLexA protein was incubated with gradient concentrations of RecA1 or RecA2 proteins in the presence of ssDNA and ATP. Reactions were stopped and visualized on a 1.2% SDS-PAGE gel stained with Coomassie brilliant blue. The bands were quantified by computerized image analysis (Bio-Rad), and the percentage of LexA fragments in the total LexA signal in every lane was used to quantify the ability of RecA to stimulate LexA autocleavage.

In the presence of DNA (dsDNA or ssDNA), the increase in the ATPase activity of RecA1 was higher than that of RecA2. For example, the ATPase activity of 1 ng RecA1 increased by 10.69 times (from 0.0586 to 0.6265 nmol Pi/μg/h) with the addition of ssDNA, while the increase of that in RecA2 was only double (from 0.1428 to 0.2857 nmol Pi/μg/h). Similarly, the addition of dsDNA increased the ATPase activities of RecA1 and RecA2 by 6.89 times (from 0.0586 to 0.4038 nmol Pi/μg/h) and 1.86 times (from 0.1428 to 0.2658 nmol Pi/μg/h), respectively.

RecA1, but Not RecA2, Has *in vitro* HR Capacity and Activates LexA Autolysis

Strand invasion or D-loop formation is a central step in HR and is one of the most common biochemical assays for characterizing the activity of RecA-type recombinase (Cox, 2003; Greene, 2016; Huang et al., 2017). We analyzed the *in vitro* recombination activities of RecA1 and RecA2 in a DNA strand recombination reaction system containing ^{32}P-ssDNA and homologous plasmid dsDNA. The reaction products were separated by agarose gel electrophoresis, and a lagged radiolabeling band appeared in the lane containing purified RecA1 but that with not RecA2 (**Figure 6C**). Furthermore, DNA strand exchange (DSE), another characteristic reaction for RecA-catalyzed HR with single-stranded circular DNA and its homologous double-stranded linear DNA, was used to identify the recombination activity of RecA1 and RecA2 (**Supplementary Figure S5**). The joined molecule DNA (jmDNA) appeared and increased as a recombinant product with the gradient addition of the RecA1 protein but not with that of the RecA2 protein. The above results indicated that RecA1, but not RecA2, has HR activity in *M. xanthus*, and these results are consistent with the *in vivo* recombination results (**Figure 5A**).

RecA promotes LexA autolysis at a specific site, thereby enabling the expression of SOS genes inhibited by LexA (Janion, 2008; Kovačič et al., 2013). We monitored the LexA cleavage activity promoted by RecA1 and RecA2, using the *M. xanthus* LexA protein as a substrate. The results showed that the LexA autolysis fragments were detected in the reaction with RecA1 but not that with RecA2 (**Figure 6D**). Thus, RecA1 participated in the LexA-dependent SOS induction reaction, and this is also consistent with the RA1 mutant losing the induction ability of the SOS gene *lexA* (**Figure 5B**).

DISCUSSION

RecA is an ATP-dependent recombinase central to DNA HR and activation of the LexA-dependent SOS response. Although the *recA* gene is duplicated in some bacterial cells, its functions have not been investigated. In *M. xanthus* DK1622, the expression of *recA1* is very low and is less than one-tenth of that of *recA2*. The two *recA* genes are both inducible by UV irradiation, but the induction of *recA2* was significantly earlier than that of *recA1* and recN (**Supplementary Figure S6**). Generally, the DNA repair genes expressed in the early and late stages of SOS are responsible for the error−free repair and maintenance processes and error-prone DNA synthesis against serious DNA damage, respectively (Kuzminov, 1999; Janion, 2008; Maslowska et al., 2019). Thus, the two RecA proteins are both involved in UV resistance, probably for different lesions caused by UV irradiation (Sinha and Häder, 2002); RecA2 is involved in the early repair processes, and RecA1 is involved in serious DNA-damage repair, i.e., post-replication repair. The deletion of *recA2* caused the mutant to have a long lag phase, but the *recA1* deletion had no significant effect on cellular growth. It is known that the growth lag phase is an adaptation period of bacterial cells to changes in temperature and nutrients in new environments (Monod, 1949; Vermeersch et al., 2019), macromolecule damage repair, and protein misfolding accumulated during cell arrest (Dukan and Nyström, 1998; Saint-Ruf et al., 2007; Rolfe et al., 2012; Bertrand, 2019), and enzyme preparation for rapid growth in the logarithmic phase (Monod, 1949; Rolfe et al., 2012). Thus, RecA2, instead of RecA1, plays a crucial role in the repair process required for cellular growth. Similar to the classic bacterial RecA, RecA1 possesses DNA recombination activity and SOS gene induction ability, which are required for survival under UV irradiation. However, RecA2 has lost its HR and SOS gene induction abilities but has evolved to play roles in the regulation of gene expression for cellular growth as well as cellular survival under oxidation pressure by hydrogen peroxide. This is the first study to clearly determine the divergent functions of duplicated *recA* genes in bacterial cells.

To obtain more clues about the potential mechanisms of RecA2 in *M. xanthus*, we compared the transcriptomes of the *recA2* mutant strain (RA2) and wild-type strain DK1622. Overall, 79 genes were found to be differentially expressed ($p_{adj.} < 0.05$) by the deletion of *recA2*, and this included 60 upregulated genes and 19 downregulated genes (**Figure 7A**; for details, refer to **Supplementary Table S3**). Gene ontology (GO) enrichment analysis based on the KEGG database showed that the differentially expressed genes (DEGs) were assigned to 30 GO terms in the categories of biological process, cellular component, and molecular function (**Figure 7B**). Obviously, the biological process DEGs formed the largest group and included 17 GO terms, followed by the molecular function (10 GO terms) and cellular component (three GO terms) groups (**Figure 7B**). The DEGs were mainly enriched in two functional regions. One is related to transport and location, including the categories of transport (14 genes), localization and establishment of localization (28 genes), transmembrane transport (five genes), and protein transmembrane transport (three genes). The other category is related to antioxidation and includes the categories of oxidoreductase activity (three genes), peroxiredoxin activity (two genes), ferric iron binding (two genes), antioxidant capacity (three genes), and catalase (one gene). These DEGs were significantly enriched in ABC transporters and several metabolism-related pathways, such as methane metabolism, biosynthesis of secondary metabolites, and metabolic pathways (**Figure 7C**). Combining this with the experimental results presented in this study, we propose that the function of *recA2* is mainly related to cellular transportation and antioxidation, which are required for the normal growth of cells.

RecA1 and RecA2 are both homologous proteins with *E. coli* RecA. They both retain the DNA-dependent ATPase activity and self-aggregation ability of *E. coli* RecA (**Figure 6B** and

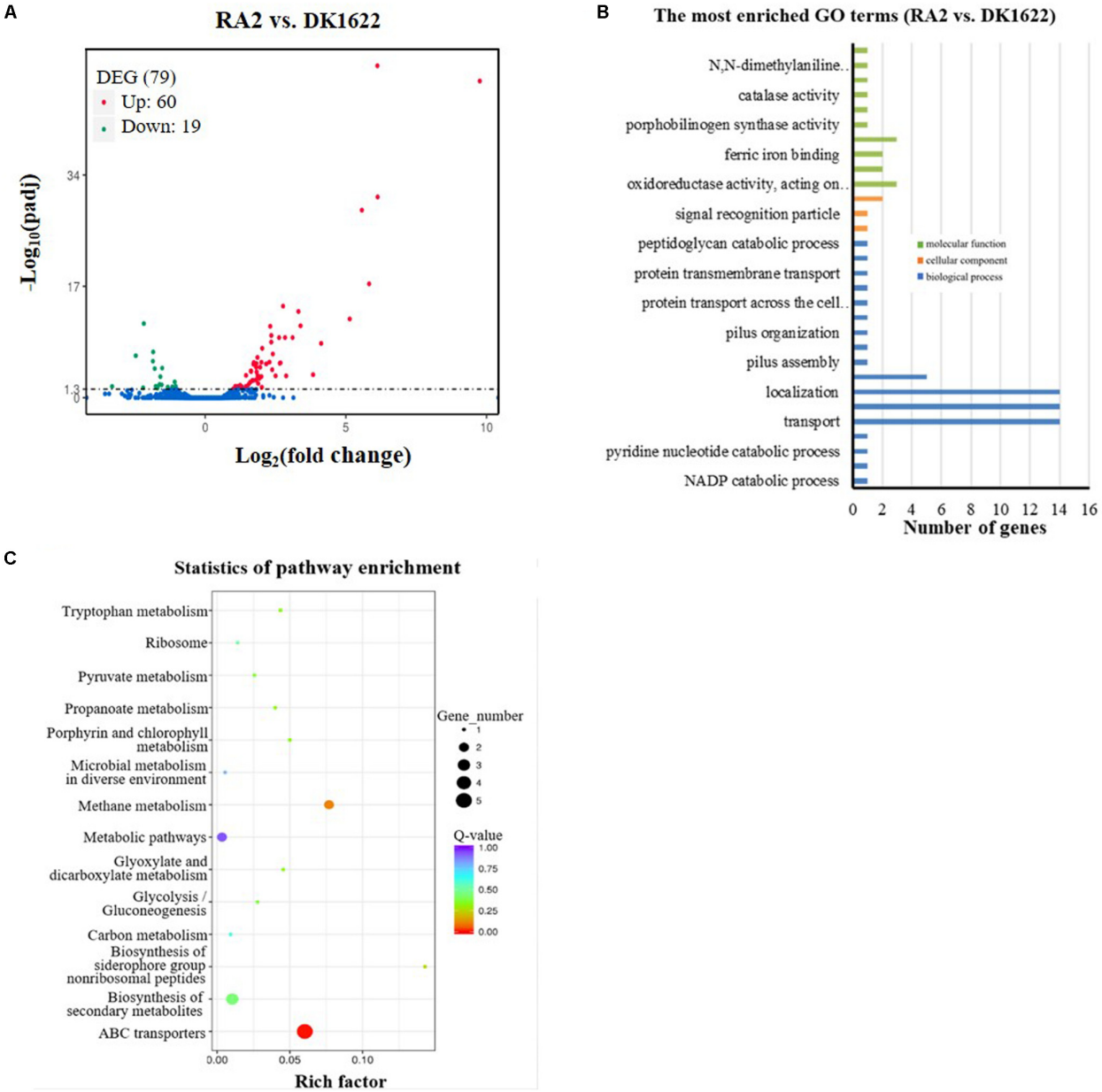


FIGURE 7 | Comparison of transcriptomes between the *recA2* mutant (RA2) and the wild-type strain (DK1622). **(A)** Volcano plot of differentially expressed gene (DEG) distributions. Red dots and green dots represent the up- and down-regulated genes with significant differences, respectively ($p_{adj.} < 0.05$). The blue dots represent the genes that have not changed significantly. **(B)** The distribution of GO category and **(C)** pathways of DGEs between RA2 and DK1622. Enriched GO is shown in three categories: biological process (blue), molecular function (green), and cellular component (orange).

Supplementary Figure S7), and the interaction between RecA1 and RecA2 may even be due to the conserved self-aggregation sites. On the other hand, these two proteins also showed significantly different activities in DNA recombination and LexA autocleavage. Amino acid sequence alignment showed that the RecA1 and RecA2 amino acid sequences are highly similar in the CAD, and are mainly varied in the NTD and CTD (**Figure 1A**). Lys23 and Arg33 in the N-terminal region are both necessary for the nucleoprotein filament of RecA–ssDNA to capture homologous dsDNA (Lee and Wang, 2009). The corresponding amino acids at the two sites are both alkaline arginine residues in RecA1, which is consistent with that in EcRecA. In RecA2, however, the amino acids at the two sites are arginine and proline, respectively (**Figure 1A**). We aligned the N-terminal sequences of 11 reported bacterial RecA proteins. The amino acids at the corresponding 23rd site are all alkaline amino acids but are less conserved at the 33rd site (**Supplementary Figure S1**). Nine RecAs, including RecA1 of *M. xanthus*, have

positively charged residues (Arg or Lys) at the 33rd site, while three RecAs, including RecA2 from *M. xanthus*, RecA from *Prevotella ruminicola* (Aminov et al., 1996), and RecA from *Borrelia burgdorferi* (Huang et al., 2017), did not have the positively charged residues at this site. RecAs with an alkaline amino acid at the 33rd site all have DNA recombination activity (Sano and Kageyama, 1987; Nussbaumer and Wohlleben, 1994; Umelo et al., 1996; Cox, 1999; Kim et al., 2002; Orillard et al., 2011; Grove et al., 2012; Carrasco et al., 2019). However, similar to RecA2, RecAs from *P. ruminicola* and *B. burgdorferi* were reported to have no anti-ultraviolet radiation ability (Aminov et al., 1996; Huang et al., 2017). The results presented in this study demonstrate that RecA2 of *M. xanthus* evolved to affect the genes for cellular transportation and antioxidation, which are obviously related to damage repair for cellular growth.

As in *E. coli* RecA (Lusetti et al., 2003; Kovačič et al., 2013), RecA 1 and RecA2 both have conserved LexA binding sites, including Gly229 and Arg243, in their C-terminal regions and 10 neighboring amino acids (**Figure 1A**), which, however, does not explain the differences between the two proteins in promoting LexA autolysis. Notably, while the three domains of EcRecA are all acidic, the NTD of RecA1 and the CAD of RecA2 are alkaline, with pI values of 9.82 and 8.40, respectively. Accordingly, RecA1 forms more negative charges on the outer side of the polymer, while RecA2 forms more negative charges on the inner side of the helical structure (**Supplementary Figure S2**). We noted that, unlike the *E. coli* LexA (EcLexA) protein, which is an acidic protein [theoretical pI is 6.23, calculated by online software (ExPASy Compute pI/Mw tool)], *M. xanthus* LexA (MxLexA) is a basic protein, and its theoretical pI is 8.77. EcLexA and MxLexA are conserved in their catalytic and DNA binding domains, and the differences between the two proteins lie mainly in the linker region (**Supplementary Figure S3**). The EcLexA linker contains more acidic amino acids (theoretical pI = 3.58), while the linker of MxLexA contains more basic amino acids (theoretical pI = 8.75). In addition, MxLexA has two more fragments flanking the linker sequence. The additional fragment at the N-terminal side destroys the β2 folding structure and further lengthens the irregular linker of MxLexA, leading to a long irregular chain containing more basic amino acids (theoretical pI = 12.01). According to the binding mode between EcLexA and EcRecA (Kovačič et al., 2013), the linker region of LexA is close to the inner groove of the RecA protein filament (**Supplementary Figure S4**). The inner helix part of RecA2 (in the CAD) is alkaline (**Supplementary Figure S2B**), which hinders MxLexA binding to the RecA filament and thus hinders its ability to promote MxLexA self-cleavage.

Myxobacteria has a relatively large genome size (9–14 Mbp) and contains many DNA repeats (Goldman et al., 2006; Andersson and Hughes, 2009; Han et al., 2013). These repetitive DNA fragments are potential substrates for RecA-catalyzed HR. Functional divergence of duplicate RecA proteins and low expression of the recombination enzyme RecA1 reduce the DNA recombinant activity without affecting other cellular repair functions in *M. xanthus* (such as the functions carried out by RecA2). In the sequenced myxobacterial genomes (**Supplementary Table S4**), all the strains, except *Anaeromyxobacter*, harbor two *recA* genes, and their amino acid sequences are highly conserved. For example, the amino acid identities of RecA1 and RecA2 of all *Myxococcus* are >89.4 and 93.6%, respectively. The *Anaeromyxobacter* strains have a single *recA* gene in their genomes; however, their genomes are small (5.0–5.2 Mbp) and possess few repetitive sequences. We propose that the functional divergence and expression regulation of duplicate RecA proteins might be a strategy for bacteria with a large number of repetitive sequences in their large genomes to avoid incorrect recombination.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

D-HS and Y-ZL designed the experiments. D-HS, Y-XW, MQ, and J-YZ performed the experiments. D-HS, X-JY, and Y-ZL analyzed the data. D-HS and Y-ZL wrote the manuscript.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (NSFC) (Grant Nos. 31670076 and 31471183), the National Key Research and Development Program (2018YFA0901700), and the Key Program of Shandong Natural Science Foundation (Grant No. ZR2016QZ002) to Y-ZL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This manuscript has been released as a Pre-Print at bioRxiv.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00140/full#supplementary-material

REFERENCES

Alexseyev, A. A., Bakhlanova, I. V., Zaitsev, E. N., and Lanzov, V. A. (1996). Genetic characteristics of new *recA* mutants of *Escherichia coli* K-12. *J. Bacteriol.* 178, 2018–2024. doi: 10.1128/jb.178.7.2018-2024.1996

Aminov, R. I., Nagamine, T., Ogata, K., Sugiura, M., Tajima, K., and Benno, Y. (1996). Cloning, sequencing and complementation analysis of the *recA* gene from *Prevotella ruminicola*. *FEMS Microbiol. Lett.* 144, 53–59.

Andersson, D. I., and Hughes, D. (2009). Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.* 43, 167–195. doi: 10.1146/annurev-genet-102108-134805

Ausubel, F. M., Brent, R., and Kingston, R. E. (1995). *Short Protocals in Molecular Biology*, 3rd Edn. New York, NY: Cold Spring Harbor Laboratory Press.

Bertrand, R. L. (2019). Lag phase is a dynamic, organized, adaptive, and evolvable period that prepares bacteria for cell division. *J. Bacteriol.* 201:e00697-618. doi: 10.1128/JB.00697-18

Bichara, M., Meier, M., Wagner, J., Cordonnier, A., and Lambert, I. B. (2011). Postreplication repair mechanisms in the presence of DNA adducts in *Escherichia coli*. *Mutat. Res* 727, 104–122. doi: 10.1016/j.mrrev.2011.04.003

Bretscher, A. P., and Kaiser, D. (1978). Nutrition of *Myxococcus xanthus*, a fruiting myxobacterium. *J. Bacteriol.* 133, 763–768. doi: 10.1128/jb.133.2.763-768.1978

Campoy, S., Fontes, M., Padmanabhan, S., Cortés, P., Llagostera, M., and Barbé, J. (2003). LexA-independent DNA damage-mediated induction of gene expression in *Myxococcus xanthus*. *Mol. Microbiol.* 49, 769–781. doi: 10.1046/j.1365-2958.2003.03592.x

Carr, A. M., and Lambert, S. (2013). Replication stress-induced genome instability: the dark side of replication maintenance by homologous recombination. *J. Mol. Biol.* 425, 4733–4744. doi: 10.1016/j.jmb.2013.04.023

Carrasco, B., Serrano, E., Martín-González, A., Moreno-Herrero, F., and Alonso, J. C. (2019). *Bacillus subtilis* MutS modulates RecA-mediated DNA strand exchange between divergent dna sequences. *Front Microbiol.* 10:237. doi: 10.3389/fmicb.2019.00237

Cloud, V., Chan, Y. L., Grubb, J., Budke, B., and Bishop, D. K. (2012). Rad51 is an accessory factor for Dmc1-mediated joint molecule formation during meiosis. *Science* 337, 1222–1225. doi: 10.1126/science.1219379

Courcelle, J., Khodursky, A., Peter, B., Brown, P. O., and Hanawalt, P. C. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* 58, 41–64.

Cox, M. M. (1999). Recombinational DNA repair in bacteria and the RecA protein. *Prog. Nucleic. Acid. Res. Mol. Biol.* 63, 311–366. doi: 10.1016/s0079-6603(08)60726-6

Cox, M. M. (2003). The bacterial RecA protein as a motor protein. *Annu Rev Microbiol.* 57, 551–577. doi: 10.1146/annurev.micro.57.030502.090953

Cox, M. M. (2007). Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* 42, 41–63. doi: 10.1080/10409230701260258

Dukan, S., and Nyström, T. (1998). Bacterial senescence: stasis results in increased and differential oxidation of cytoplasmic proteins leading to developmental induction of the heat shock regulon. *Genes Dev.* 12, 3431–3441. doi: 10.1101/gad.12.21.3431

García-Solache, M., Lebreton, F., McLaughlin, R. E., Whiteaker, J. D., Gilmore, M. S., and Rice, L. B. (2016). Homologous recombination within large chromosomal regions facilitates acquisition of β-lactam and vancomycin resistance in *Enterococcus faecium*. *Antimicrob. Agents Chemother.* 60, 5777–5786. doi: 10.1128/aac.00488-16

Goldman, B. S., Nierman, W. C., Kaiser, D., Slater, S. C., Durkin, A. S., Eisen, J. A., et al. (2006). Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15200–15205. doi: 10.1073/pnas.0607335103

Greene, E. C. (2016). DNA sequence alignment during homologous recombination. *J. Biol. Chem.* 291, 11572–11580. doi: 10.1074/jbc.R116.724807

Grove, D. E., Anne, G., Hedayati, M. A., and Bryant, F. R. (2012). Stimulation of the *Streptococcus pneumoniae* RecA protein-promoted three-strand exchange reaction by the competence-specific SsbB protein. *Biochem. Biophys. Res. Commun.* 424, 40–44. doi: 10.1016/j.bbrc.2012.06.059

Han, K., Li, Z. F., Peng, R., Zhu, L. P., Zhou, T., Wang, L. G., et al. (2013). Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci. Rep.* 3:2101. doi: 10.1038/srep02101

He, S., Chandler, M., Varani, A. M., Hickman, A. B., Dekker, J. P., and Dyda, F. (2016). Mechanisms of evolution in high-consequence drug resistance plasmids. . *mBio* 7:e01987-16.. doi: 10.1128/mBio.01987-16

Herrero-Fresno, A., Rodicio, R., Montero, I., Garcaí, P., and Rodicio, M. R. (2015). Transposition and homologous recombination drive evolution of pUO-StVR2, a multidrug resistance derivative of pSLT, the virulence plasmid specific of *Salmonella enterica* serovar Typhimurium. *Infect Genet Evol.* 29, 99–102. doi: 10.1016/j.meegid.2014.11.010

Huang, S. H., Hart, M. A., Wade, M., Cozart, M. R., McGrath, S. L., and Kobryn, K. (2017). Biochemical characterization of *Borrelia burgdorferi*'s RecA protein. *PLoS One* 12:e0187382. doi: 10.1371/journal.pone.0187382

Janion, C. (2008). Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int. J. Biol. Sci.* 4, 338–344. doi: 10.7150/ijbs.4.338

Jaszczur, M. M., Vo, D. D., Stanciauskas, R., Bertram, J. G., Sikand, A., Cox, M. M., et al. (2019). Conformational regulation of *Escherichia coli* DNA polymerase V by RecA and ATP. *PLoS Genet.* 15:e1007956. doi: 10.1371/journal.pgen.1007956

Kim, J. I., Sharma, A. K., Abbott, S. N., Wood, E. A., Dwyer, D. W., Jambura, A., et al. (2002). RecA Protein from the extremely radioresistant bacterium *Deinococcus radiodurans*: expression, purification, and characterization. *J Bacteriol.* 184, 1649–1660. doi: 10.1128/jb.184.6.1649-1660.2002

Kovačič, L., Paulič, N., Leonardi, A., Hodnik, V., Anderluh, G., Podlesek, Z., et al. (2013). Structural insight into LexA-RecA* interaction. *Nucleic Acids Res.* 41, 9901–9910. doi: 10.1093/nar/gkt744

Kuzminov, A. (1999). Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* 63, 751–813. doi: 10.1128/mmbr.63.4.751-813.1999

Lawrence, J. G., and Retchless, A. C. (2009). The interplay of homologous recombination and horizontal gene transfer in bacterial speciation. *Methods Mol Biol.* 532, 29–53. doi: 10.1007/978-1-60327-853-9_3

Lee, C. D., and Wang, T. F. (2009). The N-terminal domain of *Escherichia coli* RecA have multiple functions in promoting homologous recombination. *J. Biomed. Sci.* 16:37. doi: 10.1186/1423-0127-16-37

Lusetti, S. L., and Cox, M. M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu. Rev. Biochem.* 71, 71–100. doi: 10.1146/annurev.biochem.71.083101.133940

Lusetti, S. L., Wood, E. A., Fleming, C. D., Modica, M. J., Korth, J., Abbott, L., et al. (2003). C-terminal deletions of the Escherichia coli RecA protein. Characterization of in vivo and in vitro effects. *J. Biol. Chem.* 278:1637216380.

Maslowska, K. H., Makiela-Dzbenska, K., and Fijalkowska, I. J. (2019). The SOS system: a complex and tightly regulated response to DNA damage. *Environ. Mol. Mutagen.* 60, 368–384. doi: 10.1002/em.22267

Monod, J. (1949). The growth of bacterial cultures. *Annu. Rev. Microbiol.* 3, 371–394.

Nahrstedt, H., Schröder, C., and Meinhardt, F. (2005). Evidence for two *recA* genes mediating DNA repair in *Bacillus megaterium*. *Microbiology* 151(Pt 3), 775–787. doi: 10.1099/mic.0.27626-0

Norioka, N., Hsu, M. Y., Inouye, S., and Inouye, M. (1995). Two *recA* genes in *Myxococcus xanthus*. *J. Bacteriol.* 177, 4179–4182. doi: 10.1128/jb.177.14.4179-4182.1995

Nussbaumer, B., and Wohlleben, W. (1994). Identification, isolation and sequencing of the *recA* gene of *Streptomyces lividans* TK24. *FEMS Microbiol. Lett.* 118, 57–63. doi: 10.1016/0378-1097(94)90596-7

Orillard, E., Radicella, J. P., and Marsin, S. (2011). Biochemical and cellular characterization of *Helicobacter* pylori RecA, a protein with high-level constitutive expression. *J. Bacteriol.* 193, 6490–6497. doi: 10.1128/JB.05646-11

Prado, F. (2018). Homologous recombination: to fork and beyond. *Genes* 9, E603.

Quinet, A., Lemaçon, D., and Vindigni, A. (2017). Replication fork reversal: players and guardians. *Mol. Cell* 68, 830–833. doi: 10.1016/j.molcel.2017.11.022

Rastogi, R. P., Richa Kumar, A., Tyagi, M. B., and Sinha, R. P. (2010). Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J. Nucleic Acids* 2010:592980. doi: 10.4061/2010/592980

Richa, Sinha, R. P., and Häder, D. P. (2015). Physiological aspects of UV-excitation of DNA. *Top Curr. Chem.* 356, 203–248. doi: 10.1007/128_2014_531

Rolfe, M. D., Rice, C. J., Lucchini, S., Pin, C., Thompson, A., Cameron, A. D., et al. (2012). Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. *J. Bacteriol.* 194, 686–701. doi: 10.1128/JB.06112-11

Saint-Ruf, C., Pesut, J., Sopta, M., and Matic, I. (2007). Causes and consequences of DNA repair activity modulation during stationary phase in *Escherichia coli*. *Crit. Rev. Biochem. Mol. Biol.* 42, 259–270. doi: 10.1080/10409230701495599

Sano, Y., and Kageyama, M. (1987). The sequence and function of the *recA* gene and its protein in *Pseudomonas aeruginosa* PAO. *Mol. Gen. Genet.* 208, 412–419. doi: 10.1007/bf00328132

Schefe, J. H., Lehmann, K. E., Buschmann, I. R., Unger, T., and Funke-Kaiser, H. (2006). Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. *J. Mol. Med* 84, 901–910. doi: 10.1007/s00109-006-0097-6

Sheng, D., Liu, R., Xu, Z., Singh, P., Shen, B., and Hua, Y. (2005).). Dual negative regulatory mechanisms of RecX on RecA functions in radiation resistance, DNA recombination and consequent genome instability in *Deinococcus radiodurans*. *DNA Repair*. 4, 671–678. doi: 10.1016/j.dnarep.2005.02.006

Sinha, R. P., and Häder, D. P. (2002). UV-induced DNA damage and repair: a review. *Photochem. Photobiol. Sci.* 1, 225–236.

Story, R. M., Weber, I. T., and Steitz, T. A. (1992). The structure of the *E. coli* RecA protein monomer and polymer. *Nature* 355, 318–325. doi: 10.1038/355318a0

Ueki, T., Inouye, S., and Inouye, M. (1996). Positive-negative KG cassettes for construction of multi-gene deletions using a single drug marker. *Gene* 183, 153–157. doi: 10.1016/s0378-1119(96)00546-x

Umelo, E., Noonan, B., and Trust, T. J. (1996). Cloning, characterization and expression of the *recA* gene of *Aeromonas salmonicida*. *Gene* 175, 133–136. doi: 10.1016/0378-1119(96)00138-2

Vermeersch, L., Perez-Samper, G., Cerulus, B., Jariani, A., Gallone, B., Voordeckers, K., et al. (2019). On the duration of the microbial lag phase. *Curr. Genet.* 65, 721–727. doi: 10.1007/s00294-019-00938-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 08 April 2020
doi: 10.3389/fmicb.2020.00590

Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress

Joanna Hołówka and Jolanta Zakrzewska-Czerwińska*

Department of Molecular Microbiology, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

The bacterial chromosome must be efficiently compacted to fit inside the small and crowded cell while remaining accessible for the protein complexes involved in replication, transcription, and DNA repair. The dynamic organization of the nucleoid is a consequence of both intracellular factors (i.e., simultaneously occurring cell processes) and extracellular factors (e.g., environmental conditions, stress agents). Recent studies have revealed that the bacterial chromosome undergoes profound topological changes under stress. Among the many DNA-binding proteins that shape the bacterial chromosome structure in response to various signals, NAPs (nucleoid associated proteins) are the most abundant. These small, basic proteins bind DNA with low specificity and can influence chromosome organization under changing environmental conditions (i.e., by coating the chromosome in response to stress) or regulate the transcription of specific genes (e.g., those involved in virulence).

Keywords: stress response, nucleoid associated proteins, bacterial chromosome dynamics, bacterial chromosome compaction, host survival

OPEN ACCESS

Edited by:
Leise Riber,
University of Copenhagen, Denmark

Reviewed by:
Ole Skovgaard,
Roskilde University, Denmark
Soumitra Ghosh,
Université de Lausanne, Switzerland
Valakunja Nagaraja,
Indian Institute of Science, India

****Correspondence:***
Joanna Hołówka
joanna.holowka@uwr.edu.pl

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *19 January 2020*
Accepted: *18 March 2020*
Published: *08 April 2020*

Citation:
Hołówka J and Zakrzewska-Czerwińska J (2020) Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress. Front. Microbiol. 11:590. doi: 10.3389/fmicb.2020.00590

INTRODUCTION

Bacteria have developed a plethora of strategies to inhabit nearly every environment on Earth (Boor, 2006). To survive, bacteria must quickly adapt to changing environmental conditions. To date, dozens of group- or species-specific and universal adaptive mechanisms have been uncovered (Anderson and Kendall, 2017; Singh, 2017; Bussi and Gutierrez, 2019). Among them, changes in the architecture of the entire chromosome or particular chromosome regions (e.g., gene promoters) appear to be the most rapid and effective adaptation strategies, particularly in response to sudden stress (Boor, 2006; Morikawa et al., 2006; Trojanowski et al., 2019). Such a response is apparently universal, as it has been observed in many of the bacterial species investigated to date.

To fit the bacterial chromosome along with all associated proteins and RNA inside a tiny cell, the DNA has to be compacted more than 1000-fold (Murphy and Zimmerman, 1995). The nucleoid exhibits a multi-level hierarchical structural organization similar to that of eukaryotic chromatin (Macvanin and Adhya, 2012; Badrinarayanan et al., 2015; Verma et al., 2019; Dame et al., 2020). In the model organism, *Escherichia coli*, the 4.6-Mb chromosome is organized into four structural macrodomains (Ori, Ter, Left, and Right chromosomal arms) and the two unstructured regions, each of which consists of small (average ∼10 kb) topologically independent microdomains (Postow et al., 2004; Valens et al., 2004; Espeli et al., 2008). This hierarchical structure maintains the global nucleoid organization and ensures the accessibility of particular chromosomal regions for DNA-dependent processes, such as replication, transcription, DNA repair, and recombination. The organization of the highly compacted yet dynamic nucleoid structure reflects the input of many different factors, including molecular crowding, depletion forces, DNA supercoiling, and

nucleoid-associated proteins (NAPs) (Luijsterburg et al., 2006; de Vries, 2010; Dillon and Dorman, 2010; Jeon et al., 2017; Joyeux, 2019). The NAPs are small basic proteins that help compact the DNA into microdomains and also act as global regulators of transcription (Shahul Hameed et al., 2019). A great deal of studies indicated that NAPs play crucial roles in the ability of a bacterium to adapt to unfavorable conditions, particularly stress (Atlung and Ingmer, 1997; Nguyen et al., 2009; Kahramanoglou et al., 2011; Mangan et al., 2011; Datta et al., 2019). Under stress conditions, some NAPs can function as "rapid reaction forces" by introducing DNA topology changes that protect DNA or alter the transcriptional profile, particularly with respect to genes that are crucial for bacterial survival.

Here, we provide a mini review of the NAP-mediated rapid adaptation strategies that bacteria use to endure unfavorable conditions.

NUCLEOID DYNAMICS ARE ORCHESTRATED BY NAPs

The proper balance between chromosome compaction and the availability of chromosomal regions for the protein complexes involved in different cellular processes depends mainly on the DNA binding activity of NAPs (Krogh et al., 2018; Flores-Ríos et al., 2019). These small basic proteins can condense chromosomal DNA by bending, wrapping, and/or bridging relatively distant DNA strands (Luijsterburg et al., 2006; Dillon and Dorman, 2010). They all possess dimerization/oligomerization domains that facilitate chromosome coating and binding within the chromosomal regions to create inflexible filaments. Most NAPs show rather low sequence specificity for binding; however, their binding sites are often AT-rich, which is a characteristic feature of gene promoters (Kahramanoglou et al., 2011; Prieto et al., 2012; Odermatt et al., 2018). All bacterial species possess NAPs, some of which are unique for a given genus and/or species (Datta et al., 2019; Gehrke et al., 2019; Liu et al., 2019). The NAPs of *E. coli* are the best studied examples (Wold et al., 1996; Ali Azam et al., 1999; Ryan et al., 2002; Dillon and Dorman, 2010; Verma et al., 2019). The main NAPs include HU (heat-unstable protein), IHF (integration host factor), H-NS (histone-like nucleoid structuring protein), Lrp (leucine-responsive regulatory protein), Fis (factor for inversion stimulation), and Dps (DNA-binding protein from starved cells) (Luijsterburg et al., 2006; Wang et al., 2011). These NAPs can be divided based on their DNA-binding modes (**Figure 1**): HU, IHF, Fis, and Dps organize the chromosome by inducing bends into the DNA; H-NS can bridge two DNA strands; and in the case of Lrp, DNA is wrapped around the protein complex, enabling the joining of distant DNA strands. These DNA-binding activities of NAPs induce both topological and structural changes in the chromosomal DNA to ensure its proper compaction inside the cell. In addition to their architectural roles, NAPs are also involved in cellular processes, such as transcription (H-NS), DNA replication (HU, IHF, Fis), and DNA recombination, repair, and SOS response (HU) (Wold et al., 1996; Atlung and Ingmer, 1997; Kamashev and Rouviere-Yaniv, 2000; Ryan et al., 2002; Shahul Hameed et al., 2019). Given the variety of the functions overseen by NAPs, it is unsurprising that their expression pattern differs during growth (see **Figure 1**; Ali Azam et al., 1999; Dillon and Dorman, 2010; Verma et al., 2019). During the exponential phase of growth, the most abundant NAPs in *E. coli* include HU and Fis (Wold et al., 1996; Ryan et al., 2002; Kivisaar, 2020). Cells in the stationary phase produce NAPs that can most effectively condense the chromosome (e.g., Dps) (Calhoun and Kwon, 2011; Sato et al., 2013). Some NAPs (e.g., H-NS) are consistently expressed at a relatively low level, rendering them available to alter the expression of certain genes under a given stimulus (Shahul Hameed et al., 2019). NAPs have been shown to change the transcriptional profile of the cell (Atlung and Ingmer, 1997; Kahramanoglou et al., 2011), and this reportedly reflects their DNA-binding preferences. Recent studies have shown that, in addition to their growth-phase-dependent expression, some NAPs undergo posttranslational modifications (e.g., phosphorylation, acetylation, pupylation, succinylation) (Gupta et al., 2014; Ghosh et al., 2016; Okanishi et al., 2017; Dilweg and Dame, 2018). Acetylation and phosphorylation of basic residues (particularly those within the DNA-binding domain) will tend to neutralize or negatively shift the overall protein charge, respectively, which in turn decreases the DNA-binding activity of the modified NAP. Such additional control could be essential in the case of stress conditions, when the binding patterns of certain NAPs must be changed (Dilweg and Dame, 2018). The variety of NAPs and their balanced expression and activity regulation ensure the availability of chromosomal regions involved in cellular processes and enable the cell to adapt to various environmental and stress conditions. A rapid reaction to stress, which is crucial for the cell's ability to survive, mostly relies on NAPs DNA binding activity. By influencing gene expression and/or coating the chromosomal DNA, NAPs help the cell quickly react to changing conditions and thereby protect the DNA from damage.

NAPs EXHIBIT NUCLEOID-PROTECTING ACTIVITY UNDER STRESS CONDITIONS

Bacteria have developed numerous mechanisms to mount stress responses that enable the cell to adjust to changing conditions in various habitats (Boor, 2006; Bleuven and Landry, 2016). Saprophytic species living in soil or water are constantly subjected to potentially stressful environmental conditions, such as UV radiation, cold shock, heat shock, drying, and nutrient limitation. Some species survive by forming spores or endospores that can start a new population in a different niche and/or under more favorable conditions. Pathogens, meanwhile, have developed many sophisticated mechanisms that enable them to live inside the host cells (e.g., *Mycobacterium tuberculosis*, an etiological agent of tuberculosis, can survive within host alveolar macrophages for decades) (Bussi and Gutierrez, 2019). Most pathogenic species must face stress factors that reflect the host defenses mechanisms, such as low pH, oxidative stress, hypoxia, and limited nutrient availability (Anderson and Kendall, 2017;

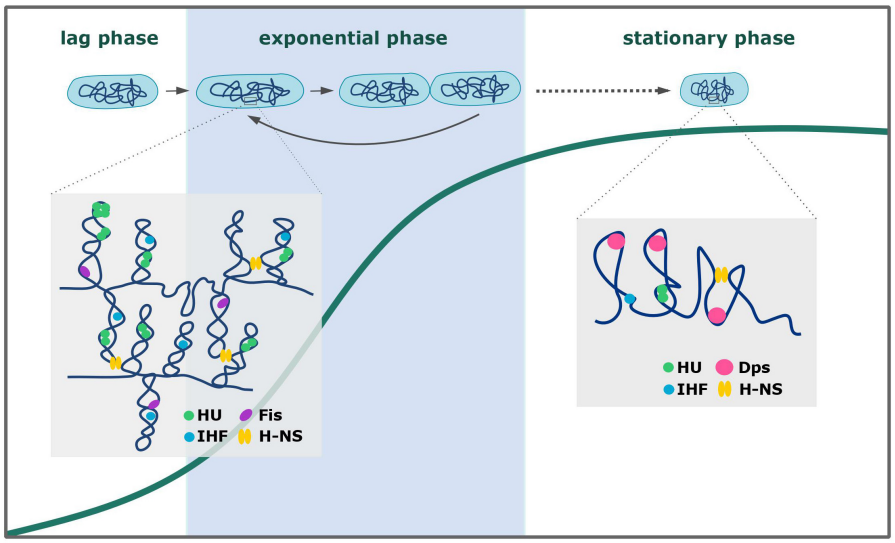
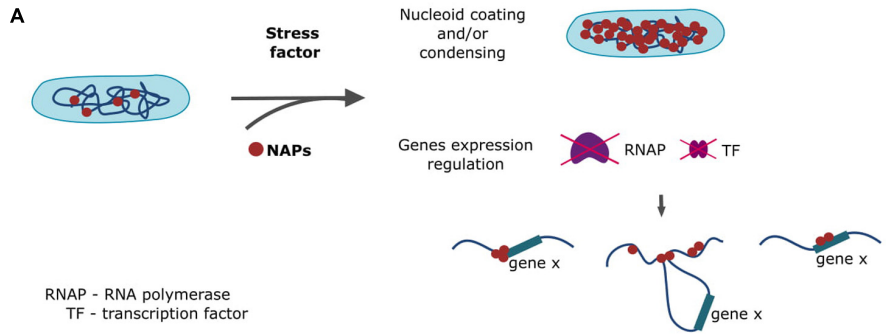


FIGURE 1 | Chromosome organization during the growth of *Escherichia coli*. The expression patterns of *E. coli* NAPs reflect the chromosome compaction level (higher in the stationary than in the exponential phase) and cellular processes that involve certain NAPs (Ali Azam et al., 1999; Luijsterburg et al., 2006; Dillon and Dorman, 2010). See text for a detailed description.

Singh, 2017; Bussi and Gutierrez, 2019; Datta et al., 2019). Beyond the systems that specifically cope with stress (e.g., the general stress response involving alternative sigma factors, the stringent response), the immediate protection comes from the NAPs (see **Figure 2**; Boor, 2006; Boutte and Crosson, 2013).

The cellular response will vary depending on the level, type, and duration of the perceived stress; such a response might range from the activation of precise mechanisms to the initiation of a global protective reaction that involves the whole nucleoid (**Figure 2A**; Atlung and Ingmer, 1997; Morikawa et al., 2006; Sato et al., 2013; Gehrke et al., 2019). The *M. smegmatis* nucleoid shrinks upon antibiotic treatment; this preserves the structure and integrity of the nucleoid and allows the cell to revive after the inhibitor removal (Trojanowski et al., 2019). Such tight chromosome compaction is also observed in the transition to the stationary phase, when the cells are shorter and there is much less room for the nucleoid (Meyer and Grainger, 2013). Some NAP family proteins can coat the whole chromosome; for example, HU can bind along the whole chromosome, although it prefers AT-rich regions and certain DNA structures (e.g., Holliday junctions, replication forks) rather than specific motifs (Kamashev and Rouviere-Yaniv, 2000; Bahloul et al., 2001). *Deinococcus radiodurans* is an extremophilic organism that is highly resistant to radiation of any type (e.g., ionizing radiation, UV light) (Makarova et al., 2001). Its genome encodes three HU protein homologs that contribute to the survival of this bacterium in unfavorable conditions (Nguyen et al., 2009). The homolog of *E. coli* HU protein encoded in the genome of pathogenic *Helicobacter pylori* also shows protective activity toward the chromosomal DNA, and a mutant strain lacking this HU-like protein exhibits increased sensitivity to oxidative and acid stress and decreased survival inside macrophages (Wang et al., 2012). The *E. coli* IHF protein shows a DNA-binding profile similar to that of HU (Azam et al., 2000; Wang et al., 2011). Both proteins exhibit a preference for AT-rich regions, but unlike HU, the IHF protein specifically recognizes 13-bp sequences with the consensus 5'-WATCAANNNNTTR-3' (Hales et al., 1994; Prieto et al., 2012). *M. tuberculosis* possesses homologs of both HU and IHF (HupB and mIHF, respectively), and these proteins are essential during the infection of macrophages (Pandey et al., 2014; Odermatt et al., 2018). Moreover, it was shown that expression of *hupB* gene increases during the infection (Kumar et al., 2011). When faced with nutrient exhaustion in their habitat, some bacteria, such as *Streptomyces*, form spores that enable them to survive. Many agents are involved in the proper switching of the life cycle; among them, HU and IHF play vital roles. In *S. coelicolor*, sIHF (IHF homolog) and HupS (HU-like protein) are required to enable the DNA to fit inside the tiny spores (spores deprived of sIHF or HupS are temperature sensitive) (Salerno et al., 2009; Swiercz et al., 2013). An HU-like protein found in the human pathogen, *Francisella tularensis* (the causative agent of tularemia), protects the DNA against free hydroxyl radicals (Stojkova et al., 2018). A similar mechanism of action is exhibited by the *Staphylococcus aureus* MgrA protein; this homolog of *E. coli* Dps coats the DNA, protecting it against oxidative stress and ensuring prolonged survival of the cell inside phagosomes (Crosby et al., 2016; Ushijima et al., 2017). It has been reported that the *M. tuberculosis* genome encodes a novel NAP, called NapM, that is also required for the pathogen to survive inside the host macrophages (Liu et al., 2019). The NapM sequence homolog from *M. smegmatis* was shown to colocalize with the *E. coli* nucleoid (Liu et al., 2016), potentially suggesting that this protein exhibits similar binding modes in pathogenic mycobacteria.

The NAPs involved in the protective DNA-coating mechanism share a few similarities, including a relatively low DNA-binding



B

Protein	Stress type	Response
Drd HU	radiation	DNA-coating protective mechanism
Hpy HU	oxidative/acidic stress	
Mtb NaM	host defences mechanisms	
Sar MgrA	oxidative stress	
Sco IHF, Sco HupS	nutrient exhaustion	
Ftn HU	oxidative stress	
Sm HU, Mtb HupB, Mtb IHF	host defences mechanisms	
Mtb Lsr2, Mtb NapA, Ddn H-NS, Ddn Fis	host defences mechanisms	Influencing genes expression
Sm H-NS, Sven Lsr2	environmental conditions	

Drd - Deinococcus radiodurans
Hpy - Helicobacter pylori
Mtb - Mycobacterium tuberculosis
Sar - Staphylococcus aureus
Sco - Streptomyces coelicolor
Ftn - Francisella tularensis
Sm - Salmonella
Ddn - Dickeya dadantii
Sven - Streptomyces venezuelae

FIGURE 2 | Involvement of NAPs in stress responses. **(A)** General mechanisms through which NAPs act in response to a stress factor (Dillon and Dorman, 2010; Meyer and Grainger, 2013; Kriel et al., 2018; Trojanowski et al., 2019). **(B)** Examples of the homologs of the canonical *E. coli* NAPs involved in the cellular response triggered upon detection of stress conditions.

specificity and a high copy number (Ali Azam et al., 1999; Dillon and Dorman, 2010; Verma et al., 2019). Homologs of the canonical "whole-chromosome binders" (i.e., HU, IHF, Dps) often possess unique structural features that ensure their effective binding along the entire chromosome. HU-like proteins in Actinobacteria (e.g., mycobacterial HupB, *S. coelicolor* HupS) have additional positively charged C-terminal domains that have been shown to stabilize the DNA-protein complexes (Salerno et al., 2009; Hołówka et al., 2017). *D. radiodurans* HU homologs have also repetitive basic residues, but within the N-terminal domain (Ghosh and Grove, 2006). The DNA-coating mechanism is activated immediately when unfavorable conditions are sensed, and helps maintain chromosomal integrity by creating a physical barrier against stress factors, such as radiation, antibiotic treatment, oxidative and acidic stress.

NAPs ALTER BASIC CELLULAR PROCESSES IN RESPONSE TO STRESS

The abilities to mount a rapid and effective response to changing environmental conditions and/or adjust the cell's metabolic activity to prolonged stress are key factors in the survival of both pathogens and saprophytes. The most "specific" stress response mechanism involving NAPs relies on their ability to influence the expression level of the certain gene(s) and/or gene cluster(s) (Kahramanoglou et al., 2011; Kriel et al., 2018; Flores-Ríos et al., 2019; Dorman et al., 2020). These small nucleoid organizers can affect transcription by inducing topological and/or structural changes in the chromosomal DNA (**Figure 2A**) that can alter the binding of RNA polymerase or transcription factors. Moreover, as mentioned earlier, NAPs often bind to AT-rich regions within promoter sequences and thereby repress gene expression. Depending on a given NAP's DNA-binding specificity and number of target DNA sequences, it can simultaneously affect the transcription of many genes/gene clusters or act as a specific switch that alters the expression levels of certain genes (Gordon et al., 2010; Prieto et al., 2012; Gehrke et al., 2019; Shahul Hameed et al., 2019). H-NS, which exhibits DNA-bridging activity (**Figure 1**), was shown to be a global transcription repressor in human pathogens, including *Salmonella enterica serovar Typhimurium*, *Vibrio cholerae*, and toxigenic strains of *E. coli* (Ayala et al., 2015; Helgesen et al., 2016; Shahul Hameed et al., 2019). Additionally, it was shown that a H-NS paralog, StpA protein cooperate with H-NS to alter virulence genes expression in uropathogenic *E. coli* strains (Müller et al., 2006).

The *E. coli* H-NS binding sites are reportedly clustered near the *ter* region, where genes connected with motility and biofilm formation are localized. Interestingly, H-NS from *Salmonella* acts as a repressor for horizontally acquired pathogenicity islands (Lucchini et al., 2006; Brunet et al., 2015). Similarly, a structural homolog of H-NS in *M. tuberculosis* (called Lsr2) is involved in the regulation of many genes, including those connected with virulence (Gordon et al., 2010). Deletion of the *lsr2* gene results in decreased growth and survival under hypoxia (Bartek et al., 2014), suggesting that Lsr2 could be a crucial agent that "switches" mycobacteria to the dormant state and enables them to endure inside host cells. The recently described NapA protein is another mycobacterial NAP that serves as a global transcription factor (Datta et al., 2019). It exhibits a preference for AT-rich regions and coats the DNA to create inflexible rods that interrupt DNA supercoiling. *M. tuberculosis* NapA regulates the expression of genes that encode virulence regulators. An Lsr2-like protein produced by another member of Actinobacteria, the saprophytic *S. venezuelae*, was shown to control genes whose products are involved in signaling and producing specialized secondary metabolites (Gehrke et al., 2019). Homologs of the *E. coli* H-NS and Fis proteins produced in the plant pathogen, *Dickeya dadantii*, influence the expression levels of the *pal* genes, which act as major virulence factors (Ouafa et al., 2012). Intriguingly, HU-like proteins found in *Salmonella* and *F. tularensis* not only create the physical protective barrier against stress factors, they also regulate genes involved in general physiology, metabolism, and virulence (**Figure 2B**; Mangan et al., 2011; Stojkova et al., 2018). The mycobacterial HupB protein regulates the expression of the *katG* gene (acting as a repressor), whose product activates the anti-tuberculosis drug, isoniazid (Niki et al., 2012; Enany et al., 2017); a *M. smegmatis* strain deprived of HupB showed increased susceptibility to this drug (Hołówka et al., 2017). Additionally, recent studies showed that the *M. tuberculosis* mIHF protein represses the expression of many genes, including those connected with pathogenesis (Odermatt et al., 2018).

In addition to their conventional architectural role and involvement in regulating gene expression, NAPs may also contribute to other cellular processes. For example, studies have shown that NAPs influence chromosome replication by binding and inducing some structural changes within the origin of chromosomal replication (*oriC*). In *E. coli*, the IHF and HU proteins facilitate the formation of the pre-replication complex, and the Fis protein prevents replication initiation (Wold et al., 1996; Ryan et al., 2002). Interestingly, expression of the *M. tuberculosis napM* gene increases upon stress; the NapM protein binds DnaA (a replication initiation protein) to inhibit chromosome replication, which in turn ensures that mycobacteria transition to the dormant state to survive inside host macrophages (Liu et al., 2019). Almost all processes involving spatial transitions of DNA strands, such as DNA repair and recombination and the topoisomerase-mediated maintenance of topological homeostasis, are based on cooperation with NAPs (e.g., HU interacts directly with topoisomerase A to alter its DNA-relaxing activity) (Shanado et al., 1998; Kamashev and Rouviere-Yaniv, 2000; Ghosh et al., 2014; Kivisaar, 2020). Overall, the low DNA-binding specificity and relatively high copy number of NAPs make them readily available and able to assist with complex cellular processes. Proper synchronization of the processes occurring inside the cell with constantly changing environmental conditions is a key element to survival under stress.

CONCLUSION

During the course of their evolution, bacteria developed the ability to rapidly adapt to constantly changing environmental conditions. Rapid reactions to many different signals, including stress factors, are crucial for the survival of both saprophytes and pathogens. As reviewed herein, NAPs ensure the very efficient and immediate response to various stimuli. These small basic proteins shape chromosomal DNA, adjusting its architecture in response to intra- and extracellular conditions. When the bacterial cell detects strong stress, NAPs (e.g., HU, Dps) generally coat and/or condense the nucleoid, creating a physical protective barrier for the DNA (Nguyen et al., 2009; Salerno et al., 2009; Pandey et al., 2014; Crosby et al., 2016; Odermatt et al., 2018). More specific NAP-related stress response mechanisms involve the ability of NAPs to regulate transcription. Upon binding, NAPs (e.g., H-NS, Fis) induce structural and/or topological DNA changes that lead to alteration of the expression levels of certain genes (Kahramanoglou et al., 2011; Brunet et al., 2015). Many genes involved in the adaptation to a new living condition, such as by formation of biofilm or alteration of motility, synthesis of secondary metabolites, and/or virulence, are regulated by NAPs. Additionally, NAPs can regulate basic cellular processes (e.g., replication initiation) in order to synchronize such processes with changing environmental conditions (Ryan et al., 2002; Datta et al., 2019). Hence, most NAPs act as the "rapid reaction forces" that enable the bacterial cell to endure under stress.

AUTHOR CONTRIBUTIONS

JH wrote the main body of the manuscript, conclusion and prepared figures. JZ-C wrote the introduction and revise the entire manuscript.

FUNDING

This work was financed by grants (OPUS 2017/25/B/NZ1/00657 and SONATINA 2018/28/C/NZ1/00128) from the National Science Center (Poland).

ACKNOWLEDGMENTS

We are very grateful to members of our laboratory for their comments on the manuscript.

REFERENCES

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

Anderson, C. J., and Kendall, M. M. (2017). *Salmonella enterica* serovar typhimurium strategies for host adaptation. *Front. Microbiol.* 8:1983. doi: 10.3389/fmicb.2017.01983

Atlung, T., and Ingmer, H. (1997). H-NS: a modulator of environmentally regulated gene expression. *Mol. Microbiol.* 24, 7–17. doi: 10.1046/j.1365-2958.1997.3151679.x

Ayala, J. C., Wang, H., Benitez, J. A., and Silva, A. J. (2015). RNA-Seq analysis and whole genome DNA-binding profile of the *Vibrio cholerae* histone-like nucleoid structuring protein (H-NS). *Genom Data* 5, 147–150. doi: 10.1016/j.gdata.2015.05.039

Azam, T. A., Hiraga, S., and Ishihama, A. (2000). Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5, 613–626. doi: 10.1046/j.1365-2443.2000.00350.x

Badrinarayanan, A., Le, T. B. K., and Laub, M. T. (2015). Bacterial chromosome organization and segregation. *Annu. Rev. Cell Dev. Biol.* 31, 171–199. doi: 10.1146/annurev-cellbio-100814-125211

Bahloul, A., Boubrik, F., and Rouviere-Yaniv, J. (2001). Roles of *Escherichia coli* histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of dnaA46ts. *Biochimie* 83, 219–229. doi: 10.1016/s0300-9084(01)01246-9

Bartek, I. L., Woolhiser, L. K., Baughn, A. D., Basaraba, R. J., Jacobs, W. R., Lenaerts, A. J., et al. (2014). *Mycobacterium tuberculosis* Lsr2 is a global transcriptional regulator required for adaptation to changing oxygen levels and virulence. *mBio* 5:e01106-14. doi: 10.1128/mBio.01106-14

Bleuven, C., and Landry, C. R. (2016). Molecular and cellular bases of adaptation to a changing environment in microorganisms. *Proc. Biol. Sci.* 283:20161458. doi: 10.1098/rspb.2016.1458

Boor, K. J. (2006). Bacterial stress responses: what doesn't kill them can make them stronger. *PLoS Biol.* 4:e0040023. doi: 10.1371/journal.pbio.0040023

Boutte, C. C., and Crosson, S. (2013). Bacterial lifestyle shapes the regulation of stringent response activation. *Trends Microbiol.* 21, 174–180. doi: 10.1016/j.tim.2013.01.002

Brunet, Y. R., Khodr, A., Logger, L., Aussel, L., Mignot, T., Rimsky, S., et al. (2015). H-NS Silencing of the *Salmonella* pathogenicity island 6-encoded Type VI secretion system limits *Salmonella enterica* serovar typhimurium interbacterial killing. *Infect. Immun.* 83, 2738–2750. doi: 10.1128/IAI.00198-15

Bussi, C., and Gutierrez, M. G. (2019). Mycobacterium tuberculosis infection of host cells in space and time. *FEMS Microbiol. Rev.* 43, 341–361. doi: 10.1093/femsre/fuz006

Calhoun, L. N., and Kwon, Y. M. (2011). Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in *Escherichia coli*: a review. *J. Appl. Microbiol.* 110, 375–386. doi: 10.1111/j.1365-2672.2010.04890.x

Crosby, H. A., Schlievert, P. M., Merriman, J. A., King, J. M., Salgado-Pabón, W., and Horswill, A. R. (2016). The *Staphylococcus aureus* global regulator MgrA modulates clumping and virulence by controlling surface protein expression. *PLoS Pathog.* 12:e1005604. doi: 10.1371/journal.ppat.1005604

Dame, R. T., Rashid, F.-Z. M., and Grainger, D. C. (2020). Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.* 21, 227–242. doi: 10.1038/s41576-019-0185-4

Datta, C., Jha, R. K., Ganguly, S., and Nagaraja, V. (2019). NapA (Rv0430), a novel nucleoid-associated protein that regulates a virulence operon in *Mycobacterium tuberculosis* in a supercoiling-dependent manner. *J. Mol. Biol.* 431, 1576–1591. doi: 10.1016/j.jmb.2019.02.029

de Vries, R. (2010). DNA condensation in bacteria: interplay between macromolecular crowding and nucleoid proteins. *Biochimie* 92, 1715–1721. doi: 10.1016/j.biochi.2010.06.024

Dillon, S. C., and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* 8, 185–195. doi: 10.1038/nrmicro2261

Dilweg, I. W., and Dame, R. T. (2018). Post-translational modification of nucleoid-associated proteins: an extra layer of functional modulation in bacteria? *Biochem. Soc. Trans.* 46, 1381–1392. doi: 10.1042/BST20180488

Dorman, C. J., Schumacher, M. A., Bush, M. J., Brennan, R. G., and Buttner, M. J. (2020). When is a transcription factor a NAP? *Curr. Opin. Microbiol.* 55, 26–33. doi: 10.1016/j.mib.2020.01.019

Enany, S., Yoshida, Y., Tateishi, Y., Ozeki, Y., Nishiyama, A., Savitskaya, A., et al. (2017). Mycobacterial DNA-binding protein 1 is critical for long term survival of Mycobacterium smegmatis and simultaneously coordinates cellular functions. *Sci. Rep.* 7:6810. doi: 10.1038/s41598-017-06480-w

Espeli, O., Mercier, R., and Boccard, F. (2008). DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. *Mol. Microbiol.* 68, 1418–1427. doi: 10.1111/j.1365-2958.2008.06239.x

Flores-Ríos, R., Quatrini, R., and Loyola, A. (2019). Endogenous and foreign nucleoid-associated proteins of bacteria: occurrence, interactions and effects on mobile genetic elements and host's biology. *Comput. Struct. Biotechnol. J.* 17, 746–756. doi: 10.1016/j.csbj.2019.06.010

Gehrke, E. J., Zhang, X., Pimentel-Elardo, S. M., Johnson, A. R., Rees, C. A., Jones, S. E., et al. (2019). Silencing cryptic specialized metabolism in *Streptomyces* by the nucleoid-associated protein Lsr2. *eLife* 8:e047691. doi: 10.7554/eLife.47691

Ghosh, S., and Grove, A. (2006). The Deinococcus radiodurans-encoded HU protein has two DNA-binding domains. *Biochemistry* 45, 1723–1733. doi: 10.1021/bi0514010

Ghosh, S., Mallick, B., and Nagaraja, V. (2014). Direct regulation of topoisomerase activity by a nucleoid-associated protein. *Nucleic Acids Res.* 42, 11156–11165. doi: 10.1093/nar/gku804

Ghosh, S., Padmanabhan, B., Anand, C., and Nagaraja, V. (2016). Lysine acetylation of the Mycobacterium tuberculosis HU protein modulates its DNA binding and genome organization. *Mol. Microbiol.* 100, 577–588. doi: 10.1111/mmi.13339

Gordon, B. R. G., Li, Y., Wang, L., Sintsova, A., van Bakel, H., Tian, S., et al. (2010). Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5154–5159. doi: 10.1073/pnas.0913551107

Gupta, M., Sajid, A., Sharma, K., Ghosh, S., Arora, G., Singh, R., et al. (2014). HupB, a nucleoid-associated protein of *Mycobacterium tuberculosis*, is modified by serine/threonine protein kinases *in vivo*. *J. Bacteriol.* 196, 2646–2657. doi: 10.1128/JB.01625-14

Hales, L. M., Gumport, R. I., and Gardner, J. F. (1994). Determining the DNA sequence elements required for binding integration host factor to two different target sites. *J. Bacteriol.* 176, 2999–3006.

Helgesen, E., Fossum-Raunehaug, S., and Skarstad, K. (2016). Lack of the H-NS protein results in extended and aberrantly positioned DNA during chromosome replication and segregation in *Escherichia coli*. *J. Bacteriol.* 198, 1305–1316. doi: 10.1128/JB.00919-15

Hołówka, J., Trojanowski, D., Ginda, K., Wojtaś, B., Gielniewski, B., Jakimowicz, D., et al. (2017). HupB is a bacterial nucleoid-associated protein with an indispensable eukaryotic-like tail. *mBio* 8:e01272-17. doi: 10.1128/mBio.01272-17

Jeon, C., Jung, Y., and Ha, B.-Y. (2017). A ring-polymer model shows how macromolecular crowding controls chromosome-arm organization in *Escherichia coli*. *Sci. Rep.* 7:11896. doi: 10.1038/s41598-017-10421-y

Joyeux, M. (2019). Preferential localization of the bacterial nucleoid. *Microorganisms* 7:204. doi: 10.3390/microorganisms7070204

Kahramanoglou, C., Seshasayee, A. S. N., Prieto, A. I., Ibberson, D., Schmidt, S., Zimmermann, J., et al. (2011). Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res.* 39, 2073–2091. doi: 10.1093/nar/gkq934

Kamashev, D., and Rouviere-Yaniv, J. (2000). The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *EMBO J.* 19, 6527–6535. doi: 10.1093/emboj/19.23.6527

Kivisaar, M. (2020). Mutation and recombination rates vary across bacterial chromosome. *Microorganisms* 8:25. doi: 10.3390/microorganisms8010025

Kriel, N. L., Gallant, J., van Wyk, N., van Helden, P., Sampson, S. L., Warren, R. M., et al. (2018). Mycobacterial nucleoid associated proteins: an added dimension in gene regulation. *Tuberculosis* 108, 169–177. doi: 10.1016/j.tube.2017.12.004

Krogh, T. J., Møller-Jensen, J., and Kaleta, C. (2018). Impact of chromosomal architecture on the function and evolution of bacterial genomes. *Front. Microbiol.* 9:2019. doi: 10.3389/fmicb.2018.02019

Kumar, M., Khan, F. G., Sharma, S., Kumar, R., Faujdar, J., Sharma, R., et al. (2011). Identification of Mycobacterium tuberculosis genes preferentially expressed

during human infection. *Microb. Pathog.* 50, 31–38. doi: 10.1016/j.micpath.2010.10.003

Liu, Y., Wang, H., Cui, T., Zhou, X., Jia, Y., Zhang, H., et al. (2016). NapM, a new nucleoid-associated protein, broadly regulates gene expression and affects mycobacterial resistance to anti-tuberculosis drugs. *Mol. Microbiol.* 101, 167–181. doi: 10.1111/mmi.13383

Liu, Y., Xie, Z., Zhou, X., Li, W., Zhang, H., and He, Z.-G. (2019). NapM enhances the survival of Mycobacterium tuberculosis under stress and in macrophages. *Commun. Biol.* 2:65. doi: 10.1038/s42003-019-0314-9

Lucchini, S., Rowley, G., Goldberg, M. D., Hurd, D., Harrison, M., and Hinton, J. C. D. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* 2:e81. doi: 10.1371/journal.ppat.0020081

Luijsterburg, M. S., Noom, M. C., Wuite, G. J. L., and Dame, R. T. (2006). The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.* 156, 262–272. doi: 10.1016/j.jsb.2006.05.006

Macvanin, M., and Adhya, S. (2012). Architectural organization in *E. coli* nucleoid. *Biochim. Biophys. Acta* 1819, 830–835. doi: 10.1016/j.bbagrm.2012.02.012

Makarova, K. S., Aravind, L., Wolf, Y. I., Tatusov, R. L., Minton, K. W., Koonin, E. V., et al. (2001). Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. *Microbiol. Mol. Biol. Rev.* 65, 44–79. doi: 10.1128/MMBR.65.1.44-79.2001

Mangan, M. W., Lucchini, S., Cróinín, T. O., Fitzgerald, S., Hinton, J. C. D., and Dorman, C. J. (2011). Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar Typhimurium. *Microbiology* 157, 1075–1087. doi: 10.1099/mic.0.046359-0

Meyer, A. S., and Grainger, D. C. (2013). The *Escherichia coli* nucleoid in stationary phase. *Adv. Appl. Microbiol.* 83, 69–86. doi: 10.1016/B978-0-12-407678-5.00002-7

Morikawa, K., Ohniwa, R. L., Kim, J., Maruyama, A., Ohta, T., and Takeyasu, K. (2006). Bacterial nucleoid dynamics: oxidative stress response in *Staphylococcus aureus*. *Genes Cells* 11, 409–423. doi: 10.1111/j.1365-2443.2006.00949.x

Müller, C. M., Dobrindt, U., Nagy, G., Emödy, L., Uhlin, B. E., and Hacker, J. (2006). Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic *Escherichia coli*. *J. Bacteriol.* 188, 5428–5438. doi: 10.1128/JB.01956-05

Murphy, L. D., and Zimmerman, S. B. (1995). Condensation and cohesion of lambda DNA in cell extracts and other media: implications for the structure and function of DNA in prokaryotes. *Biophys. Chem.* 57, 71–92. doi: 10.1016/0301-4622(95)00047-2

Nguyen, H. H., de la Tour, C. B., Toueille, M., Vannier, F., Sommer, S., and Servant, P. (2009). The essential histone-like protein HU plays a major role in *Deinococcus radiodurans* nucleoid compaction. *Mol. Microbiol.* 73, 240–252. doi: 10.1111/j.1365-2958.2009.06766.x

Niki, M., Niki, M., Tateishi, Y., Ozeki, Y., Kirikae, T., Lewin, A., et al. (2012). A novel mechanism of growth phase-dependent tolerance to isoniazid in mycobacteria. *J. Biol. Chem.* 287, 27743–27752. doi: 10.1074/jbc.M111.333385

Odermatt, N. T., Sala, C., Benjak, A., and Cole, S. T. (2018). Essential nucleoid associated protein mIHF (Rv1388) controls virulence and housekeeping genes in *Mycobacterium tuberculosis*. *Sci. Rep.* 8:14214. doi: 10.1038/s41598-018-32340-2

Okanishi, H., Kim, K., Fukui, K., Yano, T., Kuramitsu, S., and Masui, R. (2017). Proteome-wide identification of lysine succinylation in thermophilic and mesophilic bacteria. *Biochim. Biophys. Acta Proteins Proteom.* 1865, 232–242. doi: 10.1016/j.bbapap.2016.11.009

Ouafa, Z.-A., Reverchon, S., Lautier, T., Muskhelishvili, G., and Nasser, W. (2012). The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium *Dickeya dadantii*. *Nucleic Acids Res.* 40, 4306–4319. doi: 10.1093/nar/gks014

Pandey, S. D., Choudhury, M., Yousuf, S., Wheeler, P. R., Gordon, S. V., Ranjan, A., et al. (2014). Iron-regulated protein HupB of *Mycobacterium tuberculosis* positively regulates siderophore biosynthesis and is essential for growth in macrophages. *J. Bacteriol.* 196, 1853–1865. doi: 10.1128/JB.01483-13

Postow, L., Hardy, C. D., Arsuaga, J., and Cozzarelli, N. R. (2004). Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev.* 18, 1766–1779. doi: 10.1101/gad.1207504

Prieto, A. I., Kahramanoglou, C., Ali, R. M., Fraser, G. M., Seshasayee, A. S. N., and Luscombe, N. M. (2012). Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res.* 40, 3524–3537. doi: 10.1093/nar/gkr1236

Ryan, V. T., Grimwade, J. E., Nievera, C. J., and Leonard, A. C. (2002). IHF and HU stimulate assembly of pre-replication complexes at *Escherichia coli* oriC by two different mechanisms. *Mol. Microbiol.* 46, 113–124. doi: 10.1046/j.1365-2958.2002.03129.x

Salerno, P., Larsson, J., Bucca, G., Laing, E., Smith, C. P., and Flärdh, K. (2009). One of the two genes encoding nucleoid-associated HU proteins in *Streptomyces coelicolor* is developmentally regulated and specifically involved in spore maturation. *J. Bacteriol.* 191, 6489–6500. doi: 10.1128/JB.00709-09

Sato, Y. T., Watanabe, S., Kenmotsu, T., Ichikawa, M., Yoshikawa, Y., Teramoto, J., et al. (2013). Structural change of DNA induced by nucleoid proteins: growth phase-specific fis and stationary phase-specific Dps. *Biophys. J.* 105, 1037–1044. doi: 10.1016/j.bpj.2013.07.025

Shahul Hameed, U. F., Liao, C., Radhakrishnan, A. K., Huser, F., Aljedani, S. S., Zhao, X., et al. (2019). H-NS uses an autoinhibitory conformational switch for environment-controlled gene silencing. *Nucleic Acids Res.* 47, 2666–2680. doi: 10.1093/nar/gky1299

Shanado, Y., Kato, J., and Ikeda, H. (1998). *Escherichia coli* HU protein suppresses DNA-gyrase-mediated illegitimate recombination and SOS induction. *Genes Cells* 3, 511–520. doi: 10.1046/j.1365-2443.1998.00208.x

Singh, S. K. (2017). Staphylococcus aureus intracellular survival: a closer look in the process. *Virulence* 8, 1506–1507. doi: 10.1080/21505594.2017.1384896

Stojkova, P., Spidlova, P., Lenco, J., Rehulkova, H., Kratka, L., and Stulik, J. (2018). HU protein is involved in intracellular growth and full virulence of *Francisella tularensis*. *Virulence* 9, 754–770. doi: 10.1080/21505594.2018.1441588

Swiercz, J. P., Nanji, T., Gloyd, M., Guarné, A., and Elliot, M. A. (2013). A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Res.* 41, 4171–4184. doi: 10.1093/nar/gkt095

Trojanowski, D., Kołodziej, M., Hołówka, J., Müller, R., and Zakrzewska-Czerwińska, J. (2019). Watching DNA replication inhibitors in action: exploiting time-lapse microfluidic microscopy as a tool for target-drug interaction studies in *Mycobacterium*. *Antimicrob. Agents Chemother.* 63:e00739-19. doi: 10.1128/AAC.00739-19

Ushijima, Y., Ohniwa, R. L., and Morikawa, K. (2017). Identification of nucleoid associated proteins (NAPs) under oxidative stress in *Staphylococcus aureus*. *BMC Microbiol.* 17:207. doi: 10.1186/s12866-017-1114-3

Valens, M., Penaud, S., Rossignol, M., Cornet, F., and Boccard, F. (2004). Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J.* 23, 4330–4341. doi: 10.1038/sj.emboj.7600434

Verma, S. C., Qian, Z., and Adhya, S. L. (2019). Architecture of the *Escherichia coli* nucleoid. *PLoS Genet.* 15:e1008456. doi: 10.1371/journal.pgen.1008456

Wang, G., Lo, L. F., and Maier, R. J. (2012). A histone-like protein of *Helicobacter pylori* protects DNA from stress damage and aids host colonization. *DNA Repair* 11, 733–740. doi: 10.1016/j.dnarep.2012.06.006

Wang, W., Li, G.-W., Chen, C., Xie, X. S., and Zhuang, X. (2011). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* 333, 1445–1449. doi: 10.1126/science.1204697

Wold, S., Crooke, E., and Skarstad, K. (1996). The *Escherichia coli* Fis protein prevents initiation of DNA replication from oriC in vitro. *Nucleic Acids Res.* 24, 3527–3532. doi: 10.1093/nar/24.18.3527

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 15 April 2020
doi: 10.3389/fmicb.2020.00588

Chromosome Segregation Proteins as Coordinators of Cell Cycle in Response to Environmental Conditions

*Monika Pióro and Dagmara Jakimowicz**

Department of Molecular Microbiology, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

Chromosome segregation is a crucial stage of the cell cycle. In general, proteins involved in this process are DNA-binding proteins, and in most bacteria, ParA and ParB are the main players; however, some bacteria manage this process by employing other proteins, such as condensins. The dynamic interaction between ParA and ParB drives movement and exerts positioning of the chromosomal origin of replication (*oriC*) within the cell. In addition, both ParA and ParB were shown to interact with the other proteins, including those involved in cell division or cell elongation. The significance of these interactions for the progression of the cell cycle is currently under investigation. Remarkably, DNA binding by ParA and ParB as well as their interactions with protein partners conceivably may be modulated by intra- and extracellular conditions. This notion provokes the question of whether chromosome segregation can be regarded as a regulatory stage of the cell cycle. To address this question, we discuss how environmental conditions affect chromosome segregation and how segregation proteins influence other cell cycle processes.

Keywords: chromosome segregation, ParA, ParB, segrosome, cell division, cell elongation

OPEN ACCESS

Edited by:
Leise Riber,
University of Copenhagen, Denmark

Reviewed by:
Barbara Funnell,
University of Toronto, Canada
Marie-Françoise Noirot-Gros,
Argonne National Laboratory (DOE), United States
Preeti Srivastava,
Indian Institute of Technology Delhi, India

****Correspondence:***
Dagmara Jakimowicz
dagmara.jakimowicz@uwr.edu.pl

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *19 January 2020*
Accepted: *18 March 2020*
Published: *15 April 2020*

Citation:
Pióro M and Jakimowicz D (2020) Chromosome Segregation Proteins as Coordinators of Cell Cycle in Response to Environmental Conditions. Front. Microbiol. 11:588. doi: 10.3389/fmicb.2020.00588

INTRODUCTION

Bacteria must adjust their cell cycle to their environmental conditions. Unfavorable conditions such as starvation, oxidative, or osmotic stress alter the energetic state of the cell and trigger the stress signaling molecules (sigma factors, response regulators, signaling nucleotides) (Hengge, 2009; Gottesman, 2019; Latoscha et al., 2019; McLean et al., 2019). As the result, cells modify transcription, increase the generation time, completely inhibit cell division or sometimes form spores or enter dormancy (Errington, 2003; Jones et al., 2013; Heinrich et al., 2015; Desai and Kenney, 2019). While the primary cell cycle checkpoints are the initiation of replication and onset of cell division, these processes must be tightly coordinated with chromosome segregation (for recent reviews, see: Dewachter et al., 2018; Marczynski et al., 2019; Reyes-Lamothe and Sherratt, 2019; Burby and Simmons, 2020). Thus, the chromosome segregation process may link critical stages of the cell cycle.

The role of segregation proteins is to control the positioning of chromosomal (or plasmid) DNA during cell division. Importantly, in most bacterial cells, chromosome segregation begins soon after the initiation of chromosome replication and must be completed before the termination of cell division (Dewachter et al., 2018; Reyes-Lamothe and Sherratt, 2019). During chromosome replication, segregation proteins position newly duplicated chromosomal origin of replication (*oriC*) regions and ensure proper chromosome organization. Interestingly, in a number of bacterial species, positioning of the *oriC* and the pattern of chromosome organization may be modified in response to altered environmental conditions, such as limited nutrients (Wang et al., 2013; Badrinarayanan et al., 2015). Perfect examples of this modification are the profound changes in chromosome compaction observed during starvation-induced sporulation of *Bacillus subtilis* and *Streptomyces* spp. (Errington, 2001; Flärdh and Buttner, 2009; Jakimowicz and van Wezel, 2012). While in vegetatively growing *B. subtilis,* the *oriC* is shifted away from the cell pole, during the formation of endospores, which begins with asymmetric cell division, the *oriC* is anchored at the poles (Wang et al., 2014). However, even in non-sporulating bacteria, chromosome organization patterns may be altered depending on culture conditions; for example, in *Escherichia coli,* the chromosome arrangement changes from one in which the *oriC* adopts a mid-cell position in fast-growing cultures to a longitudinal in slow-growing cells in minimal media (Kleckner et al., 2014; Badrinarayanan et al., 2015). Chromosome topology is controlled by a set of proteins, predominantly topoisomerases and nucleoid-associated proteins (NAPs), whose activities were shown to be influenced by environmental and physiological factors (temperature, pH, salt concentration) (Dorman and Dorman, 2016; Dame et al., 2019). However, the mechanisms by which chromosome arrangement and segregation are adjusted to physiological state of bacterial cell only begun to emerge.

In this review, we discuss how chromosome segregation may be influenced by environmental stress, particularly nutrients limitation, and induced by this factor change of physiological conditions. Furthermore, we explore how changes in segregation protein activity may allow cells to adjust to particular conditions. To address these issues, we focus on the interactions of segregation proteins with DNA and the crosstalk between segregation proteins and their partners.

THE FUNCTIONS OF CHROMOSOME SEGREGATION PROTEINS

Chromosome segregation in bacteria is governed by a set of proteins, among which ParA and ParB are key players. ParA and ParB were first identified as plasmid segregation proteins and further studies revealed that homologs of these proteins control positioning of the chromosomal *oriC* region (Lee et al., 2003; Mierzejewska and Jagura-Burdzy, 2012; Funnell, 2016). However, ParA and ParB are not fully widespread and chromosome segregation in some bacterial species (i.e., some γ-proteobacteria, including *E. coli*, which lack *parAB* genes) exploits the activities of other proteins, such as condensins (Nolivos and Sherratt, 2014; Dewachter et al., 2018). Condensins, which compact the bulk of chromosomal DNA, also play an auxiliary function in ParAB-dependent chromosome segregation (Graumann et al., 1998; Moriya et al., 1998; Wang et al., 2015). Segregation of the terminus-proximal region (*ter*) usually requires the activities of additional proteins, such as the DNA translocase FtsK or a type II topoisomerase specialized in DNA decatenation (TopoIV) (Kato et al., 1990; Yu et al., 1998; Dewachter et al., 2018) (**Figure 1**).

The Role of ParA and ParB in *oriC* Segregation

The ParA and ParB proteins are components of the tripartite segregation system, which also involves *parS* sequences bound by ParB (*parABS* system) (Schumacher, 2008; Badrinarayanan et al., 2015). From 1 to 20 *parS* sequences, depending on the bacterial species, may be scattered over the *oriC*-proximal chromosomal region, which encompasses a range from 10 kb (in *Caulobacter crescentus*) to 200 kb in *Streptomyces coelicolor*, or even up to 650 kb in *Pseudomonas aeruginosa* (Jakimowicz et al., 2002; Livny et al., 2007; Tran et al., 2018; Kawalek et al., 2020). Upon interaction with ParB, the *parS* sequence-rich region engages in the formation of a large nucleoprotein complex named the segrosome (Funnell, 2014; Oliva, 2016). *parS* sites are often referred to as centromeric sites since they mark the chromosomal region that segregates first. Interestingly, in all studied bacterial species that use the ParABS system, ParB binding to DNA was reported to be maintained during majority of the cell cycle. In fact, the initiation of chromosome replication can be detected as the duplication of ParB complexes and the segrosomes mark the positions of the *oriC* regions throughout the whole cell cycle (Ptacin et al., 2010; Schofield et al., 2010; Shebelut et al., 2010; Harms et al., 2013; Trojanowski et al., 2015; Kois-Ostrowska et al., 2016) (**Figure 1**).

The organization of the ParB complex is still under investigation, but the studies to date have revealed that its architecture seems to be adjusted for the requirements of each particular bacterial species. Conserved structural features of ParB include the DNA-binding HTH motif in the central part of the polypeptide chain and two conserved sequences named ParB boxes in proximity to the N-terminus (Leonard et al., 2004; Schumacher and Funnell, 2005; Schumacher, 2007, 2008). While, the ParB boxes were shown to be required for segrosome formation, the N-terminal ParB domain is critical for interactions with ParA and C-terminal domain facilitates non-specific DNA binding (Autret et al., 2001; Schumacher and Funnell, 2005; Fisher et al., 2017). Non-specific interactions with DNA allow ParB to spread away from *parS* sites (Murray et al., 2006; Breier and Grossman, 2007; Kusiak et al., 2011; Chen et al., 2015; Taylor et al., 2015; Song et al., 2017). Moreover, ParB complex assembly was shown to require the bridging of protein molecules bound to spatially distant *parS* sites (Graham et al., 2014; Song et al., 2017; Soh et al., 2019). This bridging is mediated by dimerization of the arginine-rich ParB box II and was recently shown to be modulated by CTP binding in this region (Osorio-Valeriano et al., 2019; Soh et al., 2019). Notably, while in all bacterial species that

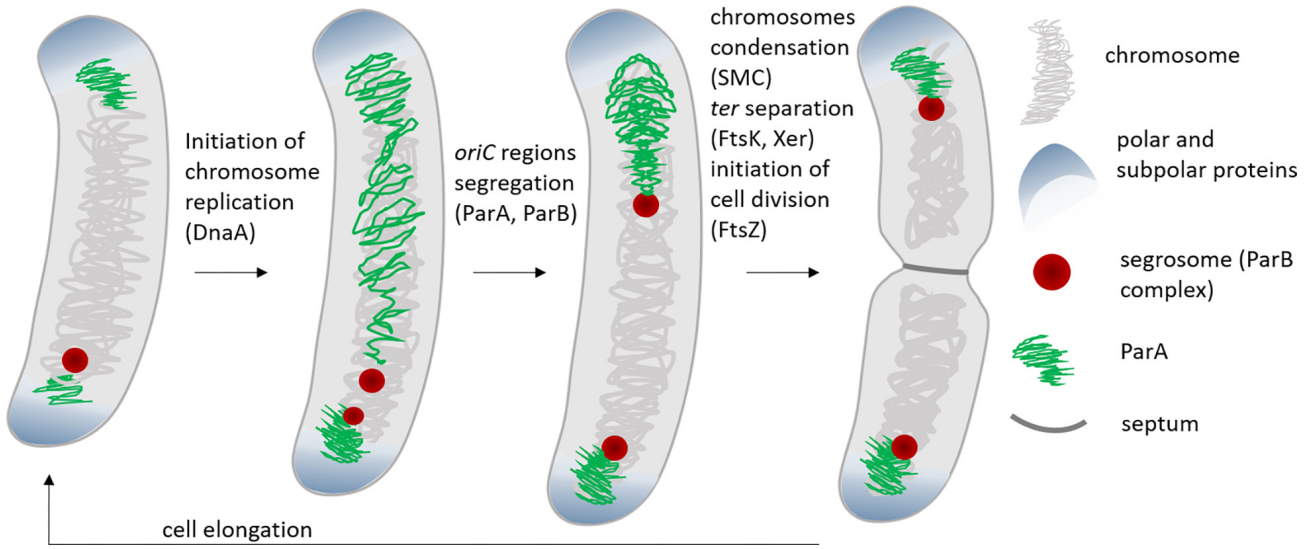


FIGURE 1 | Stages of cell cycle and chromosome segregation in bacteria with polar or subpolar *oriC* localization (e.g., *C. crescentus*, *V. cholerae*, *M. smegmatis*, *M. xanthus*). The scheme shows the stage of the cell cycle when activity of DnaA – chromosome replication initiator, ParA, ParB, SMC, and FtsK – proteins involved in chromosome segregation, as well as FtsZ- cell division initiator is required.

use the *parABS* system, ParB specifically binds *parS* sequences, the affinity and specificity of ParB toward *parS* sequences vary among bacteria, resulting in differences in ParB spreading (Jalal et al., 2019). These differences, as well as variations in the number and distribution of *parS* sites, are reflected in the species-specific architecture of the ParB complex. Nevertheless, the primary role of the ParB complex is to organize the *oriC*-proximal region of the chromosome to facilitate its movement; the ParB complex thus performs a critical step in chromosome segregation.

The driving force for the chromosomal ParB complex is provided by a P-loop ATPase - ParA. Over the last decade, the hypothesized ParA mechanism of action has changed from filament formation to the generation of a dynamic concentration gradient (Ptacin et al., 2010; Lim et al., 2014; Vecchiarelli et al., 2014; Le Gall et al., 2016). Pivotal for the gradient-based model is non-specific DNA binding by ATP-bound ParA dimers (Leonard et al., 2005; Hester and Lutkenhaus, 2007). Interaction with segrosomes stimulates ParA ATPase activity, while ATP hydrolysis triggers dimer dissociation and protein release from DNA. This generates a ParA-depleted zone in proximity to the ParB complex. Due to their high affinity for DNA-bound ParA dimers, segrosomes move away from the depletion zone toward higher ParA concentration. The directionality of the ParA gradient and ParB movement was suggested to be enhanced by interactions between ParA and the polar or subpolar proteins: TipN and PopZ in *C. crescentus*, the bactofilin complex in *Myxococcus xanthus*, HubP in *Vibrio cholerae* and DivIVA in *Mycobacterium smegmatis* (Bowman et al., 2008; Ebersbach et al., 2008; Yamaichi et al., 2012; Ginda et al., 2013; Lin et al., 2017). These interactions are critical for proper *oriC* subcellular localization (Bowman et al., 2008; Ebersbach et al., 2008; Yamaichi et al., 2012; Lin et al., 2017; Pióro et al., 2019). Thus, highly genus- or species-specific interactions play a role in the spatial coordination of chromosome segregation with other cell cycle processes and presumably adjust the segregation machinery to the requirements of the life cycle of a particular bacterium.

The Additional Roles of ParAB Proteins

In addition to their main function in moving the *oriC* region, in *B. subtilis, C. crescentus*, and *Streptococcus pneumoniae,* segrosomes were demonstrated to serve as the loading platform for the condensin complex, which is composed of SMC and the accessory proteins ScpA and ScpB (Sullivan et al., 2009; Wang et al., 2015; Tran et al., 2017). Large, rod-shaped, coiled-coil SMC proteins form dimers due to interactions within the hinge region and ATP-binding head domains. ATP hydrolysis and DNA binding induce conformational changes in the dimer that allow DNA loop extrusion, providing the basis for DNA compaction (Nolivos and Sherratt, 2014; Ganji et al., 2018; Baxter et al., 2019; Marko et al., 2019). Since binding and ATP hydrolysis are crucial for condensin activity, the efficiency of chromosome compaction induced by SMC proteins is presumably dependent on ATP levels in the cell. Importantly, SMC protein loading requires ParB bridging activity (Graham et al., 2014; Wilhelm et al., 2015). Upon loading in proximity to *oriC,* condensins spread along the chromosome, inducing its overall compaction and longitudinal arrangement.

Finally, an additional function of the ParB complex is its cooperation with proteins engaged in cell division regulation, such as MipZ in *C. crescentus* and *Rhodobacter sphaeroides* and PldP as well as FtsZ in *Corynebacterium glutamicum* (Donovan et al., 2010; Dubarry et al., 2019). Recent studies also shown that ParB also cooperate with NOC and both proteins are required to prevent cell division over nucleoid in *B. subtilis* (Hajduk et al., 2019) (**Table 1**).

TABLE 1 | Interaction between proteins engaged in chromosome segregation and their protein partners.

Microorganism	Segregation protein	Polar or subpolar protein	Replication protein	Chromosome organization protein	Cell division protein	Other cell cycle-involved protein
B. subtilis	Soj		DnaA (Murray and Errington, 2008)			
	Spo0J			SMC (Gruber and Errington, 2009)		
C. crescentus	ParA	TipN, PopZ (Ptacin et al., 2010; Schofield et al., 2010)				
	ParB	PopZ (Bowman et al., 2008; Ebersbach et al., 2008)			MipZ (Thanbichler and Shapiro, 2006)	
C. glutamicum	ParB	DivIVA (Donovan et al., 2012)			FtsZ, PldP (Donovan et al., 2010)	
M. smegmatis	ParA	DivIVA (Ginda et al., 2013)				DNA glycosylase (Huang and He, 2012)
M. xanthus	ParA	Bactofilin-PadC (Lin et al., 2017)				
R. sphaeroides	ParB				MipZ (Dubarry et al., 2019)	
S. coelicolor	ParA	Scy (Ditkowski et al., 2013)				ParJ (Ditkowski et al., 2010)
	ParB			TopA (Szafran et al., 2013)		
S. pneumoniae	ParB			SMC (Minnen et al., 2011)	CpsD (Nourikyan et al., 2015)	
V. cholerae	ParAI	HubP (Yamaichi et al., 2012)				

The variety of roles played by segregation proteins in the cell cycles of various bacterial species manifests in the plethora of phenotypes resulting from *parAB* deletion. While *parAB* genes were demonstrated to be essential in *C. crescentus* and *M. xanthus*, in a number of other bacterial species, including *B. subtilis*, *P. aeruginosa*, *M. smegmatis*, and *C. glutamicum,* elimination of ParA or ParB leads to chromosome segregation aberrations and mispositioning of the *oriC* region, eventually resulting in the formation of from 1 to 30% anucleate cells (recently comprehensively reviewed by Kawalek et al., 2020). In some bacteria (*V. cholerae* and *B. subtilis*), *parB* (but not *parA*) deletion increases the genomic content being manifested as elevated number of *oriC*s (Lee et al., 2003; Kadoya et al., 2011). Interestingly, *parAB* deletion may also lead to aberrations in the cell length (in *P. aeruginosa, M. smegmatis, C. glutamicum*) resulting from septum mispositioning or growth dysregulation (Donovan et al., 2013; Ginda et al., 2013). In some bacteria, *parAB* deletion results in more pleiotropic phenotypes, such as altered motility in *P. aeruginosa*, increased transformation competence in *S. pneumoniae,* reduced resistance to γ-radiation in *Deinococcus radiodurans*, and inhibited sporulation in *B. subtilis* (Errington, 2003; Lasocki et al., 2007; Bartosik et al., 2009; Charaka and Misra, 2012; Attaiech et al., 2015). Similarly, elimination of condensins has a bacterial species-dependent impact on chromosome organization. In *E. coli* and *B. subtilis,* the deletion of the genes encoding condensins results in a severe growth phenotype and chromosome segregation defects, while their deletion in other bacteria (*P. aeruginosa, M. smegmatis, S. coelicolor*) leads to a mild phenotype (reviewed by Nolivos and Sherratt, 2014). These observations reinforce the idea that segregation proteins are involved in multiple and varied cellular processes.

Other Proteins Involved in Chromosome Segregation

Interestingly, not all bacterial species employ the ParA and ParB proteins to segregate chromosomes. While the coccoid *S. pneumoniae* possesses a ParB homolog, it lacks ParA. In these bacteria, ParB cooperates with SMC proteins in chromosome segregation (Minnen et al., 2011). Moreover, some γ-proteobacteria, including *E. coli,* do not possess any ParA or ParB homologs. In contrast to ParAB-driven segregation, in *E. coli,* the segregation of newly replicated *oriC* regions is delayed by their cohesion. Cohesion is controlled by TopoIV, a type II topoisomerase, and SeqA, a protein involved in the regulation of replication initiation (Lu et al., 1994; Joshi et al., 2013; Dewachter et al., 2018). Interestingly, in response to DNA damage-induced stress, the SMC homolog RecN contributes to cohesion control (Vickridge et al., 2017). Moreover, in *E. coli,* additional *ori* domain-organizing factors were shown to contribute to positioning of the *oriC* region. These factors

include the *cis*-acting sites and *maoS* bound by the characteristic of *E. coli* MaoP protein as well as *migS* sites (Yamaichi and Niki, 2004; Valens et al., 2016; Dame et al., 2019). Finally, in *E. coli*, MukB, a structural SMC homolog strongly contributes to chromosome segregation (Hiraga et al., 1991; Yamazoe et al., 1999). Interestingly, in *E. coli,* in contrast to SMC in *C. crescentus* and *B. subtilis*, MukB does not cause chromosomal arms to adopt a longitudinal arrangement. MukB cooperates with TopoIV and the *ter* domain-organizing protein MatP (Nolivos et al., 2016; Lioy et al., 2018). MatP-dependent *ter* organization is a characteristic and unique feature of enterobacteria (Mercier et al., 2008). Whether there is any evolutionary advantage to abandoning the *parABS* system and adopting the another chromosome arrangement in enterobacteria has not yet been addressed. However, it is tempting to conclude that an elaborate life cycle and/or cell shape (e.g., *C. crescentus*) demand more complex chromosome segregation machinery.

As the last step of chromosome segregation, the separation of the duplicated chromosome *ter* regions is the final, critical checkpoint in this process. Interestingly, segregation of the *ter* regions was observed to be delayed in a number of bacterial species (Thiel et al., 2012). The segregation of *ter* regions requires the activity of accessory proteins, among which the chromosome translocase FtsK is the most widespread (Massey et al., 2006; Stouf et al., 2013; Crozat et al., 2014). The translocase family also includes the SpoIIIE protein, which is responsible for packaging of the chromosome into the small space of the forespore during *B. subtilis* sporulation, and TraB homologs involved in the conjugal transfer of DNA (Vogelmann et al., 2011; Thoma and Muth, 2015). As part of the divisiome, FtsK is associated with the newly formed septum via its N-terminal domain, while its C-terminal domain is involved in ATP hydrolysis-dependent DNA translocation as well as recombinase activation (Löwe et al., 2008; Grainge, 2013; Keller et al., 2016). FtsK activity is thus accompanied by DNA decatenation and recombination carried out by TopoIV and XerCD recombinase, respectively (El Sayyed et al., 2016). Replication and segregation of the *ter* region are tightly coordinated with Z-ring dynamics and the progression of cell division (Espéli et al., 2012; Adams et al., 2014).

Although the main players in the chromosome segregation have been identified, the mechanisms by which the activities of segregation proteins are regulated remain largely unexplored. Nevertheless, the emerging picture is that the process of the chromosome segregation is adjusted to the cell physiological state. Chromosome segregation may be regulated by modulation of the interaction between segregation proteins and DNA, nucleotide binding or other posttranslational protein modifications. Since the activities of various segregation proteins (ParA, SMC/MukB, FtsK, TopoIV) are dependent on the ATP hydrolysis, the overall energetic load of the cell, as manifested by its ATP level, should be considered as an important factor that modifies the efficiency of the segregation process. Importantly, the activities of the segregation proteins may be modulated due to their interactions with protein partners. Subsequently, these interacting partners may alter not only the efficiency of chromosome segregation but also other cell cycle parameters due to their engagement in cell division or cell elongation.

THE ParB COMPLEX—REGULATION OF ITS FORMATION AND ITS IMPACT ON CHROMOSOME DYNAMICS

To fulfill their functions, segregation proteins must interact with DNA; hence, the modification of their DNA affinity is critical for the regulation of their activity. Studies in various model bacteria have reported the modification of segrosome formation by the interaction of ParB with CTP or ParA (Breier and Grossman, 2007; Ginda et al., 2013; Baek et al., 2014; Donczew et al., 2016; Osorio-Valeriano et al., 2019; Soh et al., 2019) (**Figure 2**). Moreover, transcriptional regulation and posttranscriptional modifications of segregation proteins have been described.

Transcriptional Regulation of Segregation Genes

Formation of protein complexes may be controlled by transcriptional regulation. In *C. crescentus,* this mode of regulation applies to primary cell cycle coordinators, such as the replication initiator DnaA and the cell division protein FtsZ (Laub et al., 2000; Frandi and Collier, 2019). The transcriptional regulation exerted by master regulators such as CtrA and GcrA allows the functional differentiation of two daughter cells, which is characteristic of *C. crescentus* (Kirkpatrick and Viollier, 2012; Tsokos and Laub, 2012). While the stalked cell is capable of undertaking a new round of chromosome replication, in the swarmer cell, chromosome replication and cell division are inhibited. In *C. crescentus,* the transcription of genes encoding proteins involved in chromosome topology maintenance (topoisomerases and NAPs) was shown to be developmentally controlled, but the cell cycle-dependent transcriptional regulation of genes encoding segregation proteins has not been reported (Holtzendorff et al., 2004). Fluctuations in ParA and ParB levels between cell divisions cannot be excluded; however, there is no evidence of such regulatory mechanisms in any studied unicellular bacterium. Interestingly, the autoregulation of *parAB* was demonstrated in case of plasmid segregation proteins (Kwong et al., 2001). The only report showing the life-cycle associated transcriptional induction of chromosomal *parAB* genes comes from *S. coelicolor*, a mycelial bacteria that undergo sporulation in response to stress, particularly nutrients limitation. *Streptomyces* sporulation involves the conversion of multigenomic sporogenic hyphae into chains of unigenomic spores (Flärdh and Buttner, 2009). While during mycelial vegetative growth, cell divisions are rare and not accompanied by chromosome segregation, sporulation requires the synchronous segregation of multiple chromosomal copies synchronized with multiple cell divisions (Jakimowicz and van Wezel, 2012). In *S. coelicolor* and *S. venezuelae, parAB* genes, similarly as *ftsZ*, are controlled by cell cycle regulators associated with the onset of sporulation (WhiA and WhiB)

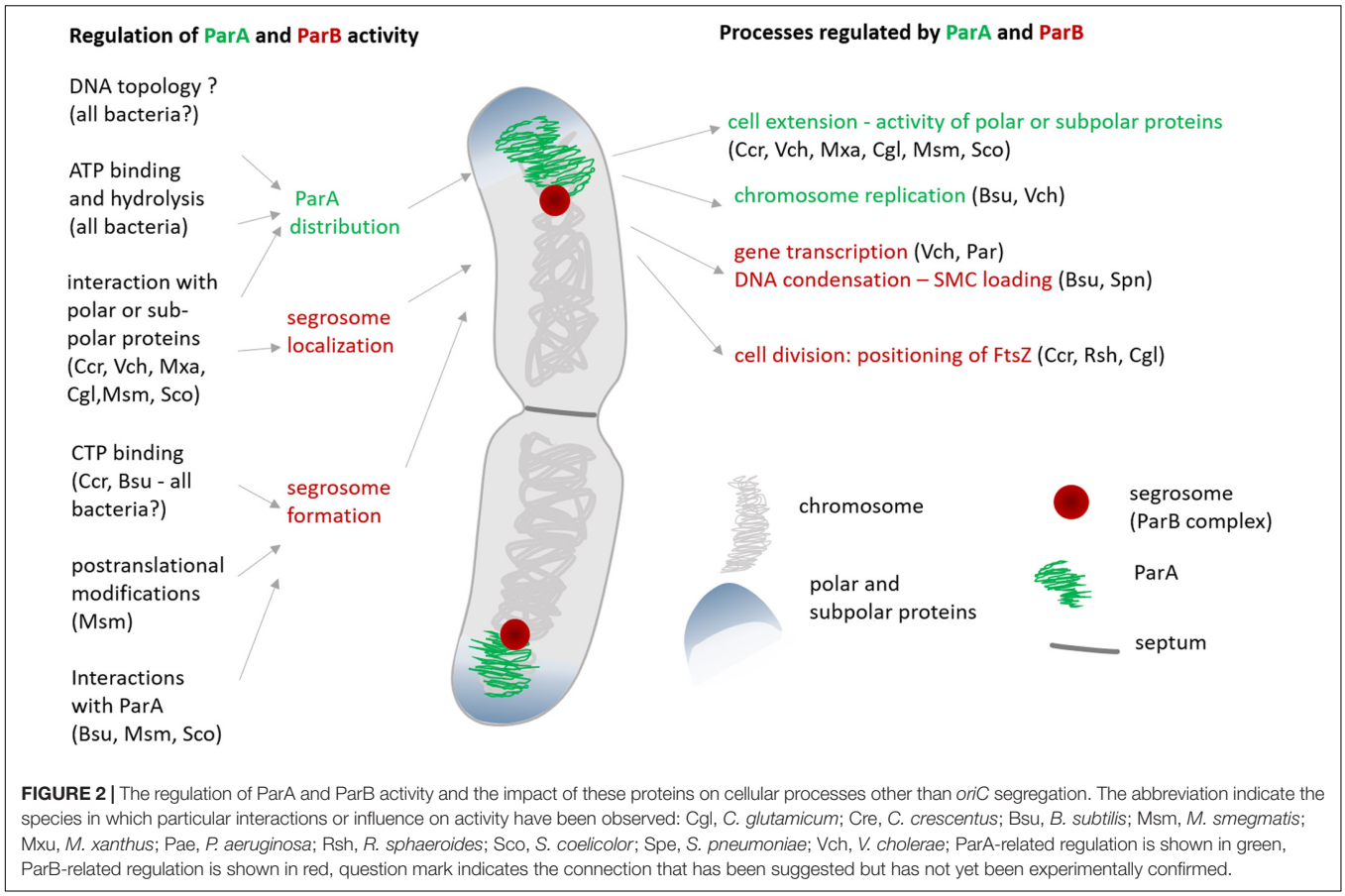


FIGURE 2 | The regulation of ParA and ParB activity and the impact of these proteins on cellular processes other than *oriC* segregation. The abbreviation indicate the species in which particular interactions or influence on activity have been observed: Cgl, *C. glutamicum*; Cre, *C. crescentus*; Bsu, *B. subtilis*; Msm, *M. smegmatis*; Mxu, *M. xanthus*; Pae, *P. aeruginosa*; Rsh, *R. sphaeroides*; Sco, *S. coelicolor*; Spe, *S. pneumoniae*; Vch, *V. cholerae*; ParA-related regulation is shown in green, ParB-related regulation is shown in red, question mark indicates the connection that has been suggested but has not yet been experimentally confirmed.

(Flärdh et al., 2000; Jakimowicz et al., 2006; Bush et al., 2013, 2016). Their upregulation allows the induction of *parAB* and *ftsZ* required for the formation of numerous ParB complexes and Z-rings, respectively, along sporogenic hyphal cells (Schwedock et al., 1997; Kim et al., 2000; Jakimowicz et al., 2005, 2007). The unprecedented transcriptional upregulation of *parAB* in *Streptomyces* fulfils the demands of their unique cell cycle and unusual mode of the chromosome segregation.

The Regulatory Role of Nucleotide Binding and Posttranslational Modifications in ParB

Little is known about the regulation of segrosome architecture; however, CTP binding by ParB was recently demonstrated to modulate interactions during ParB complex formation (Osorio-Valeriano et al., 2019; Soh et al., 2019). CTP is specifically bound within the ParB box II motif, which was previously shown to contribute to the long-range interactions (Graham et al., 2014). *ParS* binding by ParB induces CTP hydrolysis and increases the protein affinity toward the *parS* sequence, presumably leading to the complex rearrangement (Osorio-Valeriano et al., 2019; Soh et al., 2019). While the binding of CTP to ParB was demonstrated in *B. subtilis* and *M. xanthus,* the CTP interaction interface is highly conserved among ParB homologs, suggesting that this feature is preserved. Moreover, CTP binding by plasmid-encoded ParB proteins (F plasmid $ParB_F$ and P1 prophage $ParB_{P1}$) was demonstrated reinforcing the significance of nucleotide binding for complex formation (Soh et al., 2019). Since CTP is primarily involved in biosynthesis of nucleic acids and phospholipids, rather than the storage and transfer of energy, its intracellular level may reflect the cell capacity to replicate its DNA. Moreover, CTP biosynthesis is tightly regulated, and its level changes in a growth phase-dependent manner (Meng and Switzer, 2001; Jørgensen et al., 2003; Walker et al., 2004). Interestingly, in *E. coli,* one of NAP (Fis) was shown to be transcriptionally controlled by CTP, indicating the impact of this nucleotide on chromosome topology (Walker et al., 2004). Thus, the dependence of segrosome formation on subcellular CTP levels may influence functional segregation complex formation under unfavorable conditions, such as nutrients limitation, and serve as a link between cell physiology and chromosome segregation.

Furthermore, circumstantial evidence suggests that factors dependent on the cell physiological state, other than nucleotide levels, may also affect segrosome formation. Posttranslational modifications, especially phosphorylation, are well described mechanism to fine-tune the activity of the various proteins in response to changes of environmental conditions, e.g., nutrients availability (Stock et al., 1989; Bernal et al., 2014; Carabetta and Cristea, 2017; Janczarek et al., 2018). While number of nucleoid associated proteins (NAPs) including HU-like proteins were shown to be phosphorylated in *B. subtilis* and *M. tuberculosis*

and plethora of DNA organizing proteins were suggested to be target for phosphorylation, the evidence for phosphorylation of chromosome segregation proteins is limited (Gupta et al., 2014; Garcia-Garcia et al., 2016). In mycobacteria, ParB was reported to be influenced by phosphorylation, which modified protein affinity toward DNA and abolished its interaction with ParA (Baronian et al., 2015). ParB was shown to be phosphorylated *in vitro* by several eukaryotic-like Ser/Thr protein kinases whose main role is cell adaptation to changing environmental conditions (e.g., nutrient accessibility). Phosphorylation may potentially mediate fluctuations in ParB activity in relation to cell conditions; however, there is no experimental evidence of such a regulatory mechanism. Among the other posttranslational modifications, acetylation was also shown to influence the activity of numerous DNA associated proteins including topoisomerase I (in *E. coli*), DNA repair protein Ku and NAP HU (in *M. smegmatis*) (Zhou et al., 2015, 2017; Ghosh et al., 2016; Anand et al., 2017; Carabetta and Cristea, 2017), however, there are no reports of acetylation influencing directly the activity of chromosome segregation proteins ParA or ParB.

The Influence of ParA on Segresome Assembly

In some bacterial species assembly of the ParB complex was shown to be influenced by ParA. ChIP analyses of *B. subtilis*, *V. cholerae*, *S. venezuelae*, and *M. smegmatis* indicated that elimination of ParA decreased ParB binding to at least some *parS* sites (Breier and Grossman, 2007; Ginda et al., 2013; Baek et al., 2014; Donczew et al., 2016). Additionally, in *P. aeruginosa*, the impact of the ParB complex on chromosome structure was shown to be dependent on ParA (Bartosik et al., 2014). This could be explained by the lower level of ParB in *parA* deletion strains (detected in *B. subtilis* and *P. aeruginosa* but not in *S. venezuelae* and *M. smegmatis*) but may also suggest that ParA promotes ParB complex rearrangement. The latter idea is reinforced by the observed influence of ParA on ParB binding to DNA *in vitro* (Jakimowicz et al., 2007). Because ATP hydrolysis is critical for the activity of ParA, changes in intracellular ATP levels, which reflect fluctuations in the cell energetic state, may impact segrosome formation. In fact, exposure of *M. smegmatis* cells to stressful conditions modified ParA localization (Ginda et al., 2013; Pióro et al., 2019). Finally, it should be considered that environmental conditions such as increased temperature, pH or osmotic stress influence the chromosome topology (Dorman and Dorman, 2016; Qin et al., 2019), and changes of chromosome topology affect the binding of numerous DNA-interacting proteins including SMC and NAPs (Gruber and Errington, 2009; Dorman and Dorman, 2016; Tran et al., 2017; Qin et al., 2019). Consequently, the activities of segregation proteins may also be easily modified by changes in chromosome topology induced by environmental stress (**Figure 2**).

Segrosome Impact on Chromosome Topology and Gene Expression

While segrosome assembly may be adjusted in response to environmental clues, its architecture has a profound impact on chromosome structure. In *S. coelicolor,* segrosomes recruit topoisomerase I, which is required to resolve topological problems and proceed with chromosome segregation (Szafran et al., 2013). Alteration of the ParB complex architecture by changing the *parS* site position was shown to diminish *C. crescentus* fitness (Tran et al., 2018). Interestingly, in *B. subtilis* and *S. pneumoniae,* changes in the positions of *parS* sites resulted in the redistribution of ParB but had little effect on chromosome segregation and culture growth (Broedersz et al., 2014; Attaiech et al., 2015; Wang et al., 2015). However, the function of the segrosome in recruiting SMC proteins was affected by abolished ParB bridging (Gruber and Errington, 2009; Minnen et al., 2011; Graham et al., 2014; Wilhelm et al., 2015). Thus, the abovementioned studies indicate that changes in segrosome architecture influence the efficiency of chromosome compaction.

The large nucleoprotein complex formed by ParB is bound to influence chromosome topology and consequently gene expression (**Figure 2**). The first observation of segregation complex influence on gene expression was made for plasmid ParB proteins (Lynch and Wang, 1995; Rodionov et al., 1999). Considering the interspecies differences in segrosome organization, the impact of this complex on chromosome topology may be diverse. Indeed, in *S. pneumoniae*, formation of the ParB complex affects the activities of only adjacent genes, particularly the *com* operon (located 5 kb from *parS* sites), which encodes proteins involved in competence. This observation explains the increased competence of a *S. pneumoniae parB* deletion strain. Similar to *S. pneumoniae*, in *V. cholerae,* the binding of ParBI to 3 *parS* sites, results in limited ParB spreading and affects the transcription of only some of the genes in the region bound by ParB (3 of 20 genes) (Baek et al., 2014). Moreover, the transcription of several genes outside of the region bound by ParB is modified in the *parAB* deletion strain. In contrast to *S. pneumoniae* and *V. cholerae*, in *P. aeruginosa*, ParA and ParB elimination and their overexpression has been shown to affect transcription globally, influencing the expression of genes encoding stress response proteins and putative transcriptional regulators (Bartosik et al., 2014). This phenomenon was explained by *P. aeruginosa* ParB binding non-restricted to consensus *parS* sites and ability of this protein to interact with short *parS*-like motifs (Kawalek et al., 2018). This low DNA-binding specificity of ParB suggests its role in the general organization of DNA, similar to the role of NAPs. Surprisingly, in contrast to the abovementioned bacteria, in *B. subtilis,* the influence of the ParB complex on gene expression could not be detected (Breier and Grossman, 2007). Although preliminary studies suggested the involvement of *parAB* in the regulation of sporulation, this phenomenon was later shown to be independent of transcriptional regulation but was explained by the regulation of DnaA activity by ParA (see below). However, other studies reported that deletion of *parAB* in *B. subtilis* activated the SOS response by inducing a *recA* and the gene encoding the cell division inhibitor YneA (Bohorquez et al., 2018). Thus, the influence of the segrosome on the transcription of at least some genes is a common feature of the ParB complex.

COORDINATION OF THE CELL CYCLE—THE ROLE OF SEGREGATION PROTEIN INTERACTIONS

Segregation proteins interact with not only DNA and each other but also with proteins engaged in the key cell cycle processes. The ParA–DnaA interaction links chromosome segregation with chromosome replication, the interactions of ParA with polar proteins impact cell elongation, and the ParB–MipZ interaction controls the cell division (Marczynski et al., 2019) (**Figure 2**). ParA and ParB interaction partners may contribute to the chromosome segregation process; on the other hand, their activity may be controlled by ParA and/or ParB. Importantly, the DNA binding of segregation proteins modulates their availability to the partner-proteins interactions (Murray and Errington, 2008; Schofield et al., 2010; Pióro et al., 2019). Interestingly, most interactions with ParA and ParB are specific to bacterial genera, although some are more widespread and have been detected in various bacterial species.

Interactions Between Segregation and Polar or Subpolar Proteins

The interactions between segregation proteins and polar and subpolar proteins result in the specific localization of *oriC*s in *C. crescentus, V. cholerae, M. xanthus, M. smegmatis*, and *S. coelicolor*, all of which anchor the *oriC* region at their poles or subapically (Bowman et al., 2008; Ebersbach et al., 2008; Yamaichi et al., 2012; Ginda et al., 2013; Kois-Ostrowska et al., 2016; Lin et al., 2017; Pióro et al., 2019). During the asymmetric cell division of *C. crescentus*, the unidirectional chromosome segregation must be precisely controlled. The interaction between the ParB protein and the polarity factor PopZ positions the *oriC* region at the old pole before chromosome replication. PopZ is a small acidic protein that oligomerizes to form a mesh-like structure. In addition to its role in the *oriC* anchoring, its primary role is to recruit factors involved in stalk morphogenesis (Bowman et al., 2008; Ebersbach et al., 2008). Soon after the initiation of chromosome replication, one of the duplicated segrosomes is moved from the old cell pole to the new pole, and PopZ is simultaneously redistributed to form bipolar foci (Bowman et al., 2008; Ebersbach et al., 2008). Interestingly, PopZ also interacts with ParA monomers released from DNA upon ParA interaction with the ParB complex. Thus, PopZ was suggested to be involved in nucleotide exchange and the regeneration of ParA-ATP bound dimer and restoring its DNA-binding activity. Interestingly, the role of PopZ in regulating ParA activity is partially synergistic with the function of another *C. crescentus* ParA interaction partner – the coiled-coil TipN protein. TipN is mainly localized at the new pole, and the ParA–TipN interaction is critical for the ParA distribution and the directionality of segrosome movement (Lam et al., 2006; Ptacin et al., 2010). Importantly, ParA was shown to influence the function of PopZ; accumulating at the new pole ParA recruits PopZ, generating a nucleation site that initiates PopZ polymerization (Laloux and Jacobs-Wagner, 2013). It should be noted that formation of PopZ–ParA complex is dependent on availability of ParA released from nucleoid, most often by ongoing chromosome segregation (Laloux and Jacobs-Wagner, 2013). Since PopZ recruits the cell cycle regulator CtrA and its associated kinase as well as CtrA-targeting protease ClpXP (Joshi et al., 2018), the ParA control of PopZ localization possibly indicates the coordination of chromosome segregation with the global cell cycle regulation.

Similar as in *C. crescentus*, polar anchoring of *oriC* region and unidirectional chromosome segregation was described for *V. cholerae* chromosome I (the larger of the two *V. cholerae* chromosomes). In *V. cholerae,* ParAI (the ParA protein that governs the segregation of chromosome I) interacts with the polar localized protein HubP, and deletion of *hubP* abolished polarly *oriCI* positioning (Yamaichi et al., 2012). HubP also interacts with chemotactic machinery and flagellar proteins; moreover, HubP was also identified to interact with two other ATPases, ParC and FlhG (Yamaichi et al., 2012). Interestingly, in *Shewanella oneidensis* HubP homolog was also shown to be involved in chromosome segregation. Moreover, the identified in these bacteria interaction between HubP and PdeB, phosphodiesterase that controls c-di-GMP level in the cell, may indicate the potential link between cyclic nucleotide signaling and chromosome segregation (Rossmann et al., 2019). An interesting example of a bacterium in which the *oriC* is not localized at the poles but rather exhibits subpolar localization is *M. xanthus.* In this bacterium, the positioning of *oriC* is exerted by the ParA interaction with PadC, which in turn binds the bactofilin scaffold stretching from the poles (Lin et al., 2017). Only the monomeric form of ParA is recruited to the bactofilin–PadC complex, which is reminiscent of the ParA–PopZ interaction in *C. crescentus* (Lin et al., 2017). However, in case of *V. cholerae* and *M. xanthus*, there is no evidence that ParA influences the activity or localization of its interaction partners.

In the apically extending cells of actinobacteria segregation proteins also interact with polar protein complexes. The main component of the polar complex in these bacteria is the coiled-coil tropomyosin-like protein DivIVA, which recruits peptidoglycan-synthesizing machinery to the poles (Kang et al., 2008; Letek et al., 2008; Flärdh, 2010; Hammond et al., 2019). In *M. smegmatis*, DivIVA directly interacts with ParA (Ginda et al., 2013). The inhibition of this interaction was shown not only to decrease the efficiency of the chromosome segregation, but also it visibly increased the cell elongation rate, indicating ParA influence on DivIVA activity (Pióro et al., 2019). Considering that the recruitment of ParA to DivIVA was proved to compete with ParA–DNA interaction, the release of ParA from DNA upon interaction with ParB complex or, conceivably, due to changes of DNA topology, may have the impact on cell elongation rate. Markedly, DivIVA in mycobacteria is phosphorylated by the PknA kinase, the activity of which is regulated by the PknB kinase, and both PknA and PknB are essential Ser/Thr protein kinases that control growth rate and morphology (Kang et al., 2005; Jani et al., 2010; Lee et al., 2014). In response to the extracellular signals, these kinases phosphorylate regulators of central carbon metabolism and proteins involved in the stress response, transport and cell wall synthesis. It was shown that growth phase dependent DivIVA phosphorylation status regulates the rate of peptidoglycan synthesis (Jani et al., 2010). Whether the phosphorylation status of DivIVA influences its

interaction with ParA, linking environmental conditions with the segregation of chromosomes, remains to be elucidated.

Unlike *M. smegmatis*, in *S. coelicolor*, which also belongs to actinobacteria, the interaction between ParA and DivIVA was not detected, but ParA was found to directly interact with the Scy protein – the other component of a polar protein complex named the polarisome, which also includes DivIVA. During *S. coelicolor* vegetative growth, this interaction is responsible for anchoring of the apical chromosome at the tips of multigenomic hyphal cells. Importantly, the deletion of *parA* affected the rate of hyphal growth, which was explained by ParA-dependent modulation of Scy activity (Ditkowski et al., 2013; Donczew et al., 2016). Developmentally controlled ParA accumulation during sporulation leads to polarisome disassembly and inhibits hyphal elongation (Ditkowski et al., 2013). Interestingly in *S. coelicolor* ParA was shown to interact with the other segregation protein ParJ, however, the contribution of this protein to the segregation process is not fully understood (Ditkowski et al., 2010). Unlike in above described actinobacteria, in closely related *C. glutamicum* ParB directly interacts with DivIVA and this interaction positions *oriC* at the cell pole. The observation that deletion of *parB* results in impaired cell extension indicates that ParB–DivIVA interaction may impact DivIVA activity (Donovan et al., 2010, 2012). Thus, in actinobacteria, the interactions of segregation proteins with polar complexes not only contribute to chromosome segregation but also regulate cell elongation.

Interestingly, DivIVA is also involved in anchoring the *oriC* region during *B. subtilis* sporulation, though not through its direct interaction with ParAB homologs; DivIVA instead interacts with a complex containing MinD and MinJ (Kloosterman et al., 2016). Moreover, the additional RacA protein, which also specifically binds the *oriC*-proximal part of the chromosome, contributes to *oriC* anchoring in *B. subtilis* (Ben-Yehuda et al., 2005; Schumacher et al., 2016).

Interactions Between Segregation and Replication Proteins

In a number of bacterial species, ParA was also shown to be involved in the regulation of the chromosome replication. The direct interaction of ParA with DnaA was first described in *B. subtilis*, in which ParA and ParB homologs were originally identified as regulators of sporulation and called Soj and Spo0J, respectively. The elimination of Spo0J was found to inhibit sporulation and that inhibition may be counteracted by deletion of the gene encoding Soj (suppressor of Spo0J, a homolog of ParA) (Quisel et al., 1999). Later studies showed that this effect is indirect and results from Soj-dependent regulation of DnaA, which subsequently negatively regulates transcription of sporulation genes (Murray and Errington, 2008; Scholefield et al., 2011, 2012). Monomeric Soj directly interacts with DnaA and reduce its interaction with DNA inhibiting its oligomerization. In the absence of Spo0J, the DNA-bound Soj dimer is more stable, and the level of monomeric Soj available to interact with DnaA is decreased; therefore, DnaA replication activity is elevated. Thus, the function of Soj as a DnaA inhibitor depends on interaction between segregation protein and DNA. Since DNA binding by ParA homologs may be influenced by intracellular ATP level, Soj likely links the changes of the cell physiological state and environmental conditions that have the impact on cell energetic state with the DnaA replication (Murray and Errington, 2008; Scholefield et al., 2011, 2012). Additionally, in *V. cholerae*, DnaA interacts with ParA as well as ParB, while in *D. radiodurans* (another bacterium with a multipartite genome: two chromosomes and a megaplasmid), the DnaA protein interacts with ParB (Kadoya et al., 2011; Maurya et al., 2019). The involvement of the segregation proteins in replication regulation explains the increased number of *oriC* resulting from *parB* deletion.

Interactions Between Segregation and Cell Division Proteins

Chromosome segregation proteins also interact with the cell division proteins. In the α-proteobacteria *C. crescentus* and *R. sphaeroides*, the interaction between the ParB protein and MipZ (a ParA homolog) was detected (Thanbichler and Shapiro, 2006; Dubarry et al., 2019). MipZ is an inhibitor of FtsZ polymerization that exhibits dynamic localization characteristic of the ParA family of ATPases. In *C. crescentus,* MipZ forms a cloud-like structure with the lowest MipZ concentration in the middle of the cell, which restricts Z-ring formation to the cell center (Thanbichler and Shapiro, 2006). The localization of MipZ in *R. sphaeroides* is different; MipZ is situated mainly at the cell poles but also at the mid-cell position. In both *C. crescentus* and *R. sphaeroides*, the localization of MipZ depends on ParB, but unlike ParA, MipZ dimers are recruited and stabilized by ParB (Thanbichler and Shapiro, 2006; Dubarry et al., 2019). Importantly, in *C. crescentus,* the transcription level of *mipZ* changes during cell cycle progression and in response to environmental cues (e.g., nitrogen starvation) (Collier, 2019). In *C. glutamicum,* ParB also interacts with PldP – a ParA homolog involved in the regulation of cell division (Donovan et al., 2010). Moreover, in these bacteria, the direct interaction between ParB and FtsZ—a cell division initiator—was shown (Donovan et al., 2010). These interactions presumably account for the observed influence of *parAB* deletion on septum placement (Donovan and Bramkamp, 2014).

Interestingly, in *S. pneumoniae* (which lacks the ParA component of the *parABS* system), ParB interacts with CpsD, which is homologous to ParA tyrosine (BY) kinase and localizes at the site of cell division (Bender and Yother, 2001). BY-kinases are autokinases that regulate polymerization and the export of capsular polysaccharides. Inhibition of CpsD phosphorylation delayed chromosome segregation, while increased CpsD phosphorylation enhanced ParB mobility. The interaction between ParB and CpsD may coordinate chromosome segregation with capsular formation and the cell division (Nourikyan et al., 2015). Recent studies identified another ParB interacting partner in *S. pneumoniae*—the RocS protein. RocS is required for chromosome segregation but also interacts with FtsZ and CspD (Mercy et al., 2019). The above examples show that chromosome segregation and cell division are coupled due to the interactions of segregation proteins.

Some of the interactions of segregation proteins were shown to be critical under stress conditions. These include discovered

in *M. smegmatis* interaction between ParA and 3-methyladenine DNA glycosylase, a protein mainly involved in DNA repair (Huang and He, 2012). This interaction stimulates the ATPase activity of ParA and regulates cell growth and morphology independent of DNA glycosylase activity. In *B. subtilis,* ParAB was shown to cooperate with another segregation protein, WhiA, which was suggested to maintain DNA integrity (Bohorquez et al., 2018). Interestingly, double *parAB*/*whiA* deletion was lethal and could be explained by the blockade of cell division.

CONCLUDING REMARKS

Studies of last two decades have shed light on chromosome segregation, revealing the concerted actions of segregation proteins, dissecting the mechanisms of their activities and describing their interactions. However, evidence that the chromosome segregation process is adjusted to environmental conditions has only started to emerge. Environmental stress factors, such as nutrients limitation, modify cell physiology and require adjustment of the cell cycle process. The possible pathways that can be used to coordinate the cell cycle with stress response are those based on intracellular nucleotide levels and chromosome topology. In particular, the finding that the ParB–CTP interaction is a critical factor for segrosome formation opens a new avenue for the exploration of chromosome segregation regulation. Finally, the impact of polar proteins (TipN, PopZ, HubP, or DivIVA) on the activities of segregation proteins under unfavorable conditions also remains to be further investigated to identify the links between changes in cell physiology and chromosome segregation.

Furthermore, recent studies have indicated the impact of chromosome segregation proteins on other cell cycle processes. Interestingly, due to their involvement in highly species-specific interactions (including both DNA interactions during segrosome formation and protein–protein interactions), the involvement of segregation proteins in coordination of the cell cycle is diverse and species dependent. Common regulatory pathways (identified in at least two unrelated organisms) include the regulation of gene transcription, chromosome replication, and the regulation of the cell elongation and division (Marczynski et al., 2019). Notably, the availability of ParA to interact with their protein partners (DnaA, PopZ, DivIVA) depend on the ParA binding to chromosome. Since this interaction is plausibly influenced by environmental factors, it may serve as the important regulatory circuit. However, further studies are required to fully understand the complex regulatory networks behind the identified connections and the impact of external factors on the global coordination of cell cycle processes.

AUTHOR CONTRIBUTIONS

MP and DJ wrote the manuscript.

FUNDING

The authors acknowledge financial support from the National Science Center, Poland (OPUS grant 2017/27/B/NZ1/00823).

ACKNOWLEDGMENTS

We are grateful to Jolanta Zakrzewska-Czerwińska, Marcin Szafran, and Joanna Hołówka for critical reading and comments on the manuscript.

REFERENCES

Adams, D. W., Wu, L. J., and Errington, J. (2014). Cell cycle regulation by the bacterial nucleoid. *Curr. Opin. Microbiol.* 22C, 94–101. doi: 10.1016/j.mib.2014.09.020

Anand, C., Garg, R., Ghosh, S., and Nagaraja, V. (2017). A Sir2 family protein Rv1151c deacetylates HU to alter its DNA binding mode in *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* 493, 1204–1209. doi: 10.1016/j.bbrc.2017.09.087

Attaiech, L., Minnen, A., Kjos, M., Gruber, S., and Veening, J. W. (2015). The ParB-parS chromosome segregation system modulates competence development in *Streptococcus pneumoniae*. *mBio* 6:e00662-15. doi: 10.1128/mBio.00662-15

Autret, S., Nair, R., and Errington, J. (2001). Genetic analysis of the chromosome segregation protein Spo0J of *Bacillus subtilis*?: evidence for separate domains involved in DNA binding and interactions with Soj protein. *Mol. Microbiol.* 41, 743–755. doi: 10.1046/j.1365-2958.2001.02551.x

Badrinarayanan, A., Le, T. B. K., and Laub, M. T. (2015). Bacterial Chromosome Organization and Segregation. *Annu. Rev Cell Dev. Biol.* 31, 171–199. doi: 10.1146/annurev-cellbio-100814-125211

Baek, J. H., Rajagopala, S. V., and Chattoraj, D. K. (2014). Chromosome segregation proteins of *Vibrio cholerae* as transcription regulators. *mBio* 5:e01061-14. doi: 10.1128/mBio.01061-14

Baronian, G., Ginda, K., Berry, L., Cohen-Gonsaud, M., Zakrzewska-Czerwińska, J., Jakimowicz, D., et al. (2015). Phosphorylation of *Mycobacterium tuberculosis* ParB participates in regulating the ParABS chromosome segregation system. *PLoS One* 10:e0119907. doi: 10.1371/journal.pone.0119907

Bartosik, A., Mierzejewska, J., Thomas, C. M., and Jagura-Burdzy, G. (2009). ParB deficiency in *Pseudomonas aeruginosa* destabilizes the partner protein ParA and affects a variety of physiological parameters. *Microbiology* 155, 1080–1092. doi: 10.1099/mic.0.024661-0

Bartosik, A. A., Glabski, K., Jecz, P., Mikulska, S., Fogtman, A., Koblowska, M., et al. (2014). Transcriptional profiling of para and ParB mutants in actively dividing cells of an opportunistic human pathogen *Pseudomonas aeruginosa*. *PLoS One* 9:e87276. doi: 10.1371/journal.pone.0087276

Baxter, J., Oliver, A. W., and Schalbetter, S. A. (2019). Are SMC complexes loop extruding factors? Linking theory with fact. *Bioessays* 41:1800182. doi: 10.1002/bies.201800182

Bender, M. H., and Yother, J. (2001). CpsB is a modulator of capsule-associated tyrosine kinase activity in *Streptococcus pneumoniae*. *J. Biol. Chem.* 276, 47966–47974. doi: 10.1074/jbc.M105448200

Ben-Yehuda, S., Fujita, M., Liu, X. S., Gorbatyuk, B., Skoko, D., Yan, J., et al. (2005). Defining a centromere-like element in *Bacillus subtilis* by identifying the binding sites for the chromosome-anchoring protein RacA. *Mol. Cell* 17, 773–782. doi: 10.1016/j.molcel.2005.02.023

Bernal, V., Castaño-Cerezo, S., Gallego-Jara, J., Écija-Conesa, A., de Diego, T., Iborra, J. L., et al. (2014). Regulation of bacterial physiology by lysine acetylation of proteins. *N. Biotechnol.* 31, 586–595. doi: 10.1016/j.nbt.2014.03.002

Bohorquez, L. C., Surdova, K., Jonker, M. J., and Hamoen, L. W. (2018). The conserved DNA binding protein WhiA influences chromosome. *J. Bacteriol.* 200:e00633-17. doi: 10.1128/JB.00633-17

Bowman, G. R., Comolli, L. R., Zhu, J., Eckart, M., Koenig, M., Downing, K. H., et al. (2008). A polymeric protein anchors the chromosomal Origin/ParB complex at a bacterial cell pole. *Cell* 134, 945–955. doi: 10.1016/j.cell.2008.07.015

Breier, A. M., and Grossman, A. D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol. Microbiol.* 64, 703–718. doi: 10.1111/j.1365-2958.2007.05690.x

Broedersz, C. P., Wang, X., Meir, Y., Loparo, J. J., Rudner, D. Z., and Wingreen, N. S. (2014). Condensation and localization of the partitioning protein ParB on the bacterial chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8809–8814. doi: 10.1073/pnas.1402529111

Burby, P. E., and Simmons, L. A. (2020). Regulation of cell division in bacteria by monitoring genome integrity and DNA replication status. *J. Bacteriol.* 202:e00408-19. doi: 10.1128/jb.00408-19

Bush, M. J., Bibb, M. J., Chandra, G., Findlay, K. C., and Buttner, M. J. (2013). Genes required for aerial growth, cell division, and chromosome segregation are targets of whia before sporulation in *Streptomyces venezuelae*. *mBio* 4:e00684-13. doi: 10.1128/mBio.00684-13

Bush, M. J., Chandra, G., Bibb, M. J., Findlay, K. C., and Buttner, M. J. (2016). Genome-wide chromatin immunoprecipitation sequencing analysis shows that WhiB is a transcription factor that cocontrols its regulon with WhiA to initiate developmental cell division in Streptomyces. *mBio* 7:e00523-16. doi: 10.1128/mBio.00523-16

Carabetta, V. J., and Cristea, I. M. (2017). Regulation, function, and detection of protein acetylation in bacteria. *J. Bacteriol.* 199:e00107-17. doi: 10.1128/JB.00107-17

Charaka, V. K., and Misra, H. S. (2012). Functional characterization of the role of the chromosome I partitioning system in genome segregation in *Deinococcus radiodurans*. *J. Bacteriol.* 194, 5739–5748. doi: 10.1128/JB.00610-12

Chen, B.-W., Lin, M.-H. H., Chu, C.-H. H., Hsu, C.-E. E., and Sun, Y.-J. J. (2015). Insights into ParB spreading from the complex structure of Spo0J and parS. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6613–6618. doi: 10.1073/pnas.1421927112

Collier, J. (2019). Cell division control in *Caulobacter crescentus*. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 685–690. doi: 10.1016/j.bbagrm.2018.04.005

Crozat, E., Rousseau, P., Fournes, F., and Cornet, F. (2014). The FtsK Family of DNA translocases finds the ends of circles. *J. Mol. Microbiol. Biotechnol.* 24, 396–408. doi: 10.1159/000369213

Dame, R. T., Rashid, F.-Z. M., and Grainger, D. C. (2019). Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.* 12, 1–16. doi: 10.1038/s41576-019-0185-4

Desai, S. K., and Kenney, L. J. (2019). Switching lifestyles is an *in vivo* adaptive strategy of bacterial pathogens. *Front. Cell. Infect. Microbiol.* 9:421. doi: 10.3389/fcimb.2019.00421

Dewachter, L., Verstraeten, N., Fauvart, M., and Michiels, J. (2018). An integrative view of cell cycle control in *Escherichia coli*. *FEMS Microbiol. Rev.* 42, 116–136. doi: 10.1093/femsre/fuy005

Ditkowski, B., Holmes, N., Rydzak, J., Donczew, M., Bezulska, M., Ginda, K., et al. (2013). Dynamic interplay of ParA with the polarity protein, Scy, coordinates the growth with chromosome segregation in *Streptomyces coelicolor*. *Open Biol.* 3:130006. doi: 10.1098/rsob.130006

Ditkowski, B., Troc, P., Ginda, K., Donczew, M., Chater, K. F. K. F., Zakrzewska-Czerwińska, J., et al. (2010). The actinobacterial signature protein ParJ (SCO1662) regulates ParA polymerization and affects chromosome segregation and cell division during *Streptomyces sporulation*. *Mol. Microbiol.* 78, 1403–1415. doi: 10.1111/j.1365-2958.2010.07409.x

Donczew, M., Mackiewicz, P., Wrobel, A., Flardh, K., Zakrzewska-Czerwínska, J., Jakimowicz, D., et al. (2016). ParA and ParB coordinate chromosome segregation with cell elongation and division during *Streptomyces sporulation*. *Open Biol.* 6:150263. doi: 10.1098/rsob.150263

Donovan, C., and Bramkamp, M. (2014). Cell division in Corynebacterineae. *Front. Microbiol.* 5:132. doi: 10.3389/fmicb.2014.00132

Donovan, C., Schauss, A., Krämer, R., and Bramkamp, M. (2013). Chromosome segregation impacts on cell growth and division site selection in *Corynebacterium glutamicum*. *PLoS One* 8:e55078. doi: 10.1371/journal.pone.0055078

Donovan, C., Schwaiger, A., Krämer, R., and Bramkamp, M. (2010). Subcellular localization and characterization of the ParAB system from *Corynebacterium glutamicum*. *J. Bacteriol.* 192, 3441–3451. doi: 10.1128/JB.00214-10

Donovan, C., Sieger, B., Krämer, R., and Bramkamp, M. (2012). A synthetic *Escherichia coli* system identifies a conserved origin tethering factor in Actinobacteria. *Mol. Microbiol.* 84, 105–116. doi: 10.1111/j.1365-2958.2012.08011.x

Dorman, C. J., and Dorman, M. J. (2016). DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys. Rev.* 8, 89–100. doi: 10.1007/s12551-016-0238-2

Dubarry, N., Willis, C. R., Ball, G., Lesterlin, C., and Armitage, J. P. (2019). In vivo imaging of the segregation of the 2 chromosomes and the cell division proteins of rhodobacter sphaeroides reveals an unexpected role for MipZ. *mBio* 10:e2515-18. doi: 10.1128/mBio.02515-18

Ebersbach, G., Briegel, A., Jensen, G. J., and Jacobs-Wagner, C. (2008). A self-associating protein critical for chromosome attachment, division, and polar organization in caulobacter. *Cell* 134, 956–968. doi: 10.1016/j.cell.2008.07.016

El Sayyed, H., Le Chat, L., Lebailly, E., Vickridge, E., Pages, C., Cornet, F., et al. (2016). Mapping topoisomerase IV binding and activity sites on the *E. coli* Genome. *PLoS Genet.* 12:e1006025. doi: 10.1371/journal.pgen.1006025

Errington, J. (2001). Septation and chromosome segregation during sporulation in *Bacillus subtilis*. *Curr. Opin. Microbiol.* 4, 660–666. doi: 10.1016/s1369-5274(01)00266-1

Errington, J. (2003). Regulation of endospore formation in *Bacillus subtilis*. *Nat. Rev. Microbiol.* 1, 117–126. doi: 10.1038/nrmicro750

Espéli, O., Borne, R., Dupaigne, P., Thiel, A., Gigant, E., Mercier, R., et al. (2012). A MatP-divisome interaction coordinates chromosome segregation with cell division in *E. coli*. *EMBO J.* 31, 3198–3211. doi: 10.1038/emboj.2012.128

Fisher, G. L. M., Pastrana, C. L., Higman, V. A., Koh, A., Taylor, J. A., Butterer, A., et al. (2017). The structural basis for dynamic DNA binding and bridging interactions which condense the bacterial centromere. *eLife* 6:e28086. doi: 10.7554/eLife.28086

Flärdh, K. (2010). Cell polarity and the control of apical growth in Streptomyces. *Curr. Opin. Microbiol.* 13, 758–765. doi: 10.1016/j.mib.2010.10.002

Flärdh, K., and Buttner, M. J. (2009). Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat. Rev. Microbiol.* 7, 36–49. doi: 10.1038/nrmicro1968

Flärdh, K., Leibovitz, E., Buttner, M. J., and Chater, K. F. (2000). Generation of a non-sporulating strain of *Streptomyces coelicolor* A3(2) by the manipulation of a developmentally controlled ftsZ promoter. *Mol. Microbiol.* 38, 737–749. doi: 10.1046/j.1365-2958.2000.02177.x

Frandi, A., and Collier, J. (2019). Multilayered control of chromosome replication in *Caulobacter crescentus*. *Biochem. Soc. Trans.* 47, 187–196. doi: 10.1042/BST20180460

Funnell, B. E. (2014). How to build segregation complexes in bacteria: use bridges. *Genes Dev.* 28, 1140–1142. doi: 10.1101/gad.244517.114

Funnell, B. E. (2016). ParB partition proteins: complex formation and spreading at bacterial and plasmid centromeres. *Front. Mol. Biosci.* 3:44. doi: 10.3389/fmolb.2016.00044

Ganji, M., Shaltiel, I. A., Bisht, S., Kim, E., Kalichava, A., Haering, C. H., et al. (2018). Real-time imaging of DNA loop extrusion by condensin. *Science* 360, 102–105. doi: 10.1126/science.aar7831

Garcia-Garcia, T., Poncet, S., Derouiche, A., Shi, L., Mijakovic, I., and Noirot-Gros, M. F. (2016). Role of protein phosphorylation in the regulation of cell cycle and DNA-related processes in bacteria. *Front. Microbiol.* 7:184. doi: 10.3389/fmicb.2016.00184

Ghosh, S., Padmanabhan, B., Anand, C., and Nagaraja, V. (2016). Lysine acetylation of the *Mycobacterium tuberculosis* HU protein modulates its DNA binding and genome organization. *Mol. Microbiol.* 100, 577–588. doi: 10.1111/mmi.13339

Ginda, K., Bezulska, M., Ziółkiewicz, M., Dziadek, J., Zakrzewska-Czerwińska, J., and Jakimowicz, D. (2013). ParA of *Mycobacterium smegmatis* co-ordinates

chromosome segregation with the cell cycle and interacts with the polar growth determinant DivIVA. *Mol. Microbiol.* 87, 998–1012. doi: 10.1111/mmi.12146

Gottesman, S. (2019). Trouble is coming: signaling pathways that regulate general stress responses in bacteria. *J. Biol. Chem.* 294, 11685–11700. doi: 10.1074/jbc.REV119.005593

Graham, T. G. W., Wang, X., Song, D., Etson, C. M., Van Oijen, A. M., Rudner, D. Z., et al. (2014). ParB spreading requires DNA bridging. *Genes Dev.* 28, 1228–1238. doi: 10.1101/gad.242206.114

Grainge, I. (2013). Simple topology: FtsK-directed recombination at the dif site. *Biochem. Soc. Trans.* 41, 595–600. doi: 10.1042/BST20120299

Graumann, P. L., Losick, R., and Strunnikov, A. V. (1998). Subcellular localization of *Bacillus subtilis* SMC, a protein involved in chromosome condensation and segregation. *J. Bacteriol.* 180, 5749–5755. doi: 10.1128/jb.180.21.5749-5755.1998

Gruber, S., and Errington, J. (2009). Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell* 137, 685–696. doi: 10.1016/j.cell.2009.02.035

Gupta, M., Sajid, A., Sharma, K., Ghosh, S., Arora, G., Singh, R., et al. (2014). HupB, a nucleoid-associated protein of *mycobacterium tuberculosis*, is modified by serine/threonine protein kinases *in vivo*. *J. Bacteriol.* 196, 2646–2657. doi: 10.1128/JB.01625-14

Hajduk, I. V., Mann, R., Rodrigues, C. D. A., and Harry, E. J. (2019). The ParB homologs, Spo0J and Noc, together prevent premature midcell Z ring assembly when the early stages of replication are blocked in *Bacillus subtilis*. *Mol. Microbiol.* 112, 766–784. doi: 10.1111/mmi.14319

Hammond, L. R., White, M. L., and Eswara, P. J. (2019). vIVA la DivIVA! *J. Bacteriol.* [Epub ahead of print].

Harms, A., Treuner-Lange, A., Schumacher, D., and Søgaard-Andersen, L. (2013). Tracking of chromosome and replisome dynamics in *Myxococcus xanthus* reveals a novel chromosome arrangement. *PLoS Genet.* 9:e1003802. doi: 10.1371/journal.pgen.1003802

Heinrich, K., Leslie, D. J., and Jonas, K. (2015). Modulation of bacterial proliferation as a survival strategy. *Adv. Appl. Microbiol.* 92, 127–171. doi: 10.1016/bs.aambs.2015.02.004

Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Microbiol.* 7, 263–273. doi: 10.1038/nrmicro2109

Hester, C. M., and Lutkenhaus, J. (2007). Soj (ParA) DNA binding is mediated by conserved arginines and is essential for plasmid segregation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20326–20331. doi: 10.1073/pnas.0705196105

Hiraga, S., Niki, H., Imamura, R., Ogura, T., Yamanaka, K., Feng, J., et al. (1991). Mutants defective in chromosome partitioning in *E. coli*. *Res. Microbiol.* 142, 189–194. doi: 10.1016/0923-2508(91)90029-A

Holtzendorff, J., Hung, D., Brende, P., Reisenauer, A., Viollier, P. H., McAdams, H. H., et al. (2004). Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science* 304, 983–987. doi: 10.1126/science.1095191

Huang, F., and He, Z.-G. G. (2012). Characterization of a conserved interaction between DNA glycosylase and ParA in *Mycobacterium smegmatis* and *M. tuberculosis*. *PLoS One* 7:e38276. doi: 10.1371/journal.pone.0038276

Jakimowicz, D., Chater, K., and Zakrzewska-Czerwínska, J. (2002). The ParB protein of *Streptomyces coelicolor* A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome. *Mol. Microbiol.* 45, 1365–1377. doi: 10.1046/j.1365-2958.2002.03102.x

Jakimowicz, D., Gust, B., Zakrzewska-Czerwińska, J., Chater, K. F. K. F., Zakrzewska-Czerwinska, J., and Chater, K. F. K. F. (2005). Developmental-stage-specific assembly of ParB complexes in *Streptomyces coelicolor* hyphae. *J. Bacteriol.* 187, 3572–3580. doi: 10.1128/JB.187.10.3572

Jakimowicz, D., Mouz, S., Zakrzewska-Czerwińska, J., and Chater, K. F. K. F. (2006). Developmental control of a parAB promoter leads to formation of sporulation-associated ParB complexes in *Streptomyces coelicolor*. *J. Bacteriol.* 188, 1710–1720. doi: 10.1128/JB.188.5.1710

Jakimowicz, D., and van Wezel, G. P. G. P. (2012). Cell division and DNA segregation in *Streptomyces*: how to build a septum in the middle of nowhere? *Mol. Microbiol.* 85, 393–404. doi: 10.1111/j.1365-2958.2012.08107.x

Jakimowicz, D., Zydek, P., Kois, A., Zakrzewska-Czerwińska, J., and Chater, K. F. K. F. (2007). Alignment of multiple chromosomes along helical ParA scaffolding in sporulating *Streptomyces* hyphae. *Mol. Microbiol.* 65, 625–641. doi: 10.1111/j.1365-2958.2007.05815.x

Jalal, A. S. B., Pastrana, C. L., Tran, N. T., Stevenson, C. E., Lawson, D. M., Moreno-Herrero, F., et al. (2019). Structural and biochemical analyses of *Caulobacter crescentus* ParB reveal the role of its N- terminal domain in chromosome segregation Adam. *bioRxiv* [Preprint]. doi: 10.1101/816959

Janczarek, M., Vinardell, J. M., Lipa, P., and Karaś, M. (2018). Hanks-type serine/threonine protein kinases and phosphatases in bacteria: roles in signaling and adaptation to various environments. *Int. J. Mol. Sci.* 19:2872. doi: 10.3390/ijms19102872

Jani, C., Eoh, H., Lee, J. J., Hamasha, K., Sahana, M. B., Han, J.-S., et al. (2010). Regulation of polar peptidoglycan biosynthesis by Wag31 Phosphorylation in Mycobacteria. *BMC Microbiol.* 10:327. doi: 10.1186/1471-2180-10-327

Jones, T. H., Vail, K. M., and McMullen, L. M. (2013). Filament formation by foodborne bacteria under sublethal stress. *Int. J. Food Microbiol.* 165, 97–110. doi: 10.1016/j.ijfoodmicro.2013.05.001

Jørgensen, C. M., Hammer, K., and Martinussen, J. (2003). CTP limitation increases expression of CTP synthase in *Lactococcus lactis*. *J. Bacteriol.* 185, 6562–6574. doi: 10.1128/JB.185.22.6562-6574.2003

Joshi, K. K., Battle, C. M., and Chien, P. (2018). Polar localization hub protein PopZ restrains adaptor-dependent ClpXP proteolysis in *Caulobacter crescentus*. *J. Bacteriol.* 200:e00221-18. doi: 10.1128/JB.00221-18

Joshi, M. C., Magnan, D., Montminy, T. P., Lies, M., Stepankiw, N., and Bates, D. (2013). Regulation of sister chromosome cohesion by the replication fork tracking protein SeqA. *PLoS Genet.* 9:e1003673. doi: 10.1371/journal.pgen.1003673

Kadoya, R., Baek, J. H., Sarker, A., and Chattoraj, D. K. (2011). Participation of chromosome segregation protein ParAI of *Vibrio cholerae* in chromosome replication. *J. Bacteriol.* 193, 1504–1514. doi: 10.1128/JB.01067-10

Kang, C., Abbott, D. W., Park, S. T., Dascher, C. C., Cantley, L. C., and Husson, R. N. (2005). The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev.* 19, 1692–1704. doi: 10.1101/gad.1311105.nism

Kang, C.-M. M., Nyayapathy, S., Lee, J.-Y. Y., Suh, J.-W. W., and Husson, R. N. (2008). Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. *Microbiology* 154, 725–735. doi: 10.1099/mic.0.2007/0140760

Kato, J. I., Nishimura, Y., Imamura, R., Niki, H., Hiraga, S., and Suzuki, H. (1990). New topoisomerase essential for chromosome segregation in *E. coli*. *Cell* 63, 393–404. doi: 10.1016/0092-8674(90)90172-B

Kawalek, A., Bartosik, A. A., Glabski, K., and Jagura-Burdzy, G. (2018). *Pseudomonas aeruginosa* partitioning protein ParB acts as a nucleoid-associated protein binding to multiple copies of a parS-related motif. *Nucleic Acids Res.* 46, 4592–4606. doi: 10.1093/nar/gky257

Kawalek, A., Wawrzyniak, P., Bartosik, A. A., and Jagura-Burdzy, G. (2020). Rules and exceptions?: the role of chromosomal ParB in DNA segregation and other cellular processes. *Microorganisms* 8:105. doi: 10.3390/microorganisms8010105

Keller, A. N., Xin, Y., Boer, S., Reinhardt, J., Baker, R., Arciszewska, L. K., et al. (2016). Activation of Xer-recombination at dif: structural basis of the FtsKγ-XerD interaction. *Sci. Rep.* 6:33357. doi: 10.1038/srep33357

Kim, H. J., Calcutt, M. J., Schmidt, F. J., and Chater, K. F. (2000). Partitioning of the linear chromosome during sporulation of *Streptomyces coelicolor* A3(2) involves an oriC-linked parAB locus. *J. Bacteriol.* 182, 1313–1320. doi: 10.1128/jb.182.5.1313-1320.2000

Kirkpatrick, C. L., and Viollier, P. H. (2012). Decoding Caulobacter development. *FEMS Microbiol. Rev.* 36, 193–205. doi: 10.1111/j.1574-6976.2011.00309.x

Kleckner, N., Fisher, J. K., Stouf, M., White, M. A., Bates, D., and Witz, G. (2014). The bacterial nucleoid: nature, dynamics and sister segregation. *Curr. Opin. Microbiol.* 22, 127–137. doi: 10.1016/j.mib.2014.10.001

Kloosterman, T. G., Lenarcic, R., Willis, C., Roberts, D. M., and Hamoen, L. W. (2016). Complex polar machinery required for proper chromosome segregation in vegetative and sporulating cells of *Bacillus subtilis*. *Mol. Microbiol.* 101, 333–350. doi: 10.1111/mmi.13393

Kois-Ostrowska, A., Strzałka, A., Lipietta, N., Tilley, E., Zakrzewska-Czerwińska, J., Herron, P. R., et al. (2016). Unique Function of the Bacterial Chromosome Segregation Machinery in Apically Growing Streptomyces - Targeting the Chromosome to New Hyphal Tubes and its Anchorage at the Tips. *PLoS Genet.* 12:e1006488. doi: 10.1371/journal.pgen.1006488

Kusiak, M., Gapczynska, A., Plochocka, D., Thomas, C. M., Jagura-Burdzy, G., Gapczyñska, A., et al. (2011). Binding and spreading of ParB on DNA determine

its biological function in *Pseudomonas aeruginosa*. *J. Bacteriol.* 193, 3342–3355. doi: 10.1128/JB.00328-11

Kwong, S. M., Chew, C. Y., and Chit, L. P. (2001). Molecular analysis of the pRA2 partitioning region: ParB autoregulates parAB transcription and forms a nucleoprotein complex with the plasmid partition site, parS. *Mol. Microbiol.* 40, 621–633. doi: 10.1046/j.1365-2958.2001.02405.x

Laloux, G., and Jacobs-Wagner, C. (2013). Spatiotemporal control of PopZ localization through cell cycle-coupled multimerization. *J. Cell Biol.* 201, 827–841. doi: 10.1083/jcb.201303036

Lam, H., Schofield, W. B., and Jacobs-Wagner, C. (2006). A landmark protein essential for establishing and perpetuating the polarity of a bacterial cell. *Cell* 124, 1011–1023. doi: 10.1016/j.cell.2005.12.040

Lasocki, K., Bartosik, A. A., Mierzejewska, J., Thomas, C. M., and Jagura-Burdzy, G. (2007). Deletion of the parA (soj) homologue in *Pseudomonas aeruginosa* causes ParB instability and affects growth rate, chromosome segregation, and motility. *J. Bacteriol.* 189, 5762–5772. doi: 10.1128/JB.00371-07

Latoscha, A., Wörmann, M. E., and Tschowri, N. (2019). Nucleotide second messengers in streptomyces. *Microbiology* 165, 1153–1165. doi: 10.1099/mic.0.000846

Laub, M. T., McAdams, H. H., Feldblyum, T., Fraser, C. M., and Shapiro, L. (2000). Global analysis of the genetic network controlling a bacterial cell cycle. *Science* 290, 2144–2148. doi: 10.1126/science.290.5499.2144

Le Gall, A., Cattoni, D. I., Guilhas, B., Mathieu-Demazière, C., Oudjedi, L., Fiche, J.-B., et al. (2016). Bacterial partition complexes segregate within the volume of the nucleoid. *Nat. Commun.* 7:12107. doi: 10.1038/ncomms12107

Lee, J. J., Kang, C. M., Lee, J. H., Park, K. S., Jeon, J. H., and Lee, S. H. (2014). Phosphorylation-dependent interaction between a serine/threonine kinase PknA and a putative cell division protein Wag31 in *Mycobacterium tuberculosis*. *New Microbiol.* 37, 525–533.

Lee, P. S., Lin, D. C. H., Moriya, S., and Grossman, A. D. (2003). Effects of the chromosome partitioning protein Spo0J (ParB) on oriC positioning and replication initiation in *Bacillus subtilis*. *J. Bacteriol.* 185, 1326–1337. doi: 10.1128/JB.185.4.1326-1337.2003

Leonard, T. A., Butler, P. J., and Lowe, J. (2005). Bacterial chromosome segregation?: structure and DNA binding of the Soj dimer — a conserved biological switch. *EMBO J.* 24, 270–282. doi: 10.1038/sj.emboj.7600530

Leonard, T. A., Butler, P. J. G., and Löwe, J. (2004). Structural analysis of the chromosome segregation protein Spo0J from Thermus thermophilus. *Mol. Microbiol.* 53, 419–432. doi: 10.1111/j.1365-2958.2004.04133.x

Letek, M., Ordóñez, E., Vaquera, J., Margolin, W., Flärdh, K., Mateos, L. M., et al. (2008). DivIVA is required for polar growth in the MreB-lacking rod-shaped actinomycete *Corynebacterium glutamicum*. *J. Bacteriol.* 190, 3283–3292. doi: 10.1128/JB.01934-07

Lim, H. C., Surovtsev, I. V., Beltran, B. G., Huang, F., Bewersdorf, J., and Jacobs-Wagner, C. (2014). Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *eLife* 3:e02758. doi: 10.7554/eLife.02758

Lin, L., Osorio Valeriano, M., Harms, A., Søgaard-Andersen, L., and Thanbichler, M. (2017). Bactofilin-mediated organization of the ParABS chromosome segregation system in *Myxococcus xanthus*. *Nat. Commun.* 8:1817. doi: 10.1038/s41467-017-02015-z

Lioy, V. S., Cournac, A., Koszul, R., Mozziconacci, J., Espeli, O., Boccard, F., et al. (2018). Multiscale structuring of the *E. coli* chromosome by nucleoid-associated and condensin proteins. *Cell* 172, 1–13. doi: 10.1016/j.cell.2017.12.027

Livny, J., Yamaichi, Y., and Waldor, M. K. (2007). Distribution of centromere-like parS sites in bacteria: insights from comparative genomics. *J. Bacteriol.* 189, 8693–8703. doi: 10.1128/JB.01239-07

Löwe, J., Ellonen, A., Allen, M. D., Atkinson, C., Sherratt, D. J., and Grainge, I. (2008). Molecular Mechanism of Sequence-Directed DNA Loading and Translocation by FtsK. *Mol. Cell* 31, 498–509. doi: 10.1016/j.molcel.2008.05.027

Lu, M., Campbell, J. L., Boye, E., and Kleckner, N. (1994). SeqA: a negative modulator of replication initiation in *E. coli*. *Cell* 77, 413–426. doi: 10.1016/0092-8674(94)90156-2

Lynch, A. S., and Wang, J. C. (1995). SopB protein-mediated silencing of genes linked to the sopC locus of *Escherichia coli* F plasmid. *Proc. Natl. Acad. Sci. U.S.A.* 92, 1896–1900. doi: 10.1073/pnas.92.6.1896

Marczynski, G. T., Petit, K., and Patel, P. (2019). Crosstalk regulation between bacterial chromosome replication and chromosome partitioning. *Front. Microbiol.* 10:279. doi: 10.3389/fmicb.2019.00279

Marko, J. F., De Los Rios, P., Barducci, A., and Gruber, S. (2019). DNA-segment-capture model for loop extrusion by structural maintenance of chromosome (SMC) protein complexes. *Nucleic Acids Res.* 47, 6956–6972. doi: 10.1093/nar/gkz497

Massey, T. H., Mercogliano, C. P., Yates, J., Sherratt, D. J., and Löwe, J. (2006). Double-stranded DNA translocation: structure and mechanism of Hexameric FtsK. *Mol. Cell* 23, 457–469. doi: 10.1016/j.molcel.2006.06.019

Maurya, G. K., Kota, S., and Misra, H. S. (2019). Characterisation of ParB encoded on multipartite genome in *Deinococcus radiodurans* and their roles in radioresistance. *Microbiol. Res.* 223–225, 22–32. doi: 10.1016/j.micres.2019.03.005

McLean, T. C., Lo, R., Tschowri, N., Hoskisson, P. A., Al Bassam, M. M., Hutchings, M. I., et al. (2019). Sensing and responding to diverse extracellular signals: an updated analysis of the sensor kinases and response regulators of streptomyces species. *Microbiology* 165, 929–952. doi: 10.1099/mic.0.000817

Meng, Q., and Switzer, R. L. (2001). Regulation of transcription of the *Bacillus subtilis* pyrG gene, encoding cytidine triphosphate synthetase. *J. Bacteriol.* 183, 5513–5522. doi: 10.1128/JB.183.19.5513-5522.2001

Mercier, R., Petit, M. A., Schbath, S., Robin, S., El Karoui, M., Boccard, F., et al. (2008). The MatP/matS site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* 135, 475–485. doi: 10.1016/j.cell.2008.08.031

Mercy, C., Ducret, A., Slager, J., Lavergne, J. P., Freton, C., Nagarajan, S. N., et al. (2019). RocS drives chromosome segregation and nucleoid protection in *Streptococcus pneumoniae*. *Nat. Microbiol.* 4, 1661–1670. doi: 10.1038/s41564-019-0472-z

Mierzejewska, J., and Jagura-Burdzy, G. (2012). Prokaryotic ParA-ParB-parS system links bacterial chromosome segregation with the cell cycle. *Plasmid* 67, 1–14. doi: 10.1016/j.plasmid.2011.08.003

Minnen, A., Attaiech, L., Thon, M., Gruber, S., and Veening, J.-W. (2011). SMC is recruited to oriC by ParB and promotes chromosome segregation in *Streptococcus pneumoniae*. *Mol. Microbiol.* 81, 676–688. doi: 10.1111/j.1365-2958.2011.07722.x

Moriya, S., Tsujikawa, E., Hassan, A. K. M., Asai, K., Kodama, T., and Ogasawara, N. (1998). A *Bacillus subtilis* gene-encoding protein homologous to eukaryotic SMC motor protein is necessary for chromosome partition. *Mol. Microbiol.* 29, 179–187. doi: 10.1046/j.1365-2958.1998.00919.x

Murray, H., and Errington, J. (2008). Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell* 135, 74–84. doi: 10.1016/j.cell.2008.07.044

Murray, H., Ferreira, H., and Errington, J. (2006). The bacterial chromosome segregation protein Spo0J spreads along DNA from parS nucleation sites. *Mol. Microbiol.* 61, 1352–1361. doi: 10.1111/j.1365-2958.2006.05316.x

Nolivos, S., and Sherratt, D. (2014). The bacterial chromosome: architecture and action of bacterial SMC and SMC-like complexes. *FEMS Microbiol. Rev.* 38, 380–392. doi: 10.1111/1574-6976.12045

Nolivos, S., Upton, A. L., Badrinarayanan, A., Muller, J., Zawadzka, K., Wiktor, J., et al. (2016). MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nat. Commun.* 7:10466. doi: 10.1038/ncomms10466

Nourikyan, J., Kjos, M., Mercy, C., Cluzel, C., Morlot, C., Noirot-Gros, M. F., et al. (2015). Autophosphorylation of the bacterial tyrosine-kinase CpsD connects capsule synthesis with the cell cycle in *Streptococcus pneumoniae*. *PLoS Genet.* 11:e1005518. doi: 10.1371/journal.pgen.1005518

Oliva, M. A. (2016). Segrosome complex formation during DNA trafficking in bacterial cell division. *Front. Mol. Biosci.* 3:51. doi: 10.3389/fmolb.2016.00051

Osorio-Valeriano, M., Altegoer, F., Steinchen, W., Urban, S., Liu, Y., Bange, G., et al. (2019). ParB-type DNA segregation proteins Are CTP-dependent molecular switches. *Cell* 179, 1512–1524.e15. doi: 10.1016/j.cell.2019.11.015

Pióro, M., Małecki, T., Portas, M., Magierowska, I., Trojanowski, D., Sherratt, D., et al. (2019). Competition between DivIVA and the nucleoid for ParA binding promotes segrosome separation and modulates mycobacterial cell elongation. *Mol. Microbiol.* 111, 204–220. doi: 10.1111/mmi.14149

Ptacin, J. L., Lee, S. F., Garner, E. C., Toro, E., Eckart, M., Comolli, L. R., et al. (2010). A spindle-like apparatus guides bacterial chromosome segregation. *Nat. Cell Biol.* 12, 791–798. doi: 10.1038/ncb2083

Qin, L., Erkelens, A. M., Ben Bdira, F., and Dame, R. T. (2019). The architects of bacterial DNA bridges: a structurally and functionally conserved family of proteins. *Open Biol.* 9:190223. doi: 10.1098/rsob.190223

Quisel, J. D., Lin, D. C., and Grossman, A. D. (1999). Control of development by altered localization of a transcription factor in *B. subtilis*. *Mol. Cell* 4, 665–672. doi: 10.1016/s1097-2765(00)80377-9

Reyes-Lamothe, R., and Sherratt, D. J. (2019). The bacterial cell cycle, chromosome inheritance and cell growth. *Nat. Rev. Microbiol.* 17, 467–478. doi: 10.1038/s41579-019-0212-7

Rodionov, O., Łobocka, M., and Yarmolinsky, M. (1999). Silencing of genes flanking the P1 plasmid centromere. *Science* 283, 546–549. doi: 10.1126/science.283.5401.546

Rossmann, F. M., Rick, T., Mrusek, D., Sprankel, L., Dörrich, A. K., Leonhard, T., et al. (2019). The GGDEF domain of the phosphodiesterase PdeB in *Shewanella putrefaciens* mediates recruitment by the polar landmark protein HubP. *J. Bacteriol.* 201, 1–14. doi: 10.1128/JB.00534-18

Schofield, W. B., Lim, H. C., and Jacobs-Wagner, C. (2010). Cell cycle coordination and regulation of bacterial chromosome segregation dynamics by polarly localized proteins. *EMBO J.* 29, 3068–3081. doi: 10.1038/emboj.2010.207

Scholefield, G., Errington, J., and Murray, H. (2012). Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. *EMBO J.* 31, 1542–1555. doi: 10.1038/emboj.2012.6

Scholefield, G., Whiting, R., Errington, J., and Murray, H. (2011). Spo0J regulates the oligomeric state of Soj to trigger its switch from an activator to an inhibitor of DNA replication initiation. *Mol. Microbiol.* 79, 1089–1100. doi: 10.1111/j.1365-2958.2010.07507.x

Schumacher, M. (2007). Structural biology of plasmid segregation proteins. *Curr. Opin. Struct. Biol.* 17, 103–109. doi: 10.1016/j.sbi.2006.11.005

Schumacher, M. (2008). Structural biology of plasmid partition?: uncovering the molecular mechanisms of DNA segregation. *Biochem. J.* 412, 1–18. doi: 10.1042/BJ20080359

Schumacher, M., and Funnell, B. E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* 438, 516–519. doi: 10.1038/nature04149

Schumacher, M. A., Lee, J., Zeng, W., and Duke, N. H. (2016). Molecular insights into DNA binding and anchoring by the *Bacillus subtilis* sporulation kinetochore-like RacA protein. *Nucleic Acids Res.* 44, 5438–5449. doi: 10.1093/nar/gkw248

Schwedock, J., McCormick, J. R., Angert, E. R., Nodwell, J. R., and Losick, R. (1997). Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of *Streptomyces coelicolor*. *Mol. Microbiol.* 25, 847–858. doi: 10.1111/j.1365-2958.1997.mmi507.x

Shebelut, C. W., Guberman, J. M., van Teeffelen, S., Yakhnina, A. A., Gitai, Z., Van Teeffelen, S., et al. (2010). Caulobacter chromosome segregation is an ordered multistep process. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14194–14198. doi: 10.1073/pnas.1005274107

Soh, Y.-M., Davidson, I. F., Zamuner, S., Basquin, J., Bock, F. P., Taschner, M., et al. (2019). Self-organization of parS Centromeres by the ParB CTP Hydrolase. *Science* 366, 1129–1133. doi: 10.1126/science.aay3965

Song, D., Rodrigues, K., Graham, T. G. W., and Loparo, J. J. (2017). A network of cis and trans interactions is required for ParB spreading. *Nucleic Acids Res.* 45, 7106–7117. doi: 10.1093/nar/gkx271

Stock, J. B., Ninfa, A. J., and Stock, A. M. (1989). Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* 53, 450–490. doi: 10.1128/mmbr.53.4.450-490.1989

Stouf, M., Meile, J. C., and Cornet, F. (2013). FtsK actively segregates sister chromosomes in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11157–11162. doi: 10.1073/pnas.1304080110

Sullivan, N. L., Marquis, K. A., and Rudner, D. Z. (2009). Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 137, 697–707. doi: 10.1016/j.cell.2009.04.044

Szafran, M., Skut, P., Ditkowski, B., Ginda, K., Chandra, G., Zakrzewska-Czerwińska, J., et al. (2013). Topoisomerase I (TopA) is recruited to ParB complexes and is required for proper chromosome organization during *Streptomyces coelicolor* sporulation. *J. Bacteriol.* 195, 4445–4455. doi: 10.1128/JB.00798-13

Taylor, J. A., Pastrana, C. L., Butterer, A., Pernstich, C., Gwynn, E. J., Sobott, F., et al. (2015). Specific and non-specific interactions of ParB with DNA: implications for chromosome segregation. *Nucleic Acid Res.* 43, 719–731. doi: 10.1093/nar/gku1295

Thanbichler, M., and Shapiro, L. (2006). MipZ, a spatial regulator coordinating chromosome segregation with cell division in Caulobacter. *Cell* 126, 147–162. doi: 10.1016/j.cell.2006.05.038

Thiel, A., Valens, M., Vallet-Gely, I., Espéli, O., and Boccard, F. (2012). Long-range chromosome organization in *E. coli*: a site-specific system isolates the ter macrodomain. *PLoS Genet.* 8:e1002672. doi: 10.1371/journal.pgen.1002672

Thoma, L., and Muth, G. (2015). The conjugative DNA-transfer apparatus of Streptomyces. *Int. J. Med. Microbiol.* 305, 224–229. doi: 10.1016/j.ijmm.2014.12.020

Tran, N. T., Laub, M. T., and Le, T. B. K. (2017). SMC progressively aligns chromosomal arms in *Caulobacter crescentus* but is antagonized by convergent transcription. *Cell Rep.* 20, 2057–2071. doi: 10.1016/j.celrep.2017.08.026

Tran, N. T., Stevenson, C. E., Som, N. F., Thanapipatsiri, A., Jalal, A. S. B., and Le, T. B. K. (2018). Permissive zones for the centromere-binding protein ParB on the *Caulobacter crescentus* chromosome. *Nucleic Acids Res.* 46, 1196–1209. doi: 10.1093/nar/gkx1192

Trojanowski, D., Ginda, K., Pióro, M., Hołówka, J., Skut, P., Jakimowicz, D., et al. (2015). Choreography of the mycobacterium replication machinery during the cell cycle. *mBio* 6:e02125-14. doi: 10.1128/mBio.02125-14

Tsokos, C. G., and Laub, M. T. (2012). Polarity and cell fate asymmetry in *Caulobacter crescentus*. *Curr. Opin. Microbiol.* 15, 744–750. doi: 10.1016/j.mib.2012.10.011

Valens, M., Thiel, A., and Boccard, F. (2016). The MaoP/maoS Site-Specific System Organizes the Ori Region of the *E. coli* Chromosome into a Macrodomain. *PLoS Genet.* 12:e1006309. doi: 10.1371/journal.pgen.1006309

Vecchiarelli, A. G., Neuman, K. C., and Mizuuchi, K. (2014). A propagating ATPase gradient drives transport of surface-confined cellular cargo. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4880–4885. doi: 10.1073/pnas.1401025111

Vickridge, E., Planchenault, C., Cockram, C., Junceda, I. G., and Espéli, O. (2017). Management of *E. coli* sister chromatid cohesion in response to genotoxic stress. *Nat. Commun.* 8:14618. doi: 10.1038/ncomms14618

Vogelmann, J., Ammelburg, M., Finger, C., Guezguez, J., Linke, D., Flötenmeyer, M., et al. (2011). Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. *EMBO J.* 30, 2246–2254. doi: 10.1038/emboj.2011.121

Walker, K. A., Mallik, P., Pratt, T. S., and Osuna, R. (2004). The *Escherichia coli* fis promoter is regulated by changes in the levels of its transcription initiation nucleotide CTP. *J. Biol. Chem.* 279, 50818–50828. doi: 10.1074/jbc.M406285200

Wang, X., Le, T. B. K., Lajoie, B. R., Dekker, J., Laub, M. T., and Rudner, D. Z. (2015). Condensin promotes the juxtaposition of DNA flanking its loading site in *Bacillus subtilis*. *Genes Dev.* 29, 1661–1675. doi: 10.1101/gad.265876.115

Wang, X., Montero Llopis, P., and Rudner, D. Z. (2013). Organization and segregation of bacterial chromosomes. *Nat. Rev. Genet.* 14, 191–203. doi: 10.1038/nrg3375

Wang, X., Montero Llopis, P., and Rudner, D. Z. (2014). *Bacillus subtilis* chromosome organization oscillates between two distinct patterns. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12877–12882. doi: 10.1073/pnas.1407461111

Wilhelm, L., Bürmann, F., Minnen, A., Shin, H. C., Toseland, C. P., Oh, B. H., et al. (2015). SMC condensin entraps chromosomal DNA by an ATP hydrolysis dependent loading mechanism in *Bacillus subtilis*. *eLife* 4:e06659. doi: 10.7554/eLife.06659

Yamaichi, Y., Bruckner, R., Ringgaard, S., Möll, A., Ewen Cameron, D., Briegel, A., et al. (2012). A multidomain hub anchors the chromosome segregation and chemotactic machinery to the bacterial pole. *Genes Dev.* 26, 2348–2360. doi: 10.1101/gad.199869.112

Yamaichi, Y., and Niki, H. (2004). migS, a cis-acting site that affects bipolar positioning of oriC on the *Escherichia coli* chromosome. *EMBO J.* 23, 221–233. doi: 10.1038/sj.emboj.7600028

Yamazoe, M., Onogi, T., Sunako, Y., Niki, H., Yamanaka, K., Ichimura, T., et al. (1999). Complex formation of MukB, MukE and MukF proteins involved in chromosome partitioning in *Escherichia coli*. *EMBO J.* 18, 5873–5884. doi: 10.1093/emboj/18.21.5873

Yu, X. C., Weihe, E. K., and Margolin, W. (1998). Role of the C terminus of FtsK in *Escherichia coli* chromosome segregation. *J. Bacteriol.* 180, 6424–6428. doi: 10.1128/.180.23.6424-6428.1998

Zhou, Q., Zhou, Y. N., Jin, D. J., and Tse-Dinh, Y. C. (2017). Deacetylation of topoisomerase i is an important physiological function of *E. coli* CobB. *Nucleic Acids Res.* 45, 5349–5358. doi: 10.1093/nar/gkx250

Zhou, Y., Chen, T., Zhou, L., Fleming, J., Deng, J., Wang, X., et al. (2015). Discovery and characterization of Ku acetylation in *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* 362:fnu051. doi: 10.1093/femsle/fnu051

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 15 April 2020
doi: 10.3389/fmicb.2020.00534

Too Much of a Good Thing: How Ectopic DNA Replication Affects Bacterial Replication Dynamics

*Aisha H. Syeda[1], Juachi U. Dimude[2], Ole Skovgaard[3] and Christian J. Rudolph[2]**

[1] Department of Biology, University of York, York, United Kingdom, [2] Division of Biosciences, College of Health and Life Sciences, Brunel University London, Uxbridge, United Kingdom, [3] Department of Science and Environment, Roskilde University, Roskilde, Denmark

Each cell division requires the complete and accurate duplication of the entire genome. In bacteria, the duplication process of the often-circular chromosomes is initiated at a single origin per chromosome, resulting in two replication forks that traverse the chromosome in opposite directions. DNA synthesis is completed once the two forks fuse in a region diametrically opposite the origin. In some bacteria, such as *Escherichia coli*, the region where forks fuse forms a specialized termination area. Polar replication fork pause sites flanking this area can pause the progression of replication forks, thereby allowing forks to enter but not to leave. Transcription of all required genes has to take place simultaneously with genome duplication. As both of these genome trafficking processes share the same template, conflicts are unavoidable. In this review, we focus on recent attempts to add additional origins into various ectopic chromosomal locations of the *E. coli* chromosome. As ectopic origins disturb the native replichore arrangements, the problems resulting from such perturbations can give important insights into how genome trafficking processes are coordinated and the problems that arise if this coordination is disturbed. The data from these studies highlight that head-on replication–transcription conflicts are indeed highly problematic and multiple repair pathways are required to restart replication forks arrested at obstacles. In addition, the existing data also demonstrate that the replication fork trap in *E. coli* imposes significant constraints to genome duplication if ectopic origins are active. We describe the current models of how replication fork fusion events can cause serious problems for genome duplication, as well as models of how such problems might be alleviated both by a number of repair pathways as well as the replication fork trap system. Considering the problems associated both with head-on replication-transcription conflicts as well as head-on replication fork fusion events might provide clues of how these genome trafficking issues have contributed to shape the distinct architecture of bacterial chromosomes.

Keywords: replication, transcription, recG gene, termination of DNA replication, ectopic replication origins, bacterial replication dynamics, 3′ exonuclease

OPEN ACCESS

Edited by:
Torsten Waldminghaus, University of Marburg, Germany

Reviewed by:
Olivier Espeli, Centre National de la Recherche Scientifique (CNRS), France
Francois Cornet, Centre National de la Recherche Scientifique (CNRS), France

****Correspondence:***
Christian J. Rudolph
christian.rudolph@brunel.ac.uk

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *19 January 2020*
Accepted: *12 March 2020*
Published: *15 April 2020*

Citation:
Syeda AH, Dimude JU, Skovgaard O and Rudolph CJ (2020) Too Much of a Good Thing: How Ectopic DNA Replication Affects Bacterial Replication Dynamics. Front. Microbiol. 11:534. doi: 10.3389/fmicb.2020.00534

INTRODUCTION

While eukaryotic cells typically contain multiple linear chromosomes, the bacterial models studied in most detail early on, such as *Escherichia coli* and *Bacillus subtilis*, have a single chromosome with a size of roughly 5 Mbp that forms a covalently closed circle. The improved ability to sequence whole genomes has revealed considerable variations. For example, *Mycoplasma genitalium*, a sexually transmitted pathogen that can cause non-gonococcal urethritis, is one of the smallest prokaryotes capable of independent replication with a genome size of 0.58 Mbp and less than 500 genes (Taylor-Robinson and Jensen, 2011; Gnanadurai and Fifer, 2020). In strictly opportunistic or symbiotic bacteria, genomes can be even smaller: the symbiotic bacterium *Carsonella ruddii* carries a single circular chromosome containing 0.159 Mbp and is predicted to encode 182 genes (Nakabachi et al., 2006). The genome of the myxobacterium *Sorangium cellulosum*, on the other hand, contains just over 13 Mbp and is predicted to encode 9,367 coding sequences (Schneiker et al., 2007). Overall, protein-coding density of bacterial genomes is with 85–90% high (McCutcheon and Moran, 2011) and the correlation between genome size and the number of genes is surprisingly constant (Touchon and Rocha, 2016).

Many of the extensively studied bacterial models are haploid. In *E. coli*, overlapping cell cycles in fast growing cells allow an increase in genome equivalents and stationary cells contain only a single copy of the chromosome. In contrast, many other bacterial species carry multiply copies of the chromosome. *Deinococcus radiodurans* carries between four and 10 genome equivalents (Hansen, 1978), and the presence of multiple copies is thought to be one contributor to its extreme radiation resistance (Minton and Daly, 1995; Timmins and Moe, 2016). Bacteria such as *Azotobacter vinelandii* can carry up to 80 chromosome copies per cell under fast growth conditions (Nagpal et al., 1989), and tens of thousands of copies were reported for the large bacterium *Epulopiscium* (Mendell et al., 2008).

While the presence of multiple chromosome equivalents is relatively common, the presence of more than one type of chromosome is less frequent, found in about 5% of bacterial species investigated so far (Touchon and Rocha, 2016). Examples are *Vibrio cholerae* and close relatives of *Vibrio*, which usually carry two circular chromosomes (Touchon and Rocha, 2016), while *Paracoccus denitrificans*, a gram-negative soil bacterium, carries three different circular chromosomes (Winterstein and Ludwig, 1998).

While the majority of bacterial chromosomes form covalently closed circles, some bacterial species carry linear chromosomes or even a mix of circular and linear chromosomes. For example, *Agrobacterium tumefaciens* carries one circular and one linear chromosome, as well as two very large plasmids (Nester, 2015). Linear chromosomes are frequently found within the *Actinomycetales*, which includes the genus *Streptomyces* (Kirby, 2011). The normally circular *E. coli* chromosome can also be artificially linearized using the telomere system of bacteriophage N15, and the resulting cells grow stably without any observed ill effect (Cui et al., 2007; Rudolph et al., 2013).

THE BACTERIAL REPLICHORE ARRANGEMENT

Despite these considerable variations, the replichore arrangements of most bacterial genomes are straightforward. While replication of the multiple linear chromosomes in eukaryotic cells is initiated at hundreds or even thousands of origins (Leonard and Méchali, 2013), initiation sites in bacteria are restricted to a single origin per chromosome (*oriC*) (Gao and Zhang, 2008; Gao, 2015). For a bacterium such as *E. coli*, this means that the number of replisomes is restricted to two, which are recruited at the origin and proceed in opposite directions until they eventually fuse opposite the *oriC* (Masters and Broda, 1971; Prescott and Kuempel, 1972; Dimude et al., 2016). Thus, each chromosomal half or replichore (Blattner et al., 1997) is replicated by one fork in a defined directionality (**Figure 1A**). DNA replication is successfully completed once every single base pair of the chromosome is duplicated with high accuracy. However, daughter chromosomes will remain interlinked until they are resolved through post-replicative processing (Lesterlin et al., 2004; Reyes-Lamothe et al., 2012), a process that is coordinated both temporally and spatially with septum formation at mid-cell (Reyes-Lamothe et al., 2012; Zaritsky and Woldringh, 2015).

In *E. coli*, the replichore arrangement results in certain asymmetric features of the chromosomal halves. For example, the leading and lagging strands show a nucleotide composition bias, with G being overrepresented in the leading strand (Wu and Maeda, 1987; Lobry, 1996; Blattner et al., 1997). The contributions from transcription and replication toward this bias is still under debate (Francino et al., 1996; Rocha et al., 2006; Chen et al., 2016), but replication and replication-linked processes, such as cytosine deaminations, which were shown to occur preferentially in the lagging strand (Bhagwat et al., 2016), clearly contribute. The compositional bias results in a sharp transition both at the origin and the terminus near the dif dimer resolution site (Wu and Maeda, 1987; Lobry, 1996; Blattner et al., 1997; Lobry and Louarn, 2003). In addition, the KOPS 8-mer (FtsK Orienting Polar Sequences) is asymmetric, with a preference of pointing toward the dif chromosome dimer resolution site. This allows not only binding, but also the directional movement of Ftsk, which is essential for the unlinking of chromosome dimers that can arise as a result of an odd number or recombination events (Bigot, 2005; Levy et al., 2005; Barre, 2007; Sherratt et al., 2010).

Higher order genome organization appears to correlate to some extent with the replichore arrangement. In initial experiments it was observed that relatively large regions of the chromosome colocalize *in vivo*, leading to the suggestion of the existence of one macro domain that contains the origin area and a second macrodomain that contains the terminus area of the chromosome (Valens et al., 2004; Verma et al., 2019). The macrodomain structure of the chromosome was further investigated with fluorescence-microscopy and recombination-based approaches as well as, most recently, with chromosome conformation capture methods (3C), leading to the idea that the *E. coli* chromosome is divided into four macrodomains

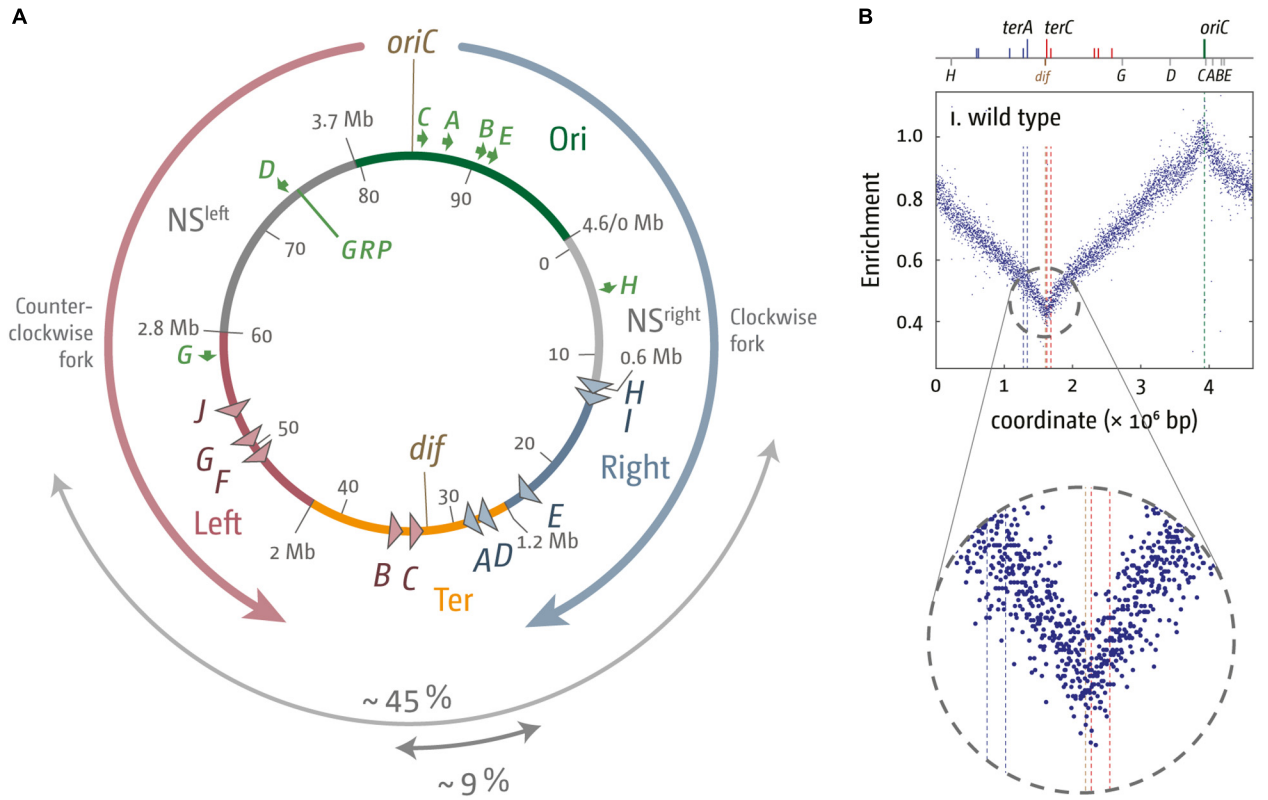


FIGURE 1 | Chromosome structure and replication dynamics in *Escherichia coli*. **(A)** Schematic representation of the *E. coli* chromosome. Two replication forks are initiated at the origin (*oriC*) move in opposite directions along the DNA and eventually approach one other and fuse within the terminus region diametrically opposed to *oriC*. A replication fork trap is formed in the terminus region via terminator sequences (*terA–J*) which are arranged as two opposed groups, with the red terminators oriented to block movement of the clockwise replication fork and the blue terminators oriented to block the anticlockwise fork. The large gray arrow highlights the total spanned area covered by ter sites, while the core termination area, defined by the four innermost ter sites, is marked by a small gray arrow. The chromosomal locations for *oriC* and the dif chromosome dimer resolution site are marked. The location of rrn operons, which are highly transcribed particularly under fast growth conditions, are shown by green arrows, with the arrow pointing in the direction in which transcribing RNA polymerase molecules travel. "GRP" indicates the location of a cluster of genes encoding ribosomal proteins, almost all of which are transcribed co-directionally with replication. Chromosomal macrodomains Ori, NS^{right}, NS^{left}, Right, Left, and Ter are shown as described in Duigou and Boccard (2017) and domain boundaries given in Mbp. Numbers on the inside are the minutes of the standard genetic map (0–100 min). **(B)** Marker frequency analysis of wild type *E. coli* cells. The number of reads (normalized against reads for a stationary phase wild type control) is plotted against the chromosomal location. A schematic representation of the *E. coli* chromosome showing positions of *oriC* (green line) and ter sites (above) as well as dif and rrn operons A–E, G, and H (below) is shown above the plotted data. The MFA raw data were taken from Rudolph et al. (2013) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome. A magnified view of the replication profile in the termination area is shown in the enlarged circle.

(Ori, Ter, Left, and Right) as well as two more flexible and non-structured regions, NS-L and NS-R, that flank the Ori macrodomain (Liu et al., 2010; Duigou and Boccard, 2017; Verma et al., 2019; **Figure 1A**).

THE TERMINATION AREA IN *ESCHERICHIA COLI*

One peculiarity of the termination area both in *E. coli* and *B. subtilis* is the ability to restrict fork movement via a "replication fork trap," a series of protein-DNA complexes that are asymmetric. An approaching fork coming from one direction can displace the bound protein and continue to traverse the chromosome, while a fork coming from the other direction will be paused and unable to proceed past the block for some time. The short DNA sequences involved are called terminator, or ter, sequences. In *E. coli* each ter sequence can be bound by a single Tus protein (terminus utilization substance), while in *B. subtilis* ter sequences are bound by an Rtp (replication termination protein) dimer. Both *E. coli* and *B. subtilis* are similar in that the ter sequences are positioned to form two opposed groups that allow replication fork complexes to enter but not exit the termination region. However, the overall size of the termination area differs significantly: while in *E. coli* ter sequences are distributed over >40% of the chromosome (**Figure 1A**), the spread is much narrower in *B. subtilis* (<10%). However, in normally growing *E. coli* cells, only the four inner-most ter sites, *terC* and *terB* on one side and *terA* and *terD* on the other, are substantially involved in the arrest of DNA replication (de Massy et al., 1987; Hill et al., 1987; Duggin and Bell, 2009). Thus, these four sites are considered to be the primary fork trap, and with

about 9% their spread is similar to the spread of ter sites in the *B. subtilis* chromosome (**Figure 1A**).

In *E. coli* MG1655, *terC* is generally the first ter/Tus complex to be encountered by a replisome (Duggin and Bell, 2009). *terC* is located almost directly opposite the origin and will arrest the replisome traversing the chromosome in clockwise orientation (**Figure 1A**). The second innermost ter site is *terA*, which is located in a slightly more asymmetric position (**Figure 1A**). The outer terminators are probably used only rarely (Griffiths and Wake, 2000; Duggin and Bell, 2009). However, it is important to note that ter/Tus complexes are not systematically involved in replication termination. This was already shown by early labeling experiments (Bouché et al., 1982) and supported more recently by high-resolution replication profiles established via deep sequencing.

High-resolution replication profiles can be generated from marker frequency analyses (MFA) by deep sequencing (Skovgaard et al., 2011; Rudolph et al., 2013; Müller et al., 2014). MFA is generated by plotting the ratio of uniquely mapped sequence reads per 1 kb window in a replicating sample relative to a non-replicating control (stationary phase wild type cells). The replication profile for rapid growing wild type cells shows the location of *oriC* as a clear maximum, while a minimum in the termination area shows the most common fork fusion point (Skovgaard et al., 2011; Rudolph et al., 2013; Ivanova et al., 2015).

The fact that replication profiles show a distinct V-shaped low point (**Figure 1B**) suggests that the majority of fork fusions in *E. coli* take place near the arithmetic mid-point. Indeed, we observed that the low point of the replication profiles in the presence and absence of a functional fork trap was in the same location (Rudolph et al., 2013; Ivanova et al., 2015; Dimude et al., 2016), suggesting that both replisomes traverse their replichores with similar speeds and fuse freely within the innermost ter sites. It appears that the fork trap is only involved in termination if one replisome is delayed at an obstacle on its way through the replichore (Duggin et al., 2008; Duggin and Bell, 2009).

A recent analysis from Galli et al. (2019) has shown that Tus-related sequences are found in most *Enterobacteriales*, in the *Pseudoalteromonas*, and in most *Aeromonadales*. In contrast, RTP-related sequences are restricted to a subgroup of the *Bacillales* (Galli et al., 2019). Indeed, sequence analysis suggests that a replication fork trap is absent in many bacterial species. This was experimentally demonstrated for the two circular *Vibrio cholerae* chromosomes (Galli et al., 2019). Similarly, no specific termination-related pause sites have been identified in eukaryotes and archaea, even though multiple replication origins per chromosome result in a much higher number of fork fusions. It appears that replication effectively terminates at random locations between origins (Duggin et al., 2011; Hawkins et al., 2013; Samson et al., 2013; Gambus, 2017).

The absence of any significant sequence or structural similarity of the components of the fork trap in *E. coli* and *B. subtilis* indicates that fork trap systems have evolved via convergent evolution (Neylon et al., 2005). If this is the case, then the system would be expected to have an important physiological function. However, early studies suggested that the inactivation of the fork trap both in *B. subtilis* and *E. coli* has very little effect on growth rate and cell morphology (Iismaa and Wake, 1987; Roecklein et al., 1991), suggesting that our understanding of the physiological role of the termination area is incomplete. We will explore possible roles of the replication fork trap later in this review.

COORDINATING REPLICATION AND TRANSCRIPTION

The combination of a single point of replication initiation with a fork trap mechanism enforces a strong directionality of replication in wild type cells, as each replichore is replicated in a defined orientation under normal conditions. It was suggested that this directionality might be advantageous (Brewer, 1988; French, 1992; Dimude et al., 2016). Replication and transcription move with very different speeds, as transcription is significantly slower than DNA replication (Vogel and Jensen, 1994; Dennis et al., 2009), and, given that both processes utilize the same template, conflicts are unavoidable. Indeed, highly transcribed genes were found to be preferentially located on the template for the leading strand in a number of bacterial species, resulting in the co-directional movement of replisomes and transcribing RNA polymerase complexes (Brewer, 1988; McLean et al., 1998; Rocha and Danchin, 2003; Evertts and Coller, 2012). In *E. coli*, global co-orientation is only just under 55%, but over 90% of genes encoding ribosomal proteins, which are particularly highly transcribed, show co-directionality of replication and transcription (Brewer, 1988; McLean et al., 1998; **Figure 1A**). A higher general co-orientation was observed in other bacteria, with more than 70% of genes being transcribed co-directionally with replication in *B. subtilis* and *Mycoplasma pneumonia*, with virtually all genes that code for ribosomal proteins being transcribed co-directionally with replication (McLean et al., 1998).

The co-directionality of highly transcribed genes and DNA replication indicates head-on encounters of replisomes with transcribing RNA polymerase complexes are particularly problematic (French, 1992; Rudolph et al., 2007a; Kim and Jinks-Robertson, 2012; McGlynn et al., 2012; Merrikh et al., 2012), even though any encounter can interfere with ongoing DNA replication (Merrikh et al., 2011; Lang and Merrikh, 2018). Indeed, it was shown in both *E. coli* and *B. subtilis* cells that replication of a highly transcribed rrn operon in an orientation opposite to normal caused significant problems (Wang et al., 2007; Boubakri et al., 2010; Srivatsan et al., 2010; De Septenville et al., 2012; Million-Weaver et al., 2015).

In eukaryotic cells, replication–transcription encounters are expected to cause similar problems. However, initially the analysis of replication and transcription directionality in human cells has revealed little overall bias, suggesting that the orientation of open reading frames might be effectively random (Necsulea et al., 2009; Hyrien, 2015), perhaps with the exception of yeast, in which a replication barrier prevents forks from entering highly transcribed ribosomal DNA repeats in a head-on orientation (Hyrien, 2000; Mirkin and Mirkin, 2007; Evertts and Coller, 2012). This view has recently changed. A recent

study showed a preference for replication initiation sites in human cells to occur in the immediate vicinity of transcription start sites, while termination of synthesis occurs at the 3′ end of genes, highlighting that the same fundamental principle of co-directionality applies in human cells (Chen et al., 2019).

CONSTRAINTS OF THE BACTERIAL REPLICHORE ARRANGEMENT

While there is a certain esthetic beauty to the straightforward bacterial replichore arrangement, this system also imposes significant constraints. If replication is initiated exclusively at a single origin, then the ability of fast growth is directly linked to the speed of chromosome duplication. Indeed, the speed of replication in *E. coli* is 650–1000 nt × s^{-1} (Pham et al., 2013), which is about 20 × faster than DNA replication in human cells (Méchali, 2010). The use of 30,000–50,000 origins in human cells can compensate for slow speed and the longer duplication time of the larger genome, and indeed, in *Xenopus laevis* and *Drosophila melanogaster*, origins are activated at very short intervals during early embryonic development (Méchali, 2010). Bacteria such as *E. coli* have to utilize overlapping rounds of DNA synthesis in order to achieve a cell duplication period that is shorter than the time required to duplicate the entire chromosome (Dewachter et al., 2018). Chromosome duplication is completed in approximately 40 min, but cells can divide every 20 min in rich medium that allows overlapping rounds of DNA replication. Indeed, under conditions where progression of ongoing DNA synthesis is blocked by DNA lesions while initiation at *oriC* can still take place, a temporary cell division period of <15 min is observed once the lesions have been eliminated, allowing all initiated forks rapidly to generate complete chromosomal copies (Rudolph et al., 2007b, 2010b).

The presence of a replication fork trap as part of the chromosome architecture in bacteria would appear to be particularly problematic in the face of obstacles to DNA replication. While replication in *E. coli* is very fast and accurate, progression of synthesis will always encounter obstacles, including stable protein-DNA complexes, secondary structures, a variety of DNA lesions and other problems (Cox, 2001; McGlynn et al., 2012; Merrikh et al., 2012). If duplication of a chromosome is restricted to two replication forks and a replication fork trap is present, such obstacles can have potentially disastrous consequences. If one fork is permanently blocked, a replication fork trap will prevent it being rescued by the second fork, as this fork will also be blocked (Dimude et al., 2016). We believe that this particular problem explains in part why replication restart proteins such as PriA are so prominent in bacteria (Dimude et al., 2016), as these proteins are essential for the re-recruitment of functional replisomes following the removal of obstacles of damage (Gabbai and Marians, 2010; Windgassen et al., 2018).

Why are bacterial chromosomes exclusively replicated using a single origin if this scenario can be problematic? In eukaryotic cells, under-replicated stretches of DNA can trigger activation of "dormant" origins that aid the completion of DNA synthesis if progression of early forks is delayed by obstacles or damage (Blow et al., 2011; Courtot et al., 2018). Whilst archaea predominantly carry circular chromosomes, at least some species utilize multiple origins to replicate their genomes (Lundgren et al., 2004; Wu et al., 2014). Thus, the consistent use of a single origin in bacteria may seem surprising, especially as gross chromosomal rearrangements can occur relatively frequently (Umenhoffer et al., 2017).

INTRODUCING A SECOND ORIGIN INTO THE *E. COLI* CHROMOSOME

Given the multiple origins per chromosome in archaea and eukaryotes, researchers have asked whether multiple origins can be utilized in bacterial chromosomes. The Sherratt lab was able to integrate a 5 kb *oriC* fragment near the lac operon into the *E. coli* chromosome, roughly in the middle of the right-hand replichore (Wang et al., 2011). To distinguish this origin from the native *oriC*, this origin was termed oriZ even though the sequence is identical to *oriC* (**Figure 2A**). Cells carrying both *oriC* and oriZ, which we will refer to as $oriC^+$ $oriZ^+$ cells, were reported to have doubling times similar to wild-type cells, and fluorescence microscopy confirmed that both origins are active and fire simultaneously (Wang et al., 2011).

In line with the fluorescence microscopy data (Wang et al., 2011) the replication profile of $oriC^+$ $oriZ^+$ cells shows a second maximum at the location of oriZ and an additional and ectopic local minimum between *oriC* and oriZ (**Figure 2B**), indicative of a second fork fusion point (Rudolph et al., 2013; Ivanova et al., 2015). However, the primary minimum of the replication profile shows a distinct step in between *terA* and *terB/C* in $oriC^+$ $oriZ^+$ cells, rather than a V-shape. As oriZ is roughly in the middle of the right-hand replichore, forks initiated at oriZ and traversing toward the termination area only have to duplicate 1/4 of the chromosome before they reach the fork trap area, while the fork initiated at *oriC* and proceeding counterclockwise has to replicate the entire replichore. Thus, within a randomly growing population there will be significantly more cells in which forks coming from oriZ will get trapped at *terC* and subsequent ter sites until the second fork reaches this area, resulting in the defined "step" between *terA* and *terC* (**Figure 2B**).

In the initial analysis, the doubling time of $oriC^+$ $oriZ^+$ and wild type cells was found to be similar (Wang et al., 2011). However, when we measured the doubling times for MG1655 and $oriC^+$ $oriZ^+$ constructs in direct comparison, we found that $oriC^+$ $oriZ^+$ cells grew slightly slower in two independent studies (∼21 min in comparison to ∼20 min in wild type cells) (Ivanova et al., 2015; Dimude et al., 2018b).

We also integrated the same 5 kb *oriC* fragment roughly into the middle of the left-hand replichore, which resulted in the generation of $oriC^+$ $oriX^+$ cells (**Figure 2A**; Dimude et al., 2018b). The replication profile of these cells proved very similar to the profile observed in $oriC^+$ $oriZ^+$ cells (**Figure 2B**). MFA analysis confirmed that oriX was active and suggests that both *oriC* and oriX fire simultaneously in the majority of cells, a result confirmed via fluorescence microscopy (Dimude et al., 2018b). Replication profiles of $oriC^+$ $oriX^+$ cells showed the

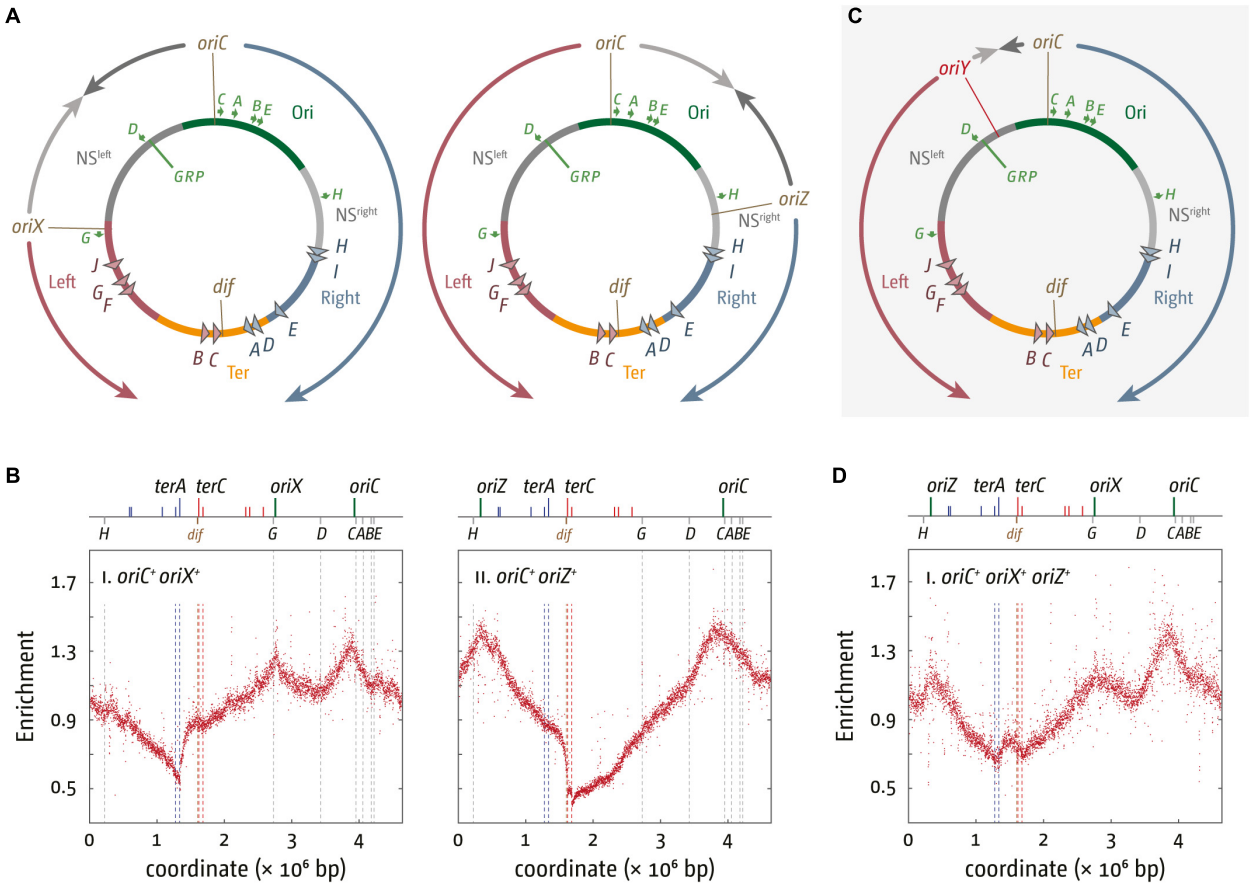


FIGURE 2 | Chromosome structure and replication dynamics in *E. coli* cells with additional ectopic replication origins. **(A)** Integration sites of 5 kb *oriC* fragments into pheA upstream of the rrnG operon, termed oriX, and near the lacZYA operon, termed oriZ (Wang et al., 2011; Ivanova et al., 2015; Dimude et al., 2018b). All genetic and structural elements shown are as described in **Figure 1**. **(B)** Marker frequency analysis of *E. coli oriC*$^{+}$ oriX^{+} and *oriC*$^{+}$ oriZ^{+} cells. The number of reads (normalized against reads for a stationary phase wild type control) is plotted against the chromosomal location. A schematic representation of the *E. coli* chromosome showing positions of *oriC*, oriX and oriZ (green lines) and ter sites (above) as well as dif and rrn operons A–E, G, and H (below) is shown above the plotted data. The MFA raw data were taken from Dimude et al. (2018b) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome. **(C)** Integration site of a 5 kb *oriC* fragment, termed oriY, into malT, upstream of the rrnD operon. See text for details. **(D)** Marker frequency analysis in *E. coli oriC*$^{+}$ oriX^{+} oriZ^{+} cells. The number of reads (normalized against reads for a stationary phase wild type control) is plotted against the chromosomal location. A schematic representation of the *E. coli* chromosome showing positions of *oriC*, oriX, and oriZ (green lines) and ter sites (all above) as well as dif and rrn operons A–E, G, and H (all below) is shown above the plotted data. The MFA raw data were taken from Dimude et al. (2018b) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome.

same general features as the profiles from *oriC*$^{+}$ oriZ^{+} cells, including a step in the termination area which is located at *terA* (Dimude et al., 2018b).

As already observed for *oriC*$^{+}$ oriZ^{+} cells, we again found the doubling times for *oriC*$^{+}$ oriX^{+} constructs to be slightly longer (∼22 min vs. ∼19.5 min for wild type cells) (Dimude et al., 2018b), providing additional confirmation that the introduction of an additional ectopic origin interferes with genome duplication and/or segregation.

While the integration of a second ectopic origin proved relatively unproblematic in both replichores aside from the mild growth defect, other attempts were less successful. The integration of a plasmid-derived origin that could be induced with IPTG at a location ∼450 kb away from *oriC* was successful, but if this origin was active, it repressed activity of *oriC* (Kouzminova and Kuzminov, 2008). In another study, integration of a shorter *oriC* fragment in two chromosomal locations, one roughly equivalent to the oriZ position while the second was closer to the termination area (1.6 Mbp), did not result in any detectable initiation at the ectopic origins (Milbredt et al., 2016). The authors suggested that origin activity might be influenced by the presence of flanking genes (Milbredt et al., 2016), which would explain why the longer 5 kb *oriC* region stretch developed in the Sherratt lab (Wang et al., 2011) proved active. However, our attempts to integrate the same 5 kb *oriC* fragment into the *malT* gene at 76.5 min, approximately 1/4 into the left-hand replichore, to generate *oriC*$^{+}$ oriY^{+} cells (**Figure 2C**), proved unsuccessful. We had little difficulty getting chromosomal integrations displaying the correct antibiotic resistance. However, the oriY was not active

and PCR analysis of two independent oriY constructs showed that the *oriC* core elements were either truncated or completely absent (Dimude et al., 2018b). The difference of the truncations observed suggests that they are spontaneous mutations, arising perhaps because of a toxicity caused by an active origin being integrated in this precise location. Given that the integration of the antibiotic resistance marker occurred without any problem, it appears that the integration of an ectopic sequence in this location is unproblematic.

INTRODUCING THREE ORIGINS INTO THE *E. COLI* CHROMOSOME

We went on to generate an *oriC*$^+$ oriX$^+$ oriZ$^+$ strain with three origins, which proved unproblematic. However, the replication profile of this construct revealed a surprising detail: the peak heights of ectopic origins oriZ and oriX were reduced in comparison to the peak height of the native *oriC* (**Figure 2D**; Dimude et al., 2018b). This indicates that both ectopic origins, oriX and oriZ, are used less frequently than the native *oriC*, a result that contrasts with both double-origin constructs where the peak heights of the native *oriC* and the ectopic origin were very similar (Dimude et al., 2018b). Replication profiles are population-based, and for this reason allow little insight into origin usage in single cells. To directly visualize active replisomes in oriX$^+$ *oriC*$^+$ oriZ$^+$ cells we used YPet-DnaN, a fluorescently tagged version of the β sliding clamp. Previously we observed that the signal in double-origin cells produced defined foci, as described before (Wang et al., 2011). In contrast, foci in triple-origin cells were much less defined. The analysis of foci, which are not only in close proximity but, in addition, not particularly well defined, proved rather difficult. However, we observed some cells with three separate foci, indicating that all three origins are active in these cells. However, the replication profiles clearly show a reduced activity of the ectopic origins in comparison to *oriC*, as the peak height of both ectopic origins is lower than the peak height of *oriC*. The difference in peak heights suggests that in some cells only two origins are active, but as the *oriC* peak is the highest it indicates that in these cells one of the two active origins is always the native *oriC*, whereas the ectopic origin is either oriX or oriZ.

We also observed that cultures of triple-origin cells showed an increase of cells with no foci. This could be due to a frequent failure of ongoing replication. Alternatively, it could highlight a failure to initiate replication. For example, a threshold concentration of the DnaA initiator protein is required for successful initiation (Boye et al., 2000). An increase in the number of origins will lead to an increase in the number of DnaA binding sites, which will cause a drop in the concentration of free DnaA. In a fraction of cells this drop might result in none of the origins being activated, as observed. No such effect was observed in any of the double-origin constructs (Wang et al., 2011; Ivanova et al., 2015; Dimude et al., 2018b), indicating that levels of free DnaA must be high enough to allow simultaneous initiation if two origins in the vast majority of cells. Thus, we currently do not know the precise molecular effects that cause formation of cells with no foci.

The fact that *oriC* activity is highest in triple-origin cells (**Figure 2D**) demonstrates that the capacity for *oriC* being active is highest in its native location, highlighting the importance of genome organization in the vicinity of *oriC*, and the importance of the location of *oriC* itself. We are only just beginning to appreciate the complexity of the three-dimensional structure of the nucleoid in bacterial cells. Indeed, changes of the *oriC* position were shown to alter the position of the Right and Left chromosomal macrodomains, highlighting that the position of *oriC* has a significant effect on chromosome organization (Duigou and Boccard, 2017). In addition, global gene order is surprisingly conserved between closely related prokaryotic species (Tamames, 2001). This order will get disrupted if additional origins are introduced into the chromosome, and we are only now starting to appreciate the effects this might have. Finally, the toxicity caused by oriY integration supports the idea that either the precise location of an active origin or the relative position of two active origins to each other can have strong effects (Dimude et al., 2018b), as observed (Kouzminova and Kuzminov, 2008).

DNA REPLICATION IN CELLS WITHOUT ACTIVE REPLICATION ORIGINS

The initiation of DNA synthesis at defined origins is a universal feature found in bacteriophages and viruses, prokaryotes, archaea, and eukaryotic cells (Costa et al., 2013). However, cells can survive without an active origin of replication. A recent study from the Allers lab (Hawkins et al., 2013) reported that *Haloferax volcanii*, a halophilic archaeon that grows in high salt environments under high osmotic pressure (Mullakhanbhai and Larsen, 1975), can not only tolerate deletion of all chromosomal origins, but grows with a doubling time faster than that of wild type cells (Hawkins et al., 2013). *Haloferax* cells contain a main chromosome, three secondary chromosomes, and a plasmid. High-resolution MFA revealed that the main chromosome is replicated from three origins, with a laboratory isolate showing a fourth, ori-pHV4, which is located in an integrated plasmid (Hawkins et al., 2013).

The deletion of single origins resulted in only mild growth rate reductions (Hawkins et al., 2013), as observed in other archaea (Samson et al., 2013; Wu et al., 2014). In contrast, deletion of multiple origins resulted in improved growth rates, and a derivative in which all replication origins were deleted grew faster than wild type cells, an effect that appears to be driven by recombination-dependent replication (Hawkins et al., 2013), replication that initiates at recombination intermediates (Hawkins et al., 2013; Michel and Bernander, 2014).

The ability to grow in the absence of replication origins is not a new finding. Kogoma and coworkers discovered that DNA intermediates involved in transcription (R-loops) and recombination (D-loops) can act as initiation points for DNA replication in *E. coli* (Kogoma and von Meyenburg, 1983). This type of synthesis was called constitutive stable DNA replication,

or cSDR (Kogoma, 1997). DNA synthesis observed following DNA damage is a second type of stable DNA replication. This type requires induction of the SOS DNA damage response and was termed induced SDR (iSDR) (Kogoma, 1997).

Kogoma and co-workers described that cSDR in *E. coli* cells lacking RNase HI is persistent enough to allow successful cellular replication in the absence of an active *oriC* (Kogoma, 1997). It was suggested that the initiation at R-loops is the main driver of chromosome replication in these cells, because RNase HI specifically degrades RNA from DNA:RNA hybrids (de Massy et al., 1984; Kogoma, 1997; Tadokoro and Kanaya, 2009). In line with this idea, cSDR is also found in cells lacking the *topA* gene, which encodes for topoisomerase I. Topoisomerase I relaxes negative supercoiling to prevent the persistence of DNA-RNA hybrids. Consequently, cells lacking topoisomerase I show hyper-negative supercoiling, increased levels of R-loops, and cSDR (Brochu et al., 2018). R-loops can also arise when transcription fails to terminate. In *E. coli*, Rho-dependent transcription termination acts as a surveillance mechanism to keep pervasive transcription in check, which may otherwise lead to the formation of R-loops (Leela et al., 2013). Such R-loops may provide nucleating points for cSDR. Indeed, in strains mutated for *rho*, plasmids with a ColE1-like replication origin, which relies on R-loop formation for the initiation of synthesis, undergo runaway plasmid replication, and a combination of *rho* with other genes involved in R-loop removal caused synthetic lethality (Harinarayanan and Gowrishankar, 2003).

In recent studies, replication profiles revealed in more detail the locations where cSDR is initiated, which are reasonably well-defined, including one particularly strong site roughly 500–600 kb clockwise from *oriC* at ∼4.5 Mbp, as well as a peak of synthesis in the termination area (Maduike et al., 2014; Dimude et al., 2015; Veetil et al., 2020). Despite a detailed analysis of the locations of initiation sites, the precise molecular mechanism that triggers the initiation of DNA synthesis in these defined locations is not fully understood (Maduike et al., 2014; Dimude et al., 2015; Veetil et al., 2020). But the synthesis observed is strong enough to allow continuous replication of the entire chromosome in the absence of *oriC* firing, and cells lacking the *rnhA* gene, which encodes for RNase HI, can tolerate the deletion of the entire *oriC* area (Kogoma, 1997; Dimude et al., 2015). However, growth of ΔrnhA cells in the absence of *oriC* firing is slow and growth of dnaA(ts) ΔrnhA cells at restrictive temperature is sensitive to rich medium, such as LB broth (Kogoma, 1997). This broth-sensitivity can be partially alleviated by an rpoB*35 allele (Dimude et al., 2015), a point mutation in the *rpoB* gene which encodes for the β subunit for RNA polymerase and which destabilizes ternary RNA polymerase complexes (Trautinger et al., 2005; Rudolph et al., 2007a). In addition, cells lacking RNase HI were reported to be synthetically lethal when also missing the homologous recombination proteins RecBCD (Itaya and Crouch, 1991), and this synthetic lethality can again be partially suppressed by an rpoB*35 (called rpo* hereafter for simplicity) point mutation (Dimude et al., 2015), indicating that replication-transcription conflicts are a strong contributor to these effects.

RecBCD is involved in homologous recombination and is a key component needed for the processing of double-stranded DNA ends (Singleton et al., 2004). It binds to blunt or near-blunt double-stranded DNA (dsDNA) substrates (Dillingham and Kowalczykowski, 2008). RecB and RecD are both helicases, but they have different polarities: RecB is a 3′ to 5′ helicase, while RecD translocates in 5′ to 3′ direction (Dillingham and Kowalczykowski, 2008). Available dsDNA ends will be unwound and very rapidly degraded by the RecBCD complex (Dillingham and Kowalczykowski, 2008; Wiktor et al., 2018) until a chi site is reached (Smith, 2012). Chi sites are asymmetric octamers which can inhibit the degradation of the 3′ end by RecBCD while degradation of the 5′ end proceeds. Thus, upon reaching a chi site, degradation by RecBCD is modified so that a 3′ ssDNA overhang suitable for the loading of RecA recombinase is produced (Singleton et al., 2004).

It has become clear that RecBCD is very important for the resolution of intermediates that arise from replication-transcription conflicts (Syeda et al., 2016). RecBCD proved to be essential for the viability of fast-growing *E. coli* cells, in which one of the rrn operons was artificially inverted to force head-on replication-transcription encounters (De Septenville et al., 2012). The fact that Δ*oriC* ΔrnhA cells are broth sensitive and that ΔrecB ΔrnhA cells are synthetically lethal, with both effects being partially alleviated by an rpo* point mutation (Dimude et al., 2015), strongly suggests that DNA synthesis triggered at R-loops in cells lacking RNase HI in chromosomal areas away from *oriC* suffers from collisions with transcribing RNA polymerase complexes and requires processing by DNA repair and recombination proteins. Similarly, cells lacking Dam methylase, which has a role in strand-discrimination for methyl-directed mismatch repair, can grow in the absence of a functional origin, an effect that is likely to be caused by recombination-dependent replication triggered at now undirected MMR repair sites (Raghunathan et al., 2019). Analogously to cells lacking Rnase HI, an rpo* point mutation is one important factor that is required for Δdam cells to grow in the absence of *oriC* firing (Raghunathan et al., 2019).

REPLICATION OBSTACLES IN CELLS CARRYING THE ECTOPIC REPLICATION ORIGIN ORIZ

While the deletion of all origins in *Haloferax* appears to allow faster growth of cells, at least under laboratory conditions (Hawkins et al., 2013), the same is not the case in bacteria such as *E. coli* and *B. subtilis*. Cells being forced to use initiation sites other than *oriC*, such as ΔrnhA cells in the absence of *oriC* firing, suffer considerable problems. Indeed, previous studies in *B. subtilis* where DNA replication initiated exclusively at an ectopic origin showed a substantial delay of replication at highly transcribed rrn operons encountered in an orientation opposite to normal (Wang et al., 2007; Srivatsan et al., 2010).

In Δ*oriC* oriZ^{+} cells the chromosome is replicated exclusively from the ectopic oriZ, very similar to the described situation in *B. subtilis*. It was therefore a surprise when Wang et al. (2011)

reported that Δ*oriC* oriZ$^+$ cells grew with a doubling time very similar to that of wild type cells. Indeed, when we re-generated a Δ*oriC* oriZ$^+$ construct, we found its doubling time to be over 40 min. Δ*oriC* oriZ$^+$ cells seriously struggle to grow and rapidly accumulate suppressor mutations that allow faster growth (Ivanova et al., 2015).

The replication profile of Δ*oriC* oriZ$^+$ cells revealed two major obstacles to replication. The asymmetry of the replichore arrangement is even more extreme in Δ*oriC* oriZ$^+$ cells than in *oriC*$^+$ oriZ$^+$ cells, as the fork traversing counterclockwise has to replicate 3/4 of the entire chromosome. Consequently, the "step" within *terA* and *terC* is strongly pronounced (Ivanova et al., 2015). But replication initiated at oriZ and traversing the chromosome counter-clockwise also encounters the highly transcribed rrnH and rrnCABE operons in an orientation opposite to normal (**Figure 2A**), resulting in significant problems (Ivanova et al., 2015), in line with results in *B. subtilis* (Wang et al., 2007, 200; Srivatsan et al., 2010). A clear prediction of these observations is that the slow growth phenotype of Δ*oriC* oriZ$^+$ cells should be suppressed by two classes of mutations: the inactivation of the replication fork trap as well as any mutation that causes a reduction of the severity of conflicts between replication and transcription. This is indeed what we observed. The slow growth phenotype of Δ*oriC* oriZ$^+$ cells was partially suppressed by the inactivation of the replication fork trap (Δtus) and an rpo* point mutation (Ivanova et al., 2015). However, the fast growth of the original Δ*oriC* oriZ$^+$ construct by Wang et al. (2011) was caused by a suppressor mutation that solved the problem in a far more elegant way: their fast growing Δ*oriC* oriZ$^+$ strain carried a substantial inversion. This inversion spanned, with the exception of rrnH, almost the entire remaining portion of the chromosome that would have been replicated in the wrong orientation from oriZ, including the entire rrnCABE operon cluster. Thus, the problem in these cells was solved simply by the re-alignment of replication and transcription (**Figure 3A**; Ivanova et al., 2015), strongly supporting to the notion that avoiding head-on collisions has significantly contributed to shaping the distinct architecture of bacterial chromosomes.

This idea is further supported by the fact that a variety of different repair systems are present in cells dedicated to dealing with tightly bound DNA-protein complexes. In *E. coli,* a variety of helicases promote fork progression through tightly bound nucleoprotein complexes, including Rep, UvrD, and DinG (Guy et al., 2009; Boubakri et al., 2010; Atkinson et al., 2011). Rep is considered an accessory replicative helicase because Rep physically associates with the replicative helicase, DnaB (Bochman et al., 2010; Brüning et al., 2014; Syeda et al., 2019). Chromosome duplication takes almost twice as long in Δrep cells than in wild type cells (Lane and Denhardt, 1975; Guy et al., 2009; Atkinson et al., 2011). In addition, enzymes involved in homologous recombination play an important role in assuring that replication forks move successfully through highly transcribed areas (Cox, 2001; Dillingham and Kowalczykowski, 2008; Boubakri et al., 2010; De Septenville et al., 2012; Michel et al., 2018), as already discussed above.

Why are a variety of repair systems needed to deal with stably bound DNA-protein complexes? When we investigated viability and replication profiles of cells lacking Rep helicase, we found a much-increased origin/terminus ratio (cf. **Figures 4Ai,ii**; Dimude et al., 2018a). The replication profiles reveal no specific areas that appear problematic. As replication profiles are population-based, this observation suggests that Rep acts on average at sites relatively evenly distributed throughout the chromosome. However, the replication profile of *oriC*$^+$ oriZ$^+$ Δrep cells revealed that the progression of DNA replication is very effectively blocked by rrn operons encountered in a head-on orientation, as indicated by the rather abrupt change of the replication gradient at rrnH (cf. **Figures 4Bi,ii**). Indeed, Δ*oriC* oriZ$^+$ Δrep cells are inviable (Dimude et al., 2018a). Viability is restored by an rpo* mutation in which replication-transcription conflicts are lessened, and the replication profiles show that synthesis can indeed proceed (Dimude et al., 2018a).

rrn operons encountered in a head-on orientation in cells lacking RecBCD block replication even more severely than in cells lacking Rep, and there is no indication of replisomes proceeding past rrnH, the first rrn operon encountered (**Figure 4Ci**). Δ*oriC* oriZ$^+$ ΔrecB cells are inviable unless an rpo* point mutation is present, but even then, cells can only survive in minimal medium, in which a reduced growth rate means a slower doubling time and a reduced demand for rRNA in comparison to growth in rich medium. They remain synthetically lethal in LB and our replication profiles show that replication proceeds past rrnH with a low frequency, low speed, or both (Dimude et al., 2018a).

Replication profiles of *oriC*$^+$ oriZ$^+$ ΔrecB cells also showed a much-reduced peak height of oriZ, while firing of *oriC* appeared to be unaffected (**Figure 4Ci**). Peak height was restored in cells also lacking the exonuclease SbcCD (**Figure 4Cii**), indicating that extensive SbcCD-dependent degradation takes place in the absence of RecBCD at replication forks arrested at highly transcribed rrn operons (Dimude et al., 2018a).

Taken together, the data currently available suggest that replication-transcription conflicts can trigger different type of arrested forks, depending, for example, on the level of transcription. Indeed, it was shown that the mode of protein displacement of nucleoprotein complexes by RecBCD helicase/exonuclease varies depending on overall protein density (Terakawa et al., 2017). The different types of arrested of perhaps even collapsed replisomes then will require different types of processing that have to take place (2016; Dimude et al., 2018a). Replication coming from oriZ will encounter several genes that are transcribed in an orientation opposite to normal, and both co-directional as well as head-on conflicts are problematic (Merrikh et al., 2011, 2012; Lang and Merrikh, 2018). Nevertheless, there is no indication of any substantial block to replication in ΔrecB cells until the first highly transcribed region is reached. It appears that Rep helicase is sufficient to facilitate replisome progression through these areas. But when forks encounter a rrn operon in an orientation opposite to normal the situation differs significantly. The intermediates generated in

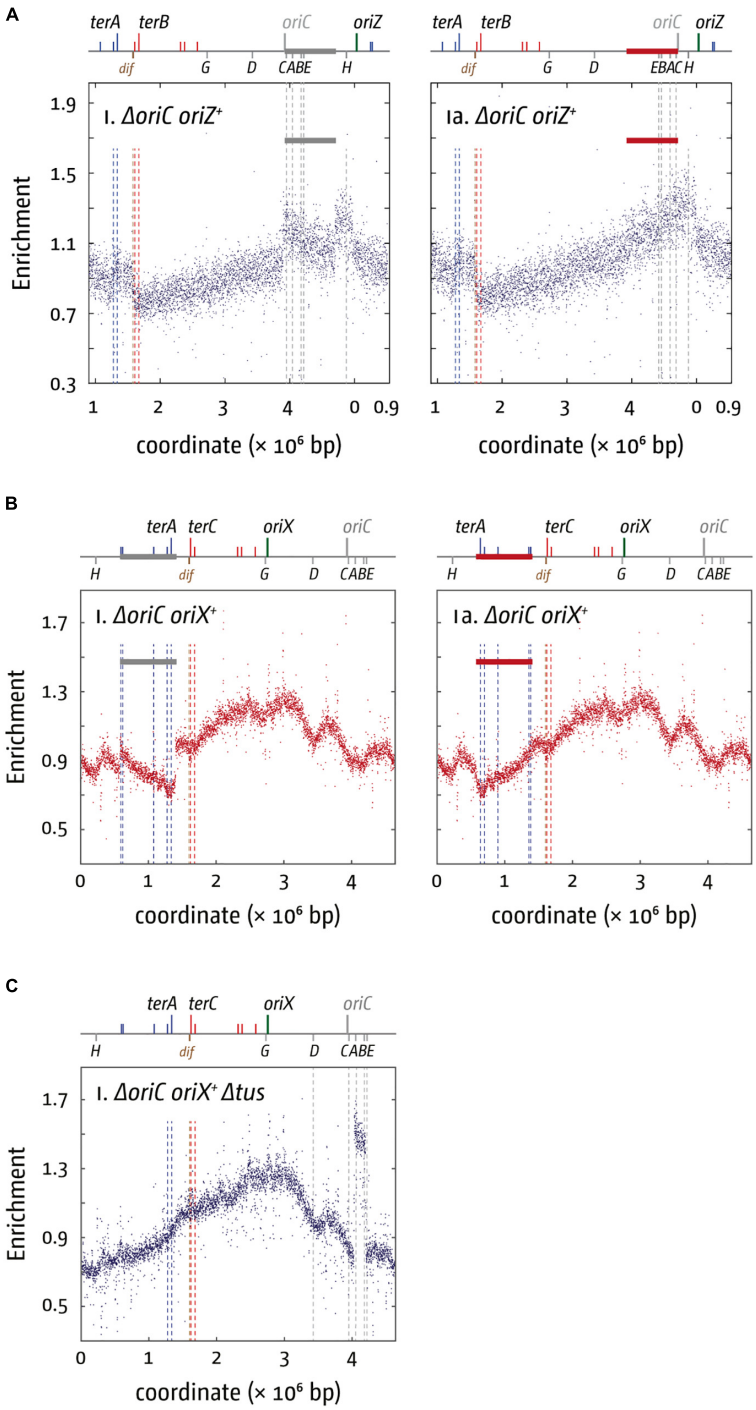


FIGURE 3 | Chromosomal rearrangements in *E. coli* cells replicating from a single ectopic replication origin. **(A)** Replication profiles of *E. coli* cells with a single ectopic replication origin. Shown is the marker frequency analysis of *E. coli* Δ*oriC* oriZ^{+} cells. The number of reads (normalized against reads for a stationary phase wild type control) is plotted against the chromosomal location. A schematic representation of the *E. coli* chromosome showing positions of *oriC* (gray to indicate the deletion) and oriZ (green line) and ter sites (above) as well as dif and rrn operons A–E, G, and H (below) is shown above the plotted data. A clear discontinuity of the profile can be seen in **(panel i)** (marked by a gray bar), which is due to a large inversion, as highlighted by the continuous replication profile that results if the area highlighted (red bar indicates the inverted area) is inverted. The MFA raw data were taken from Ivanova et al. (2015) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome. **(B)** Replication profiles of *E. coli* Δ*oriC* oriX^{+} cells. A clear discontinuity of the profile can be seen in **panel i** (marked by a gray bar), which is due to a large inversion, as highlighted by the continuous replication profile that results if the area highlighted (red bar indicates the inverted area) is inverted. The MFA raw data were taken from Dimude et al. (2018b) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome. **(C)** Replication profiles of *E. coli* Δ*oriC* oriX^{+} Δtus cells. A clear discontinuity of the replication profile can be seen between the rrn operons A and B, which is due to a duplication of the entire region. See text for details.

this situation appear to be accessible to degradation by SbcCD (Dimude et al., 2018a) and other nucleases such as RecJ (De Septenville et al., 2012) and are extensively resected. In addition, ΔuvrD Δrep rpo* cells can only survive in the presence of both RecBCD and RecA, suggesting that the loading of RecA by RecBCD is required for the continuation of DNA replication (Syeda et al., 2016). Thus, both the failing to load RecA and the extensive resection contribute to *rrnH* being such a severe block in *oriC*$^+$ oriZ$^+$ ΔrecB cells (Dimude et al., 2018a). In contrast, no obvious resection is observed in cells lacking Rep (Dimude et al., 2018a).

It was suggested that DinG is an additional protein that is involved in aiding the progression of replication through highly transcribed areas of the chromosome (Baharoglu et al., 2010; Boubakri et al., 2010). Indeed, we were able to show that Δ*oriC* oriZ$^+$ ΔdinG cells are synthetically lethal, an effect robustly suppressed by a rpo* point mutation. This result supports the idea that DinG is involved in underpinning replication of highly transcribed areas in *E. coli*. However, DinG is unable to directly promote replisome movement through stalled transcription complexes *in vitro*, and the replication profile of *oriC*$^+$ oriZ$^+$ ΔdinG cells do not reveal any abnormalities at rrnH (Hawkins et al., 2019), much in contrast to cells lacking either Rep or RecB (Dimude et al., 2018a). Thus, it seems that DinG might have an indirect effect in resolving replication-transcription encounters, potentially via its ability to unwind RNA:DNA hybrids (Voloshin and Camerini-Otero, 2007).

REPLICATION OBSTACLES IN CELLS CARRYING THE ECTOPIC REPLICATION ORIGIN ORIX

The results described so far strongly support the idea that replication-transcription conflicts are an important factor that have contributed to shaping the structure of bacterial chromosomes. In line with this idea, replication-transcription conflicts came up again when we tried to generate Δ*oriC* oriX$^+$ cells. One rationale of integrating an ectopic replication origin into the left-hand replichore in the first place was the fact that there is only a single rrn operon (rrnD) between the integration location and *oriC*, together with a cluster of highly transcribed genes that code for ribosomal proteins. Thus, we speculated that replication-transcription conflicts might be less severe in this particular construct, whereas replication in Δ*oriC* oriZ$^+$ cells has to overcome 5 highly transcribed rrn operons. However, the fact that the rpo* mutation improved doubling times of various Δ*oriC* oriX$^+$ constructs suggests that conflicts still have a considerable impact (Dimude et al., 2018b).

Our studies in oriX cells revealed that, beside replication-transcription conflicts, the replication fork trap severely impacts on genome duplication in *oriC*$^+$ oriX$^+$ and Δ*oriC* oriX$^+$ cells. In fact, similar to the situation in Δ*oriC* oriZ$^+$ cells initially described (Wang et al., 2011; Ivanova et al., 2015), we found that our Δ*oriC* oriX$^+$ construct contained a large inversion. This inversion spanned all blocking ter sites and flipped them into permissive orientation. Thus, the inversion allows replication to proceed unhindered (**Figure 3B**), demonstrating the impact of the replication fork trap on replication progression (Dimude et al., 2018b). The inversion also re-aligns the direction of replication and transcription in the way it is in *oriC* cells, and both the replication fork trap and replication-transcription conflicts might be an important factor here. However, if transcription generally interferes with replication, a prediction is that for both *oriC*$^+$ oriX$^+$ Δtus and *oriC*$^+$ oriZ$^+$ Δtus cells forks escaping the termination area should be slowed down, as their progression into the opposite replichore would force an increased number of head-on collisions. If forks escaping the termination area are slower than forks coming from the native *oriC*, the fork fusion point should be shifted from the location equidistant to both origins toward the termination area. However, this is not what we observed. In *oriC*$^+$ oriX$^+$ Δtus cells the fork fusion point was close to the arithmetic mid-point between *oriC* and oriX and only slightly shifted toward the termination area (∼20 kb) (Dimude et al., 2018b), while for *oriC*$^+$ oriZ$^+$ Δtus cells forks terminated 60 kb in the direction of *oriC* (Ivanova et al., 2015; Dimude et al., 2016). We do not have any direct information about the speed of individual forks, but these results suggest that the forks leaving the termination area and traveling in the wrong orientation have, on average, a similar speed to the forks coming from *oriC* (oriX) or are even slightly faster (oriZ) (Ivanova et al., 2015; Dimude et al., 2016, 2018b), similar to the situation observed in *Vibrio cholerae* where replication forks simply fused opposite the origin even when the origin was moved to an ectopic location (Galli et al., 2019).

A clue for an additional factor that might contribute to replication dynamics and genome structure comes from the observation that one of our Δ*oriC* oriX$^+$ Δtus constructs had acquired a spontaneous duplication of the chromosomal stretch containing rrn operons *A* and *B* (**Figure 3C**). Highly transcribed genes tend to be located in relative vicinity to the origin. In fast growing cells this area can be in a ratio of four to one relative to the termination area or even higher. The increased number of gene copies results in a gene dosage effect (Jin et al., 2012). The rrn operons *CABE* and *D* are all located in close proximity to *oriC*, causing an increased gene dosage in fast-growing cells (Jin et al., 2012). If, however, the origin is shifted from its original location into the left-hand replichore, rrn operons CABE and H are all in quite a distance from the active origin, which results in a lower copy number. This effect will be less pronounced in Δ*oriC* oriZ$^+$ cells, because the location of oriZ is in close proximity to rrnH and the rrnCABE cluster. It was reported before that inactivation of up to three of the rrn operons in *E. coli* caused significant upregulation of the remaining rrn operons, thereby compensating for the reduced copy number (Condon et al., 1993). However, especially if multiple rrn operons are affected, a reduced growth rate was observed (Condon et al., 1993).

Chromosomal replication starting exclusively at oriX will transfer especially the rrnCABE cluster and rrnH into a completely different chromosomal environment, as movement of the *oriC* position was shown to alter the position of the chromosomal macrodomains (Duigou and Boccard, 2017). Indeed, it was shown that expression of a reporter cassette under

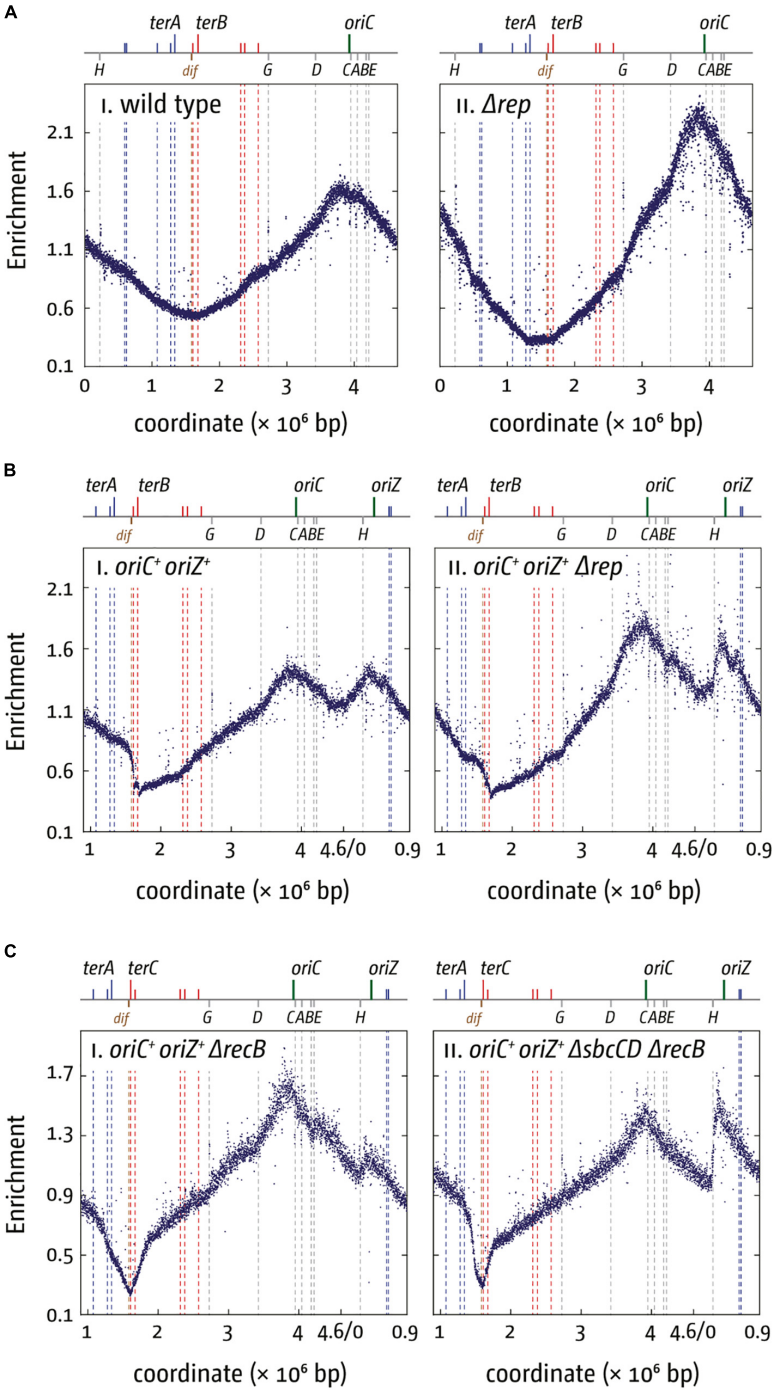


FIGURE 4 | Replication dynamics and cell viability in cells with one or two active replication origins lacking either Rep helicase or RecBCD exonuclease. **(A)** Cells lacking Rep helicase show an increased origin/terminus ratio than wild type cells, indicating that replication fork progression is significantly slowed. The replication profiles are generated by plotting the number of sequence reads (normalized against reads for a stationary phase wild type control) against their chromosomal location. The schematic representation of the *E. coli* chromosome above each panel shows the positions of the two origins, *oriC* and oriZ, and ter sites (above) as well as the dif chromosome dimer resolution site and rrn operons A–E, G, and H (below). **(B)** Replication fork progression is blocked at the highly transcribed rrnH operon replicated in a direction opposite to normal in *oriC*$^{+}$ oriZ^{+} cells lacking Rep helicase. Please note that the chromosomal coordinates are shifted in comparison to panel **(A)** so that *oriC* and oriZ next to each other. **(C)** Replication fork progression is arrested at rrnH if replication proceeds in an orientation opposite to normal, and oriZ peak height is much reduced in cells lacking RecBCD exonuclease **(panel i)**. oriZ peak height is restored if SbcCD is missing in addition to RecBCD **(panel ii)**. See text for details. For an in-depth discussion of the underrepresentation of sequence reads in the termination area please refer to Wendel et al. (2014), Sinha et al. (2017, 2018), and Dimude et al. (2018a). All raw data in panels **(A–C)** are taken from Dimude et al. (2018a) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome. As for panel **(B)**, please note that the chromosomal coordinates for panel **(C)** are shifted in comparison to panel **(A)** so that *oriC* and oriZ are next to each other.

control of the lac promoter showed a 300-fold variation in transcription levels depending on its precise integration location into the chromosome (Bryant et al., 2014; Scholz et al., 2019), and displacement of pleiotropic genes were indeed shown to affect the phenotype and competitive growth fitness of cells (Gerganova et al., 2015). If rrnCABE and H are less transcribed in Δ*oriC* oriX$^+$ cells this might contribute to the explanation why we struggled particularly with the generation of Δ*oriC* oriX$^+$ cells, and it would suggest that the observed duplication of rrnA and B are indeed beneficial to the competitive fitness of our Δ*oriC* oriX$^+$ Δtus construct (Dimude et al., 2018b). It might also explain why the inversion found in the initial Δ*oriC* oriZ$^+$ construct generated in the Sherratt Lab (Wang et al., 2011) is a particularly efficient suppressor, as it not only realigns replication and transcription, but also brings the rrnCABE cluster back into close proximity of the only active origin, oriZ (Ivanova et al., 2015).

MAKING SENSE OF THE REPLICATION FORK TRAP

The gross chromosomal rearrangement in Δ*oriC* oriX$^+$ cells that flipped all ter sites from blocking into permissive orientation strongly highlights the constraint imposed by such a replication fork trap on genome duplication (Dimude et al., 2018b). Any arrest of one of the two forks cannot be alleviated by simply waiting until the second fork arrives, as this fork will be blocked by the fork trap. However, ter/Tus complexes are not systematically involved when replication forks fuse. Early labeling experiments (Bouché et al., 1982), and more recently MFA (Rudolph et al., 2013; Ivanova et al., 2015; Dimude et al., 2016), indicates that in wild type *E. coli* cells the majority of forks fuse close to the arithmetic mid-point, somewhere between the dif chromosome dimer resolution site and *terC* (Rudolph et al., 2013; Ivanova et al., 2015; Dimude et al., 2016). Thus on a population basis both the clockwise and counterclockwise fork appear to move normally with similar speeds, which results in a fusion of two freely moving replisomes within the innermost ter sites, at least under laboratory conditions. It seems that the fork trap mostly comes into play upon a delay of one of the two forks at an obstacle, such as a nucleoprotein complex or a DNA lesion.

If the replication fork trap is not systematically involved in termination what might be its physiological role? Together with a single origin of replication, it certainly contributes to strictly maintaining replicational directionality within the two replichores, and it was suggested that a fork trap is important to maintain the co-directionality of transcription and replication (Brewer, 1988; Rudolph et al., 2007a). Given the strong impact of replication-transcription clashes described above and elsewhere, and the many repair pathways dealing with such conflicts, this will be an important factor (McGlynn et al., 2012; Merrikh et al., 2012; Lang and Merrikh, 2018).

However, it appears that many bacterial species do not utilize a dedicated fork trap (Galli et al., 2019). And, in *E. coli*, genome-wide co-directionality of replication and transcription is only approximately 55% (McLean et al., 1998). The vast majority of highly transcribed genes are transcribed co-directionally with replication (McLean et al., 1998), but all rrn operons and the majority of genes encoding for ribosomal proteins are in relative proximity to the origin (Jin et al., 2012; Dimude et al., 2016). Thus, a fork escaping the fork trap in *E. coli* would have to proceed for about 1 Mbp (1/4 of the chromosome) before any of the genes transcribed at very high levels would be reached (**Figure 1A**). Indeed, as highlighted above, replication in the vicinity of the termination area appears to proceed with speeds similar to replication coming from *oriC*, both in the left- and right-hand replichore (Ivanova et al., 2015; Dimude et al., 2018b). This observation does not rule out replication-transcription conflicts, as forks coming from *oriC* might also suffer from delays and we did not directly measure fork speed. However, for tRNA genes, which are highly transcribed under fast growth conditions and which are more globally distributed throughout the chromosome, we found a co-directionality of replication and transcription in the origin-proximal half of the chromosome only. In the origin-distal half relative orientation of replication and transcription is much more variable. Indeed, we were surprised to find a mild bias toward the head-on orientation for replication coming from *oriC* (Dimude et al., 2016). Thus, while avoiding clashes between replication and transcription is important, it remains debatable whether avoiding such clashes is the main purpose of the fork trap in *E. coli*.

Is the absence of the replication fork trap causing any phenotypes which might shed light on its physiological role? When working with oriX$^+$ and oriZ$^+$ strains we noticed that deletion of tus consistently caused a mild growth defect (Ivanova et al., 2015; Dimude et al., 2018b). This suggests that restricting fork movement in the termination area somehow facilitates replication completion or successful chromosome segregation or both.

One process that is uniquely happening in the termination area is the fusion of the two replication forks. Could this process itself, or some unwanted side effect, be responsible for the observed delay? Various experimental approaches have shown that an absence of functional ter/Tus complexes can result in replication still occurring when it is meant to stop. Because replication continues to occur when a complete copy of the DNA is generated, we call this continued synthesis over-replication, as it over-replicates molecules that are already fully replicated. This was observed for plasmid R1 in *E. coli*. R1 is replicated unidirectionally by a single fork until it gets arrested at a single ter/Tus complex close to the plasmid origin (Nordström, 2006). Inactivation of this stopping point for replication allows synthesis to proceed into an already replicated area, and this was shown to result in the accumulation of branched DNA structures, rolling circle replication intermediates and the formation of plasmid multimers (Krabbe et al., 1997). It was suggested that, upon reaching an already replicated area, the replicative helicase of the fork might displace already existing nascent strands. The resulting intermediates can then serve as substrates at which additional synthesis can proceed (Krabbe et al., 1997). Similarly, it was shown *in vitro* that for a plasmid substrate functional ter/Tus complexes efficiently prevented over-replication and the formation of complex intermediates (Hiasa and Marians, 1994), a

result that was recently confirmed in a reaction using the elegant "replication chain reaction" (Hasebe et al., 2018). These results indicate that the fork trap can prevent unwanted over-replication that is linked to termination of DNA synthesis.

Similarly, it was found that Δtus cells showed chromosomal over-replication, even though at a low level (Markovitz, 2005). This effect was exacerbated by point mutations in DNA polymerase I (Markovitz, 2005), which has a prominent role in the repair of DNA damage and the maturation of Okazaki fragments (Kurth and O'Donnell, 2009), leading to the suggestion that Pol I might be involved in bringing DNA replication to a successful conclusion in the terminus region (Markovitz, 2005). Results from *B. subtilis* suggest that the absence of the Rtp terminator protein can result in the formation of an increased number of chromosomal dimers (Lemon et al., 2001; Duggin et al., 2008). Since over-replication results in the generation of double-stranded DNA ends accessible to homologous recombination (**Figure 6** and below), the increased formation of chromosome dimers could be a result of problems with fusing replisomes, similar to the situation in *E. coli*.

An even stronger effect of a Δtus mutation was found in cells lacking RecG helicase. The replication profile of ΔrecG cells shows a peak of over-replication within the four innermost ter sites (cf. **Figures 5Ai,ii**; Rudolph et al., 2013; Wendel et al., 2014; Dimude et al., 2015, 2016; Midgley-Smith et al., 2018b). Indeed, this over-replication can support growth in the absence of a functional origin if a) a functional replication fork trap is absent (Δtus) and b) replication-transcription conflicts resulting from forks leaving the termination area and proceeding in an orientation opposite to normal are alleviated (rpo*) (Rudolph et al., 2013). In the absence of *oriC* activity ΔrecG Δtus rpo* cells show a replication profile that is inverted: the *oriC* area shows a low-point of the profile while, rather paradoxically, the highest point of the profile is observed in the termination area where forks normally fuse to end DNA synthesis (**Figure 5B**; Rudolph et al., 2013; Dimude et al., 2015).

Our genetic analysis of the over-replication in ΔrecG cells suggests that it is triggered by intermediates which are similar to those proposed for replication of plasmid R1 (Krabbe et al., 1997; Rudolph et al., 2013; Dimude et al., 2015, 2016; Lloyd and Rudolph, 2016; Midgley-Smith et al., 2018b). We believe that upon fusion of two replication forks an intermediate is generated that allows either the continuation of synthesis or the re-recruitment of new forks. The over-replication in ΔrecG cells strictly requires the ability of the main restart protein PriA to process a 3′ flap structure (Rudolph et al., 2013). In addition, we observed that over-replication also occurs in cells lacking 3′ exonucleases Exo I, Exo VII, and SbcCD (Rudolph et al., 2010a, 2013; Midgley-Smith et al., 2018a). These results indicate that a 3′ flap might be a central intermediate. We have proposed that such a 3′ flap might arise upon the fusion of two forks by the displacement of the nascent leading strand of one of the two forks by the replicative helicase of the other (**Figure 6B**). 3′ flaps were shown to be a very good substrate for RecG helicase *in vitro* (McGlynn and Lloyd, 2001; Tanaka and Masai, 2006; Rudolph et al., 2010b; Bianco, 2015) and, in its presence, would be rapidly converted into 5′ flaps or, alternatively, degraded by 3′ exonucleases (**Figure 6B**; Rudolph et al., 2013; Dimude et al., 2016; Midgley-Smith et al., 2018a). If a 3′ flap remains unprocessed, PriA might gain access and re-recruit a replisome (**Figure 6C**), leading to the observed over-replication of the termination area. However, such newly initiated synthesis would generate double-stranded DNA ends (**Figure 6C**). dsDNA ends will be rapidly processed by RecBCD and RecA, resulting in the formation of a D-loop (Rudolph et al., 2009, 2010a, 2013; Dimude et al., 2016), another substrate at which PriA can establish a functional replisome (**Figure 6D**). Progression of forks established in this way will proceed until they get blocked at a ter/Tus complex (**Figure 6D**).

Rather than by fork fusion events themselves, might the over-replication be caused by a cryptic origin that is normally suppressed, or by the increased occurrence of R-loops within the termination area, as recently suggested (Kuzminov, 2016)? While remaining a possibility, it is unlikely for a number of reasons. Firstly, we observed that linearization of the chromosome within the termination area much reduced the over-replication both in ΔrecG cells and in cells lacking 3′ exonucleases (Rudolph et al., 2013; Dimude et al., 2015; Midgley-Smith et al., 2018a). While linearization would prevent two replisomes from fusing, it will not interfere with the activity of a cryptic origin, and we have indeed observed that linearization of the chromosome in cells lacking RNase HI does not abolish the R-loop-driven over-replication in the termination area (Dimude et al., 2015). Secondly, we observed that over-replication in the termination area is dramatically exacerbated in ΔrecG cells if oriZ is introduced (**Figure 5C**; Rudolph et al., 2013; Midgley-Smith et al., 2018b). It is not clear how integration of an origin ∼1 Mbp away from the termination area should cause such a dramatic increase in activity of either a cryptic origin or a hot spot for R-loop formation, while it clearly changes fork fusion events (Midgley-Smith et al., 2018a,b). Thirdly, we were recently able to demonstrate that over-replication can indeed be triggered outside of the termination area in *oriC*$^+$ oriZ$^+$ ΔrecG cells. In *oriC*$^+$ oriZ$^+$ cells, a second fork fusion event takes place in an ectopic location, and we were able to show that this ectopic fork fusion event can also trigger over-replication (Midgley-Smith et al., 2018b). Thus the over-replication is not location-bound but can be observed in other chromosomal contexts if forks are forced to fuse in this area. In addition, it is also not clear how proteins such as 3′ exonucleases and DNA polymerase I, would be involved in suppressing a cryptic origin or R-loops (Markovitz, 2005; Rudolph et al., 2013; Midgley-Smith et al., 2018a,b). Taken together, we prefer the idea that fork fusion intermediates are responsible for triggering the over-replication observed, as it fits the available data much better than a cryptic origin or a R-loop hotspot.

If so, might the fork trap provide a defined chromosomal region where termination intermediates and the resulting over-replication can be contained and quickly and safely processed to bring DNA replication to an accurate conclusion? The termination area was found to be a recombination hotspot (Horiuchi et al., 1994), a result that would be easily explained if over-replication intermediates, that can arise

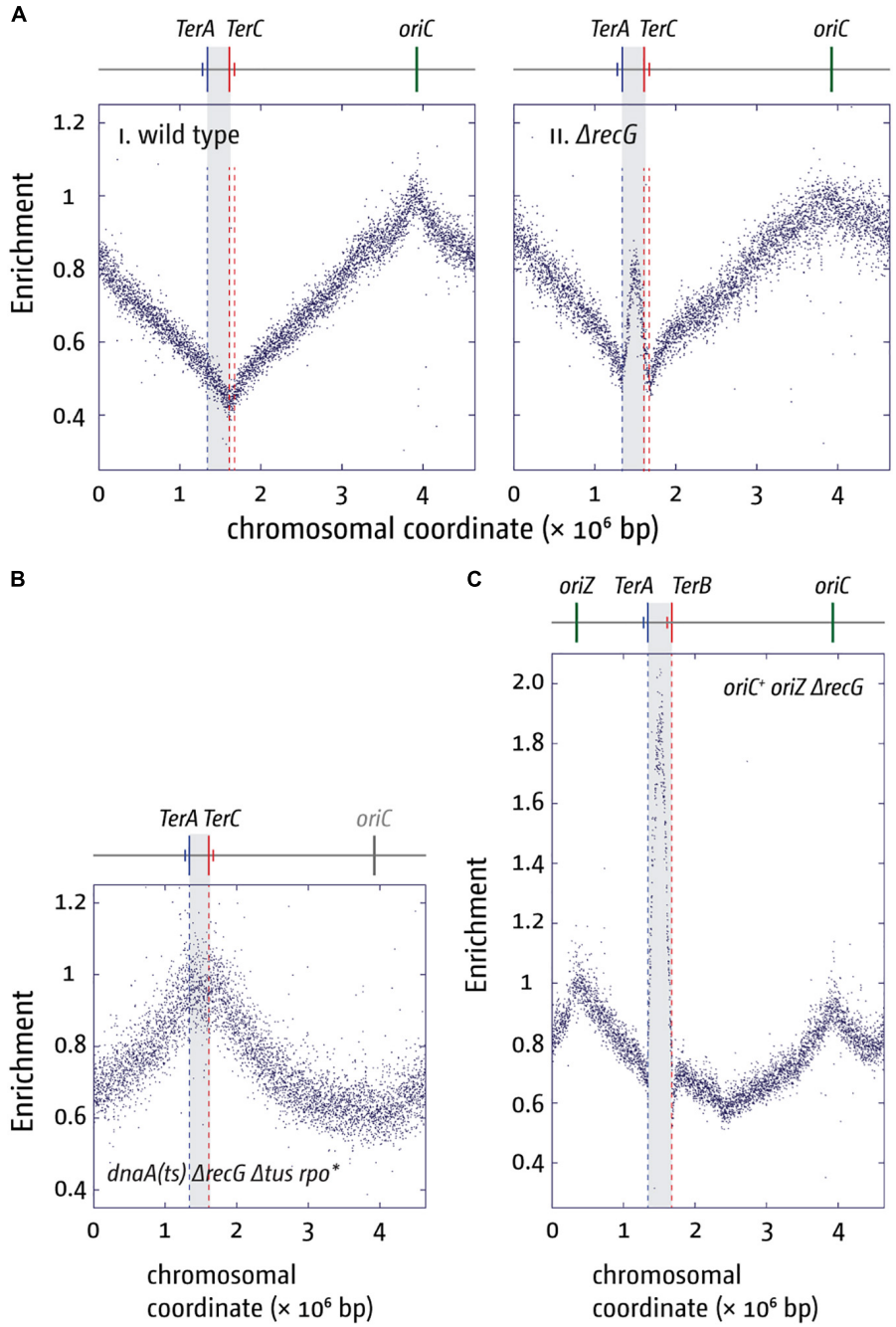


FIGURE 5 | Over-replication in the termination area in the absence of RecG helicase. **(A)** Replication profiles of *E. coli* cells in exponential phase. Cells were grown at 37°C. The number of reads (normalized against the reads for a stationary wild type control) is plotted against the chromosomal coordinate. Positions of *oriC* (green line) and primary ter sites are shown above the plotted data with red and blue lines representing the left and right replichore, as depicted in **Figure 1A**. The termination area between the innermost ter sites is highlighted in light gray. **(B)** Marker frequency analysis of a ΔrecG Δtus rpo* strain that carries a temperature-sensitive allele of the main replication initiator protein DnaA. The strain was grown at 42°C to inactivate DnaA(ts) and therefore prevent *oriC* firing. **(C)** Marker frequency analysis of chromosome replication in *oriC*$^{+}$ oriZ^{+} strain in the absence of RecG. Strains were grown at 37°C. The raw data in panels **(A–C)** were taken from Rudolph et al. (2013) and re-plotted to allow changes the scale of the plots, if necessary, and to highlight specific schematic features of the *E. coli* chromosome.

occasionally despite the presence of all processing factors in wild type cells, would trigger increased levels of recombination (Rudolph et al., 2013; Dimude et al., 2016; Midgley-Smith et al., 2018a,b). Increased recombination frequencies as well as chromosomal over-replication contribute significantly to genomic instability (Finkel et al., 2007; Blow and Gillespie, 2008;

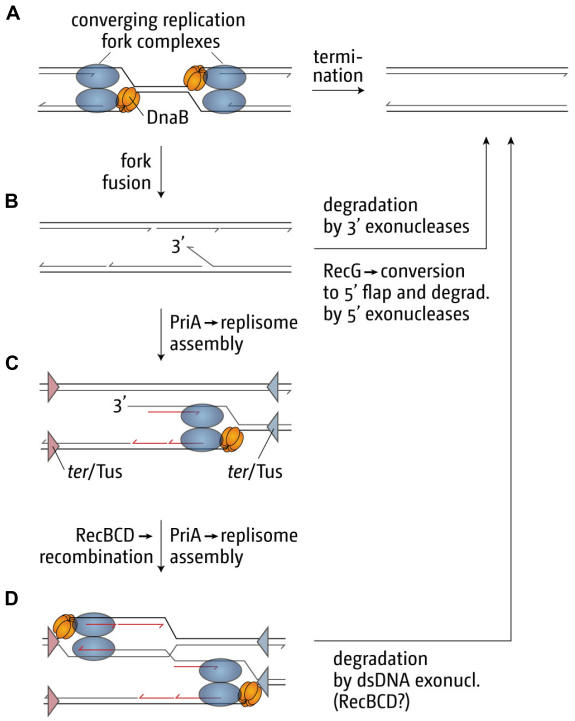


FIGURE 6 | Illustration of how replication fork fusions might trigger over-replication in the termination area and how this is normally prevented by proteins such as RecG and/or 3′ exonucleases. The fusion of two replisomes **(A)** can result in the formation of key intermediates, such as a 3′ single-stranded DNA flap **(B)**, which can be processed by restart proteins such as PriA **(C,D)** if it is not removed or degraded. ter/Tus complexes are shown in panels **(C,D)** as triangles. The blue ter/Tus complexes are oriented such that they would block synthesis initiated within the termination area and moving counterclockwise, while the red ter/Tus complexes would block clockwise synthesis. As these complexes are permissive for the forks coming from *oriC* in panel **(A)** they have been excluded for simplicity. Note that, while the formation of a 3′ flap can occur at both forks, only one such reaction was shown for simplicity. See text for details.

Alexander and Orr-Weaver, 2016; Tomasetti et al., 2017), highlighting why a fork trap might be beneficial. However, the relatively mild phenotype of cells lacking a fork trap system highlights that this effect is in addition to the various processing factors that are involved in the processing of fork fusion intermediates, such as RecG, 3′ exonucleases and DNA polymerase I. These proteins seem to be able to deal with occurring intermediates efficiently, which might explain why a fork trap system is found in only a limited number of bacterial species (Galli et al., 2019). Nevertheless, in species in which the opportunity arose (Galli et al., 2019), the fork trap system might have been a welcome addition, and as highlighted before, the effect on the doubling time, although small, is measurable (Ivanova et al., 2015; Dimude et al., 2018b).

The importance of the proteins dealing with fork fusion intermediates is highlighted by the synthetic lethality of cells that lack combinations of the proteins involved. Cells lacking both RecG and 3′ exonucleases are synthetically lethal (Rudolph et al., 2010a), as are cells lacking both RecG and DNA polymerase I (Hong et al., 1995; Zhang et al., 2010; Upton et al., 2014). Furthermore, we found that Δ*oriC* oriZ^{+} ΔrecG cells are synthetically lethal, an effect that is suppressed by the inactivation of the replication fork trap, suggesting that the lethality is caused by the vastly exacerbated levels of over-replication in the termination area (Midgley-Smith et al., 2018b). However, over-replication is not triggered by forks arrested at ter/Tus complexes, as it is still observed in ΔrecG cells lacking Tus terminator protein (Rudolph et al., 2013; Midgley-Smith et al., 2018b).

CONCLUDING REMARKS

While genome sizes and certain structural aspects of bacterial genomes show considerable variability, all bacterial chromosomes investigated so far have in common that they are duplicated by two replication forks initiated at a single origin (Gao and Zhang, 2008; Gao, 2015). Additional active origins can be introduced into the chromosome, but in the existing chromosome structure they always cause a disadvantage, such as a mild growth defect (oriX, oriZ) (Ivanova et al., 2015; Dimude et al., 2018b), silencing of one of the active origins (Kouzminova and Kuzminov, 2008) or causing some sort of toxicity to cells (oriY) (Dimude et al., 2018b). The observed problems are, at least in part, caused by genome trafficking problems, such as conflicts between replication and transcription, tightly bound protein-DNA complexes such as ter/Tus complexes, and other related issues, highlighting these processes as likely contributors of the overall structure of bacterial chromosomes, as suggested in a variety of other studies. Indeed, the finding that replication and transcription are aligned in human cells as well via the positioning of origins relative to highly transcribed genes (Chen et al., 2019) suggests that this is a very universal feature of nucleic acid metabolism. However, while the strict replichore arrangement in bacteria allows for an easy way to co-align replication and transcription, the results in human cells demonstrate that this can also be achieved in more complex environment where hundreds of origins are active.

Another process that might have contributed to shaping the landscape of bacterial chromosomes is the fusion of two converging replication forks. Work done by our lab as well as others has identified a surprising number of proteins that are involved in preventing over-replication in the termination area (Krabbe et al., 1997; Markovitz, 2005; Rudolph et al., 2013; Wendel et al., 2014, 2018; Midgley-Smith et al., 2018a,b), and we suggest that this is large number is needed for the processing of intermediates that arise directly as a result of forks fusions (Rudolph et al., 2009, 2010a; Dimude et al., 2016; Lloyd and Rudolph, 2016; **Figure 6**). Indeed, the lethality observed when multiple of these processing activities are removed from cells (Hong et al., 1995; Rudolph et al., 2010a; Zhang et al., 2010; Upton et al., 2014; Midgley-Smith et al., 2018b) highlights the importance of dealing with such intermediates. If the fusion of two forks can have harmful consequences, one easy way to limit these events is simply by reducing the number of origins. Having precisely one origin allows not only the easy co-orientation of

highly transcribed genes with DNA synthesis, but also reduces the number of fork fusion events to exactly one under normal conditions. Proteins and the replication fork trap allow then for the quick and efficient processing of potentially harmful fork intermediates (Dimude et al., 2016; Midgley-Smith et al., 2018a,b). The mild phenotypes of cells lacking a fork trap suggests that the various proteins involved can deal with fork fusion intermediates quite efficiently. Thus, acquiring the fork trap from a plasmid (Galli et al., 2019) might have been a welcome additional help to deal with these events, but it is not essential, explaining perhaps why many other bacterial species do not utilize a fork trap mechanism.

This hypothesis might help to explain why a transition from strictly single to both single and multiple origins took place in archaea. In both archaea and eukaryotic cells, the replicative helicase has the opposite polarity to the replicative helicase in bacteria (Tuteja and Tuteja, 2004; Costa and Onesti, 2008, 2009; Sakakibara et al., 2009) and encircles the single stranded leading strand template (Bai et al., 2017). Okazaki fragments in eukaryotes are much shorter than in prokaryotes (Burgers, 2009), allowing the replicative helicase to simply unwind either one or perhaps even more un-ligated Okazaki fragments. But even if any strand displacement would occur upon the merging of two forks, this would result in the generation of a 5′ flap which would be processed by a 5′ nuclease, such as the flap endo nuclease FEN-1 (Liu et al., 2004; Balakrishnan and Bambara, 2013). Thus, if the difference in the polarity of the replicative helicase alleviates potentially serious problems that arise as a result of fork fusions, it might at least in part explain the difference in origin dosage.

This does not mean that the fusion of forks is unproblematic in eukaryotic cells. On the contrary, recent work has highlighted that replisome disassembly is highly choreographed and that multiple accessory proteins, such as the helicases Rrm3 and Pif1, are necessary for bringing replication to an accurate conclusion (Steinacher et al., 2012; Maric et al., 2014; Moreno et al., 2014; Dewar et al., 2015; Moreno and Gambus, 2015; Dewar and Walter, 2017; Gambus, 2017; Deegan et al., 2019), highlighting that we have only just started to understand the mechanisms and regulation of fork fusion events both in bacteria and eukaryotes.

AUTHOR CONTRIBUTIONS

AS, JD, OS, and CR contributed to writing the manuscript.

FUNDING

The work in the Rudolph lab was supported by Research Grants BB/K015729/1 and BB/N014995/1 from the Biotechnology and Biological Sciences Research Council to CR.

ACKNOWLEDGMENTS

The authors would like to thank Steve Busby for his critical reading of the manuscript and to the referees for their constructive comments.

REFERENCES

Alexander, J. L., and Orr-Weaver, T. L. (2016). Replication fork instability and the consequences of fork collisions from rereplication. *Genes Dev.* 30, 2241–2252. doi: 10.1101/gad.288142.116

Atkinson, J., Gupta, M. K., Rudolph, C. J., Bell, H., Lloyd, R. G., and McGlynn, P. (2011). Localization of an accessory helicase at the replisome is critical in sustaining efficient genome duplication. *Nucleic Acids Res.* 39, 949–957. doi: 10.1093/nar/gkq889

Baharoglu, Z., Lestini, R., Duigou, S., and Michel, B. (2010). RNA polymerase mutations that facilitate replication progression in the rep uvrD recF mutant lacking two accessory replicative helicases. *Mol. Microbiol.* 77, 324–336. doi: 10.1111/j.1365-2958.2010.07208.x

Bai, L., Yuan, Z., Sun, J., Georgescu, R., O'Donnell, M. E., and Li, H. (2017). Architecture of the *Saccharomyces cerevisiae* Replisome. *Adv. Exp. Med. Biol.* 1042, 207–228. doi: 10.1007/978-981-10-6955-0_10

Balakrishnan, L., and Bambara, R. A. (2013). Flap endonuclease 1. *Annu. Rev. Biochem.* 82, 119–138. doi: 10.1146/annurev-biochem-072511-122603

Barre, F.-X. (2007). FtsK and SpoIIIE: the tale of the conserved tails. *Mol. Microbiol.* 66, 1051–1055. doi: 10.1111/j.1365-2958.2007.05981.x

Bhagwat, A. S., Hao, W., Townes, J. P., Lee, H., Tang, H., and Foster, P. L. (2016). Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2176–2181. doi: 10.1073/pnas.1522325113

Bianco, P. R. (2015). I came to a fork in the DNA and there was RecG. *Prog. Biophys. Mol. Biol.* 117, 166–173. doi: 10.1016/j.pbiomolbio.2015.01.001

Bigot (2005). KOPS: DNA motifs that control *E-coli* chromosome segregation by orienting the FtsK translocase. *EMBO J.* 24, 3770–3780. doi: 10.1038/sj.emboj.7600835

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462. doi: 10.1126/science.277.5331.1453

Blow, J. J., Ge, X. Q., and Jackson, D. A. (2011). How dormant origins promote complete genome replication. *Trends Biochem. Sci.* 36, 405–414. doi: 10.1016/j.tibs.2011.05.002

Blow, J. J., and Gillespie, P. J. (2008). Replication licensing and cancer–a fatal entanglement? *Nat. Rev. Cancer* 8, 799–806. doi: 10.1038/nrc2500

Bochman, M. L., Sabouri, N., and Zakian, V. A. (2010). Unwinding the functions of the Pif1 family helicases. *DNA Repair* 9, 237–249. doi: 10.1016/j.dnarep.2010.01.008

Boubakri, H., de Septenville, A. L., Viguera, E., and Michel, B. (2010). The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo. *EMBO J.* 29, 145–157. doi: 10.1038/emboj.2009.308

Bouché, J. P., Gélugne, J. P., Louarn, J., Louarn, J. M., and Kaiser, K. (1982). Relationships between the physical and genetic maps of a 470 x 10(3) base-pair region around the terminus of *Escherichia coli* K12 DNA replication. *J. Mol. Biol.* 154, 21–32. doi: 10.1016/0022-2836(82)90414-4

Boye, E., Løbner-Olesen, A., and Skarstad, K. (2000). Limiting DNA replication to once and only once. *EMBO Rep.* 1, 479–483. doi: 10.1093/embo-reports/kvd116

Brewer, B. J. (1988). When polymerases collide: replication and the transcriptional organization of the *E. coli* chromosome. *Cell* 53, 679–686. doi: 10.1016/0092-8674(88)90086-4

Brochu, J., Vlachos-Breton, É., Sutherland, S., Martel, M., and Drolet, M. (2018). Topoisomerases I and III inhibit R-loop formation to prevent unregulated replication in the chromosomal Ter region of *Escherichia coli*. *PLoS Genet.* 14:e1007668. doi: 10.1371/journal.pgen.1007668

Brüning, J.-G., Howard, J. L., and McGlynn, P. (2014). Accessory replicative helicases and the replication of protein-bound DNA. *J. Mol. Biol.* 426, 3917–3928. doi: 10.1016/j.jmb.2014.10.001

Bryant, J. A., Sellars, L. E., Busby, S. J. W., and Lee, D. J. (2014). Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res.* 42, 11383–11392. doi: 10.1093/nar/gku828

Burgers, P. M. J. (2009). Polymerase dynamics at the eukaryotic DNA replication fork. *J. Biol. Chem.* 284, 4041–4045. doi: 10.1074/jbc.R800062200

Chen, W.-H., Lu, G., Bork, P., Hu, S., and Lercher, M. J. (2016). Energy efficiency trade-offs drive nucleotide usage in transcribed regions. *Nat. Commun.* 7:11334. doi: 10.1038/ncomms11334

Chen, Y.-H., Keegan, S., Kahli, M., Tonzi, P., Fenyö, D., Huang, T. T., et al. (2019). Transcription shapes DNA replication initiation and termination in human cells. *Nat. Struct. Mol. Biol.* 26, 67–77. doi: 10.1038/s41594-018-0171-0

Condon, C., French, S., Squires, C., and Squires, C. L. (1993). Depletion of functional ribosomal RNA operons in *Escherichia coli* causes increased expression of the remaining intact copies. *EMBO J.* 12, 4305–4315. doi: 10.1002/j.1460-2075.1993.tb06115.x

Costa, A., Hood, I. V., and Berger, J. M. (2013). Mechanisms for Initiating Cellular DNA Replication. *Annu. Rev. Biochem.* 82, 25–54. doi: 10.1146/annurev-biochem-052610-094414

Costa, A., and Onesti, S. (2008). The MCM complex: (just) a replicative helicase? *Biochem. Soc. Trans.* 36, 136–140. doi: 10.1042/BST0360136

Costa, A., and Onesti, S. (2009). Structural biology of MCM helicases. *Crit. Rev. Biochem. Mol. Biol.* 44, 326–342. doi: 10.1080/10409230903186012

Courtot, L., Hoffmann, J.-S., and Bergoglio, V. (2018). The protective role of dormant origins in response to replicative stress. *Int. J. Mol. Sci.* 19:E3569. doi: 10.3390/ijms19113569

Cox, M. M. (2001). Recombinational DNA repair of damaged replication forks in *Escherichia coli*: questions. *Annu. Rev. Genet.* 35, 53–82. doi: 10.1146/annurev.genet.35.102401.090016

Cui, T., Moro-oka, N., Ohsumi, K., Kodama, K., Ohshima, T., Ogasawara, N., et al. (2007). *Escherichia coli* with a linear genome. *EMBO Rep.* 8, 181–187. doi: 10.1038/sj.embor.7400880

de Massy, B., Béjar, S., Louarn, J., Louarn, J. M., and Bouché, J. P. (1987). Inhibition of replication forks exiting the terminus region of the *Escherichia coli* chromosome occurs at two loci separated by 5 min. *Proc. Natl. Acad. Sci. U.S.A.* 84, 1759–1763. doi: 10.1073/pnas.84.7.1759

de Massy, B., Fayet, O., and Kogoma, T. (1984). Multiple origin usage for DNA replication in sdrA(rnh) mutants of Escherichia coli K-12. Initiation in the absence of *oriC*. *J. Mol. Biol.* 178, 227–236. doi: 10.1016/0022-2836(84)90141-4

De Septenville, A. L., Duigou, S., Boubakri, H., and Michel, B. (2012). Replication fork reversal after replication-transcription collision. *PLoS Genet.* 8:e1002622. doi: 10.1371/journal.pgen.1002622

Deegan, T. D., Baxter, J., Ortiz Bazán, M. Á., Yeeles, J. T. P., and Labib, K. P. M. (2019). Pif1-family helicases support fork convergence during DNA replication termination in eukaryotes. *Mol. Cell* 74, 231–244.e9. doi: 10.1016/j.molcel.2019.01.040

Dennis, P. P., Ehrenberg, M., Fange, D., and Bremer, H. (2009). Varying Rate of RNA Chain Elongation during rrn Transcription in *Escherichia coli*. *J. Bacteriol.* 191, 3740–3746. doi: 10.1128/JB.00128-09

Dewachter, L., Verstraeten, N., Fauvart, M., and Michiels, J. (2018). An integrative view of cell cycle control in *Escherichia coli*. *FEMS Microbiol. Rev.* 42, 116–136. doi: 10.1093/femsre/fuy005

Dewar, J. M., Budzowska, M., and Walter, J. C. (2015). The mechanism of DNA replication termination in vertebrates. *Nature* 525, 345–350. doi: 10.1038/nature14887

Dewar, J. M., and Walter, J. C. (2017). Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* 18, 507–516. doi: 10.1038/nrm.2017.42

Dillingham, M. S., and Kowalczykowski, S. C. (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.* 72, 642–671. doi: 10.1128/MMBR.00020-08

Dimude, J. U., Midgley-Smith, S. L., and Rudolph, C. J. (2018a). Replication-transcription conflicts trigger extensive DNA degradation in *Escherichia coli* cells lacking RecBCD. *DNA Repair* 70, 37–48. doi: 10.1016/j.dnarep.2018.08.002

Dimude, J. U., Midgley-Smith, S. L., Stein, M., and Rudolph, C. J. (2016). Replication termination: containing fork fusion-mediated pathologies in *Escherichia coli*. *Genes* 7:40. doi: 10.3390/genes7080040

Dimude, J. U., Stein, M., Andrzejewska, E. E., Khalifa, M. S., Gajdosova, A., Retkute, R., et al. (2018b). Origins left, right, and centre: increasing the number of initiation sites in the *Escherichia coli* chromosome. *Genes* 9:376. doi: 10.3390/genes9080376

Dimude, J. U., Stockum, A., Midgley-Smith, S. L., Upton, A. L., Foster, H. A., Khan, A., et al. (2015). The consequences of replicating in the wrong orientation: bacterial chromosome duplication without an active replication origin. *mBio* 6:e01294-15. doi: 10.1128/mBio.01294-15

Duggin, I. G., and Bell, S. D. (2009). Termination structures in the *Escherichia coli* chromosome replication fork trap. *J. Mol. Biol.* 387, 532–539. doi: 10.1016/j.jmb.2009.02.027

Duggin, I. G., Dubarry, N., and Bell, S. D. (2011). Replication termination and chromosome dimer resolution in the archaeon Sulfolobus solfataricus. *EMBO J.* 30, 145–153. doi: 10.1038/emboj.2010.301

Duggin, I. G., Wake, R. G., Bell, S. D., and Hill, T. M. (2008). The replication fork trap and termination of chromosome replication. *Mol. Microbiol.* 70, 1323–1333. doi: 10.1111/j.1365-2958.2008.06500.x

Duigou, S., and Boccard, F. (2017). Long range chromosome organization in *Escherichia coli*: the position of the replication origin defines the non-structured regions and the Right and Left macrodomains. *PLoS Genet.* 13:e1006758. doi: 10.1371/journal.pgen.1006758

Evertts, A. G., and Coller, H. A. (2012). Back to the origin: reconsidering replication, transcription, epigenetics, and cell cycle control. *Genes Cancer* 3, 678–696. doi: 10.1177/1947601912474891

Finkel, T., Serrano, M., and Blasco, M. A. (2007). The common biology of cancer and ageing. *Nature* 448, 767–774. doi: 10.1038/nature05985

Francino, M. P., Chao, L., Riley, M. A., and Ochman, H. (1996). Asymmetries generated by transcription-coupled repair in enterobacterial genes. *Science* 272, 107–109. doi: 10.1126/science.272.5258.107

French, S. (1992). Consequences of replication fork movement through transcription units *in vivo*. *Science* 258, 1362–1365. doi: 10.1126/science.1455232

Gabbai, C. B., and Marians, K. J. (2010). Recruitment to stalled replication forks of the PriA DNA helicase and replisome-loading activities is essential for survival. *DNA Repair* 9, 202–209. doi: 10.1016/j.dnarep.2009.12.009

Galli, E., Ferat, J.-L., Desfontaines, J.-M., Val, M.-E., Skovgaard, O., Barre, F.-X., et al. (2019). Replication termination without a replication fork trap. *Sci. Rep.* 9:8315. doi: 10.1038/s41598-019-43795-2

Gambus, A. (2017). Termination of eukaryotic replication forks. *Adv. Exp. Med. Biol.* 1042, 163–187. doi: 10.1007/978-981-10-6955-0_8

Gao, F. (2015). Bacteria may have multiple replication origins. *Front. Microbiol.* 6:324. doi: 10.3389/fmicb.2015.00324

Gao, F., and Zhang, C.-T. (2008). Ori-Finder: a web-based system for finding oriC s in unannotated bacterial genomes. *BMC Bioinformatics* 9:79. doi: 10.1186/1471-2105-9-79

Gerganova, V., Berger, M., Zaldastanishvili, E., Sobetzko, P., Lafon, C., Mourez, M., et al. (2015). Chromosomal position shift of a regulatory gene alters the bacterial phenotype. *Nucleic Acids Res.* 43, 8215–8226. doi: 10.1093/nar/gkv709

Gnanadurai, R., and Fifer, H. (2020). *Mycoplasma genitalium*: a review. *Microbiology* 166, 21–29. doi: 10.1099/mic.0.000830

Griffiths, A. A., and Wake, R. G. (2000). Utilization of subsidiary chromosomal replication terminators in *Bacillus subtilis*. *J. Bacteriol.* 182, 1448–1451. doi: 10.1128/jb.182.5.1448-1451.2000

Guy, C. P., Atkinson, J., Gupta, M. K., Mahdi, A. A., Gwynn, E. J., Rudolph, C. J., et al. (2009). Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol. Cell* 36, 654–666. doi: 10.1016/j.molcel.2009.11.009

Hansen, M. T. (1978). Multiplicity of genome equivalents in the radiation-resistant bacterium *Micrococcus radiodurans*. *J. Bacteriol.* 134, 71–75. doi: 10.1128/jb.134.1.71-75.1978

Harinarayanan, R., and Gowrishankar, J. (2003). Host factor titration by chromosomal R-loops as a mechanism for runaway plasmid replication in transcription termination-defective mutants of *Escherichia coli*. *J. Mol. Biol.* 332, 31–46. doi: 10.1016/s0022-2836(03)00753-8

Hasebe, T., Narita, K., Hidaka, S., and Su'etsugu, M. (2018). Efficient arrangement of the replication fork trap for *in vitro* propagation of monomeric circular DNA in the chromosome-replication cycle reaction. *Life* 8:43. doi: 10.3390/life8040043

Hawkins, M., Dimude, J. U., Howard, J. A. L., Smith, A. J., Dillingham, M. S., Savery, N. J., et al. (2019). Direct removal of RNA polymerase barriers to replication by accessory replicative helicases. *Nucleic Acids Res.* 47, 5100–5113. doi: 10.1093/nar/gkz170

Hawkins, M., Malla, S., Blythe, M. J., Nieduszynski, C. A., and Allers, T. (2013). Accelerated growth in the absence of DNA replication origins. *Nature* 503, 544–547. doi: 10.1038/nature12650

Hiasa, H., and Marians, K. J. (1994). Tus prevents overreplication of oriC plasmid DNA. *J. Biol. Chem.* 269, 26959–26968.

Hill, T. M., Henson, J. M., and Kuempel, P. L. (1987). The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc. Natl. Acad. Sci. U.S.A.* 84, 1754–1758. doi: 10.1073/pnas.84.7.1754

Hong, X., Cadwell, G. W., and Kogoma, T. (1995). *Escherichia coli* RecG and RecA proteins in R-loop formation. *EMBO J.* 14, 2385–2392. doi: 10.1002/j.1460-2075.1995.tb07233.x

Horiuchi, T., Fujimura, Y., Nishitani, H., Kobayashi, T., and Hidaka, M. (1994). The DNA replication fork blocked at the Ter site may be an entrance for the RecBCD enzyme into duplex DNA. *J. Bacteriol.* 176, 4656–4663. doi: 10.1128/jb.176.15.4656-4663.1994

Hyrien, O. (2000). Mechanisms and consequences of replication fork arrest. *Biochimie* 82, 5–17. doi: 10.1016/s0300-9084(00)00344-8

Hyrien, O. (2015). Peaks cloaked in the mist: The landscape of mammalian replication origins. *J. Cell Biol.* 208, 147–160. doi: 10.1083/jcb.201407004

Iismaa, T. P., and Wake, R. G. (1987). The normal replication terminus of the *Bacillus subtilis* chromosome, terC, is dispensable for vegetative growth and sporulation. *J. Mol. Biol.* 195, 299–310. doi: 10.1016/0022-2836(87)90651-6

Itaya, M., and Crouch, R. J. (1991). A combination of RNase H (rnh) and recBCD or sbcB mutations in *Escherichia coli* K12 adversely affects growth. *Mol. Gen. Genet.* 227, 424–432. doi: 10.1007/bf00273933

Ivanova, D., Taylor, T., Smith, S. L., Dimude, J. U., Upton, A. L., Mehrjouy, M. M., et al. (2015). Shaping the landscape of the *Escherichia coli* chromosome: replication-transcription encounters in cells with an ectopic replication origin. *Nucleic Acids Res.* 43, 7865–7877. doi: 10.1093/nar/gkv704

Jin, D. J., Cagliero, C., and Zhou, Y. N. (2012). Growth rate regulation in *Escherichia coli*. *FEMS Microbiol. Rev.* 36, 269–287. doi: 10.1111/j.1574-6976.2011.00279.x

Kim, N., and Jinks-Robertson, S. (2012). Transcription as a source of genome instability. *Nat. Rev. Genet.* 13, 204–214. doi: 10.1038/nrg3152

Kirby, R. (2011). Chromosome diversity and similarity within the Actinomycetales. *FEMS Microbiol. Lett.* 319, 1–10. doi: 10.1111/j.1574-6968.2011.02242.x

Kogoma, T. (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol. Mol. Biol. Rev.* 61, 212–238. doi: 10.1128/.61.2.212-238.1997

Kogoma, T., and von Meyenburg, K. (1983). The origin of replication, oriC, and the dnaA protein are dispensable in stable DNA replication (sdrA) mutants of *Escherichia coli* K-12. *EMBO J.* 2, 463–468. doi: 10.1002/j.1460-2075.1983.tb01445.x

Kouzminova, E. A., and Kuzminov, A. (2008). Patterns of chromosomal fragmentation due to uracil-DNA incorporation reveal a novel mechanism of replication-dependent double-stranded breaks. *Mol. Microbiol.* 68, 202–215. doi: 10.1111/j.1365-2958.2008.06149.x

Krabbe, M., Zabielski, J., Bernander, R., and Nordström, K. (1997). Inactivation of the replication-termination system affects the replication mode and causes unstable maintenance of plasmid R1. *Mol. Microbiol.* 24, 723–735. doi: 10.1046/j.1365-2958.1997.3791747.x

Kurth, I., and O'Donnell, M. (2009). Replisome Dynamics during Chromosome Duplication. *EcoSal Plus* 3:10.1128/ecosallus.4.4.2. doi: 10.1128/ecosalplus.4.4.2

Kuzminov, A. (2016). Chromosomal replication complexity: a novel DNA metrics and genome instability factor. *PLoS Genet.* 12:e1006229. doi: 10.1371/journal.pgen.1006229

Lane, H. E., and Denhardt, D. T. (1975). The rep mutation. IV. Slower movement of replication forks in Escherichia coli rep strains. *J. Mol. Biol.* 97, 99–112.

Lang, K. S., and Merrikh, H. (2018). The clash of macromolecular titans: replication-transcription conflicts in bacteria. *Annu. Rev. Microbiol.* 72, 71–88. doi: 10.1146/annurev-micro-090817-062514

Leela, J. K., Syeda, A. H., Anupama, K., and Gowrishankar, J. (2013). Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 258–263. doi: 10.1073/pnas.1213123110

Lemon, K. P., Kurtser, I., and Grossman, A. D. (2001). Effects of replication termination mutants on chromosome partitioning in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 212–217. doi: 10.1073/pnas.011506098

Leonard, A. C., and Méchali, M. (2013). DNA replication origins. *Cold Spring Harb. Perspect. Biol.* 5:a010116. doi: 10.1101/cshperspect.a010116

Lesterlin, C., Barre, F.-X., and Cornet, F. (2004). Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol. Microbiol.* 54, 1151–1160. doi: 10.1111/j.1365-2958.2004.04356.x

Levy, O., Ptacin, J. L., Pease, P. J., Gore, J., Eisen, M. B., Bustamante, C., et al. (2005). Identification of oligonucleotide sequences that direct the movement of the *Escherichia coli* FtsK translocase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17618–17623. doi: 10.1073/pnas.0508932102

Liu, X., Wang, X., Reyes-Lamothe, R., and Sherratt, D. (2010). Replication-directed sister chromosome alignment in *Escherichia coli*. *Mol. Microbiol.* 75, 1090–1097. doi: 10.1111/j.1365-2958.2009.06791.x

Liu, Y., Kao, H.-I., and Bambara, R. A. (2004). Flap endonuclease 1: a central component of DNA metabolism. *Annu. Rev. Biochem.* 73, 589–615. doi: 10.1146/annurev.biochem.73.012803.092453

Lloyd, R. G., and Rudolph, C. J. (2016). 25 years on and no end in sight: a perspective on the role of RecG protein. *Curr. Genet.* 62, 827–840. doi: 10.1007/s00294-016-0589-z

Lobry, J. R. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* 13, 660–665. doi: 10.1093/oxfordjournals.molbev.a025626

Lobry, J. R., and Louarn, J.-M. (2003). Polarisation of prokaryotic chromosomes. *Curr. Opin. Microbiol.* 6, 101–108. doi: 10.1016/S1369-5274(03)00024-9

Lundgren, M., Andersson, A., Chen, L., Nilsson, P., and Bernander, R. (2004). Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7046–7051. doi: 10.1073/pnas.0400656101

Maduike, N. Z., Tehranchi, A. K., Wang, J. D., and Kreuzer, K. N. (2014). Replication of the *Escherichia coli* chromosome in RNase HI-deficient cells: multiple initiation regions and fork dynamics. *Mol. Microbiol.* 91, 39–56. doi: 10.1111/mmi.12440

Maric, M., Maculins, T., De Piccoli, G., and Labib, K. (2014). Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. *Science* 346:1253596. doi: 10.1126/science.1253596

Markovitz, A. (2005). A new in vivo termination function for DNA polymerase I of *Escherichia coli* K12. *Mol. Microbiol.* 55, 1867–1882. doi: 10.1111/j.1365-2958.2005.04513.x

Masters, M., and Broda, P. (1971). Evidence for the bidirectional replications of the *Escherichia coli* chromosome. *Nat. New Biol.* 232, 137–140. doi: 10.1038/newbio232137a0

McCutcheon, J. P., and Moran, N. A. (2011). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10, 13–26. doi: 10.1038/nrmicro2670

McGlynn, P., and Lloyd, R. G. (2001). Rescue of stalled replication forks by RecG: Simultaneous translocation on the leading and lagging strand templates supports an active DNA unwinding model of fork reversal and Holliday junction formation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8227–8234. doi: 10.1073/pnas.111008698

McGlynn, P., Savery, N. J., and Dillingham, M. S. (2012). The conflict between DNA replication and transcription. *Mol. Microbiol.* 85, 12–20. doi: 10.1111/j.1365-2958.2012.08102.x

McLean, M. J., Wolfe, K. H., and Devine, K. M. (1998). Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. *J. Mol. Evol.* 47, 691–696. doi: 10.1007/pl00006428

Méchali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat. Rev. Mol. Cell Biol.* 11, 728–738. doi: 10.1038/nrm2976

Mendell, J. E., Clements, K. D., Choat, J. H., and Angert, E. R. (2008). Extreme polyploidy in a large bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6730–6734. doi: 10.1073/pnas.0707522105

Merrikh, H., Machón, C., Grainger, W. H., Grossman, A. D., and Soultanas, P. (2011). Co-directional replication-transcription conflicts lead to replication restart. *Nature* 470, 554–557. doi: 10.1038/nature09758

Merrikh, H., Zhang, Y., Grossman, A. D., and Wang, J. D. (2012). Replication-transcription conflicts in bacteria. *Nat. Rev. Microbiol.* 10, 449–458. doi: 10.1038/nrmicro2800

Michel, B., and Bernander, R. (2014). Chromosome replication origins: do we really need them? *BioEssays* 36, 585–590. doi: 10.1002/bies.201400003

Michel, B., Sinha, A. K., and Leach, D. R. F. (2018). Replication fork breakage and restart in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 82:e00013-18. doi: 10.1128/MMBR.00013-18

Midgley-Smith, S. L., Dimude, J. U., and Rudolph, C. J. (2018a). A role for 3′ exonucleases at the final stages of chromosome duplication in *Escherichia coli*. *Nucleic Acids Res.* 47, 1847–1860. doi: 10.1093/nar/gky1253

Midgley-Smith, S. L., Dimude, J. U., Taylor, T., Forrester, N. M., Upton, A. L., Lloyd, R. G., et al. (2018b). Chromosomal over-replication in *Escherichia coli* recG cells is triggered by replication fork fusion and amplified if replichore symmetry is disturbed. *Nucleic Acids Res.* 46, 7701–7715. doi: 10.1093/nar/gky566

Milbredt, S., Farmani, N., Sobetzko, P., and Waldminghaus, T. (2016). DNA replication in engineered *Escherichia coli* genomes with extra replication origins. *ACS Synth. Biol.* 5, 1167–1176. doi: 10.1021/acssynbio.6b00064

Million-Weaver, S., Samadpour, A. N., and Merrikh, H. (2015). Replication Restart after Replication-Transcription Conflicts Requires RecA in *Bacillus subtilis*. *J. Bacteriol.* 197, 2374–2382. doi: 10.1128/JB.00237-15

Minton, K. W., and Daly, M. J. (1995). A model for repair of radiation-induced DNA double-strand breaks in the extreme radiophile *Deinococcus radiodurans*. *BioEssays* 17, 457–464. doi: 10.1002/bies.950170514

Mirkin, E. V., and Mirkin, S. M. (2007). Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.* 71, 13–35. doi: 10.1128/MMBR.00030-06

Moreno, S. P., Bailey, R., Campion, N., Herron, S., and Gambus, A. (2014). Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science* 346, 477–481. doi: 10.1126/science.1253585

Moreno, S. P., and Gambus, A. (2015). Regulation of Unperturbed DNA Replication by Ubiquitylation. *Genes* 6, 451–468. doi: 10.3390/genes6030451

Mullakhanbhai, M. F., and Larsen, H. (1975). *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch. Microbiol.* 104, 207–214. doi: 10.1007/bf00447326

Müller, C. A., Hawkins, M., Retkute, R., Malla, S., Wilson, R., Blythe, M. J., et al. (2014). The dynamics of genome replication using deep sequencing. *Nucleic Acids Res.* 42:e3. doi: 10.1093/nar/gkt878

Nagpal, P., Jafri, S., Reddy, M. A., and Das, H. K. (1989). Multiple chromosomes of *Azotobacter vinelandii*. *J. Bacteriol.* 171, 3133–3138. doi: 10.1128/jb.171.6.3133-3138.1989

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., et al. (2006). The 160-kilobase genome of the bacterial endosymbiont Carsonella. *Science* 314:267. doi: 10.1126/science.1134196

Necsulea, A., Guillet, C., Cadoret, J.-C., Prioleau, M.-N., and Duret, L. (2009). The relationship between DNA replication and human genome organization. *Mol. Biol. Evol.* 26, 729–741. doi: 10.1093/molbev/msn303

Nester, E. W. (2015). Agrobacterium: nature's genetic engineer. *Front. Plant Sci.* 5:730. doi: 10.3389/fpls.2014.00730

Neylon, C., Kralicek, A. V., Hill, T. M., and Dixon, N. E. (2005). Replication termination in *Escherichia coli*: structure and antihelicase activity of the Tus-Ter complex. *Microbiol. Mol. Biol. Rev.* 69, 501–526. doi: 10.1128/MMBR.69.3.501-526.2005

Nordström, K. (2006). Plasmid R1–replication and its control. *Plasmid* 55, 1–26. doi: 10.1016/j.plasmid.2005.07.002

Pham, T. M., Tan, K. W., Sakumura, Y., Okumura, K., Maki, H., and Akiyama, M. T. (2013). A single-molecule approach to DNA replication in *Escherichia coli* cells demonstrated that DNA polymerase III is a major determinant of fork speed. *Mol. Microbiol.* 90, 584–596. doi: 10.1111/mmi.12386

Prescott, D. M., and Kuempel, P. L. (1972). Bidirectional replication of the chromosome in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 69, 2842–2845. doi: 10.1073/pnas.69.10.2842

Raghunathan, N., Goswami, S., Leela, J. K., Pandiyan, A., and Gowrishankar, J. (2019). A new role for *Escherichia coli* Dam DNA methylase in prevention of aberrant chromosomal replication. *Nucleic Acids Res.* 47, 5698–5711. doi: 10.1093/nar/gkz242

Reyes-Lamothe, R., Nicolas, E., and Sherratt, D. J. (2012). Chromosome replication and segregation in bacteria. *Annu. Rev. Genet.* 46, 121–143. doi: 10.1146/annurev-genet-110711-155421

Rocha, E. P. C., and Danchin, A. (2003). Gene essentiality determines chromosome organisation in bacteria. *Nucleic Acids Res.* 31, 6570–6577. doi: 10.1093/nar/gkg859

Rocha, E. P. C., Touchon, M., and Feil, E. J. (2006). Similar compositional biases are caused by very different mutational effects. *Genome Res.* 16, 1537–1547. doi: 10.1101/gr.5525106

Roecklein, B., Pelletier, A., and Kuempel, P. (1991). The tus gene of *Escherichia coli*: autoregulation, analysis of flanking sequences and identification of a complementary system in *Salmonella* typhimurium. *Res. Microbiol.* 142, 169–175. doi: 10.1016/0923-2508(91)90026-7

Rudolph, C. J., Dhillon, P., Moore, T., and Lloyd, R. G. (2007a). Avoiding and resolving conflicts between DNA replication and transcription. *DNA Repair* 6, 981–993. doi: 10.1016/j.dnarep.2007.02.017

Rudolph, C. J., Mahdi, A. A., Upton, A. L., and Lloyd, R. G. (2010a). RecG protein and single-strand DNA exonucleases avoid cell lethality associated with PriA helicase activity in *Escherichia coli*. *Genetics* 186, 473–492. doi: 10.1534/genetics.110.120691

Rudolph, C. J., Upton, A. L., Briggs, G. S., and Lloyd, R. G. (2010b). Is RecG a general guardian of the bacterial genome? *DNA Repair* 9, 210–223. doi: 10.1016/j.dnarep.2009.12.014

Rudolph, C. J., Upton, A. L., and Lloyd, R. G. (2007b). Replication fork stalling and cell cycle arrest in UV-irradiated *Escherichia coli*. *Genes Dev.* 21, 668–681. doi: 10.1101/gad.417607

Rudolph, C. J., Upton, A. L., and Lloyd, R. G. (2009). Replication fork collisions cause pathological chromosomal amplification in cells lacking RecG DNA translocase. *Mol. Microbiol.* 74, 940–955. doi: 10.1111/j.1365-2958.2009.06909.x

Rudolph, C. J., Upton, A. L., Stockum, A., Nieduszynski, C. A., and Lloyd, R. G. (2013). Avoiding chromosome pathology when replication forks collide. *Nature* 500, 608–611. doi: 10.1038/nature12312

Sakakibara, N., Kelman, L. M., and Kelman, Z. (2009). Unwinding the structure and function of the archaeal MCM helicase. *Mol. Microbiol.* 72, 286–296. doi: 10.1111/j.1365-2958.2009.06663.x

Samson, R. Y., Xu, Y., Gadelha, C., Stone, T. A., Faqiri, J. N., Li, D., et al. (2013). Specificity and function of archaeal DNA replication initiator proteins. *Cell Rep.* 3, 485–496. doi: 10.1016/j.celrep.2013.01.002

Schneiker, S., Perlova, O., Kaiser, O., Gerth, K., Alici, A., Altmeyer, M. O., et al. (2007). Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat. Biotechnol.* 25, 1281–1289. doi: 10.1038/nbt1354

Scholz, S. A., Diao, R., Wolfe, M. B., Fivenson, E. M., Lin, X. N., and Freddolino, P. L. (2019). High-resolution mapping of the *Escherichia coli* chromosome reveals positions of high and low transcription. *Cell Syst.* 8, 212–225.e9. doi: 10.1016/j.cels.2019.02.004

Sherratt, D. J., Arciszewska, L. K., Crozat, E., Graham, J. E., and Grainge, I. (2010). The *Escherichia coli* DNA translocase FtsK. *Biochem. Soc. Trans.* 38, 395–398. doi: 10.1042/BST0380395

Singleton, M. R., Dillingham, M. S., Gaudier, M., Kowalczykowski, S. C., and Wigley, D. B. (2004). Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* 432, 187–193. doi: 10.1038/nature02988

Sinha, A. K., Durand, A., Desfontaines, J.-M., Iurchenko, I., Auger, H., Leach, D. R. F., et al. (2017). Division-induced DNA double strand breaks in the chromosome terminus region of *Escherichia coli* lacking RecBCD DNA repair enzyme. *PLoS Genet.* 13:e1006895. doi: 10.1371/journal.pgen.1006895

Sinha, A. K., Possoz, C., Durand, A., Desfontaines, J.-M., Barre, F.-X., Leach, D. R. F., et al. (2018). Broken replication forks trigger heritable DNA breaks in the terminus of a circular chromosome. *PLoS Genet.* 14:e1007256. doi: 10.1371/journal.pgen.1007256

Skovgaard, O., Bak, M., Løbner-Olesen, A., and Tommerup, N. (2011). Genome-wide detection of chromosomal rearrangements, indels, and mutations in circular chromosomes by short read sequencing. *Genome Res.* 21, 1388–1393. doi: 10.1101/gr.117416.110

Smith, G. R. (2012). How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view. *Microbiol. Mol. Biol. Rev.* 76, 217–228. doi: 10.1128/MMBR.05026-11

Srivatsan, A., Tehranchi, A., MacAlpine, D. M., and Wang, J. D. (2010). Co-orientation of replication and transcription preserves genome integrity. *PLoS Genet.* 6:e1000810. doi: 10.1371/journal.pgen.1000810

Steinacher, R., Osman, F., Dalgaard, J. Z., Lorenz, A., and Whitby, M. C. (2012). The DNA helicase Pfh1 promotes fork merging at replication termination sites to ensure genome stability. *Genes Dev.* 26, 594–602. doi: 10.1101/gad.184663.111

Syeda, A. H., Atkinson, J., Lloyd, R. G., and McGlynn, P. (2016). The Balance between recombination enzymes and accessory replicative helicases in facilitating genome duplication. *Genes* 7:E42. doi: 10.3390/genes7080042

Syeda, A. H., Wollman, A. J. M., Hargreaves, A. L., Howard, J. A. L., Brüning, J.-G., McGlynn, P., et al. (2019). Single-molecule live cell imaging of Rep reveals the dynamic interplay between an accessory replicative helicase and the replisome. *Nucleic Acids Res.* 47, 6287–6298. doi: 10.1093/nar/gkz298

Tadokoro, T., and Kanaya, S. (2009). Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *FEBS J.* 276, 1482–1493. doi: 10.1111/j.1742-4658.2009.06907.x

Tamames, J. (2001). Evolution of gene order conservation in prokaryotes. *Genome Biol.* 2:RESEARCH0020.

Tanaka, T., and Masai, H. (2006). Stabilization of a stalled replication fork by concerted actions of two helicases. *J. Biol. Chem.* 281, 3484–3493. doi: 10.1074/jbc.M510979200

Taylor-Robinson, D., and Jensen, J. S. (2011). *Mycoplasma genitalium*: from Chrysalis to Multicolored Butterfly. *Clin. Microbiol. Rev.* 24, 498–514. doi: 10.1128/CMR.00006-11

Terakawa, T., Redding, S., Silverstein, T. D., and Greene, E. C. (2017). Sequential eviction of crowded nucleoprotein complexes by the exonuclease RecBCD molecular motor. *Proc. Natl. Acad. Sci. U.S.A.* 114, E6322–E6331. doi: 10.1073/pnas.1701368114

Timmins, J., and Moe, E. (2016). A Decade of Biochemical and Structural Studies of the DNA Repair Machinery of *Deinococcus radiodurans*: Major Findings, Functional and Mechanistic Insight and Challenges. *Comput. Struct. Biotechnol. J.* 14, 168–176. doi: 10.1016/j.csbj.2016.04.001

Tomasetti, C., Li, L., and Vogelstein, B. (2017). Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* 355, 1330–1334. doi: 10.1126/science.aaf9011

Touchon, M., and Rocha, E. P. C. (2016). Coevolution of the Organization and Structure of Prokaryotic Genomes. *Cold Spring Harb. Perspect. Biol.* 8:a018168. doi: 10.1101/cshperspect.a018168

Trautinger, B. W., Jaktaji, R. P., Rusakova, E., and Lloyd, R. G. (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol. Cell* 19, 247–258. doi: 10.1016/j.molcel.2005.06.004

Tuteja, N., and Tuteja, R. (2004). Prokaryotic and eukaryotic DNA helicases. Essential molecular motor proteins for cellular machinery. *Eur. J. Biochem.* 271, 1835–1848. doi: 10.1111/j.1432-1033.2004.04093.x

Umenhoffer, K., Draskovits, G., Nyerges, Á., Karcagi, I., Bogos, B., Tímár, E., et al. (2017). Genome-Wide Abolishment of Mobile Genetic Elements Using Genome Shuffling and CRISPR/Cas-Assisted MAGE Allows the Efficient Stabilization of a Bacterial Chassis. *ACS Synth. Biol.* 6, 1471–1483. doi: 10.1021/acssynbio.6b00378

Upton, A. L., Grove, J. I., Mahdi, A. A., Briggs, G. S., Milner, D. S., Rudolph, C. J., et al. (2014). Cellular location and activity of *Escherichia coli* RecG proteins shed light on the function of its structurally unresolved C-terminus. *Nucleic Acids Res.* 42, 5702–5714. doi: 10.1093/nar/gku228

Valens, M., Penaud, S., Rossignol, M., Cornet, F., and Boccard, F. (2004). Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J.* 23, 4330–4341. doi: 10.1038/sj.emboj.7600434

Veetil, R. T., Malhotra, N., Dubey, A., and Seshasayee, A. S. N. (2020). Laboratory evolution experiments help identify a predominant region of constitutive stable DNA Replication Initiation. *mSphere* 5:e00939-19. doi: 10.1128/mSphere.00939-19

Verma, S. C., Qian, Z., and Adhya, S. L. (2019). Architecture of the *Escherichia coli* nucleoid. *PLoS Genet.* 15:e1008456. doi: 10.1371/journal.pgen.1008456

Vogel, U., and Jensen, K. F. (1994). The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol.* 176, 2807–2813. doi: 10.1128/jb.176.10.2807-2813.1994

Voloshin, O. N., and Camerini-Otero, R. D. (2007). The DinG protein from *Escherichia coli* is a structure-specific helicase. *J. Biol. Chem.* 282, 18437–18447. doi: 10.1074/jbc.M700376200

Wang, J. D., Berkmen, M. B., and Grossman, A. D. (2007). Genome-wide coorientation of replication and transcription reduces adverse effects on replication in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5608–5613. doi: 10.1073/pnas.0608999104

Wang, X., Lesterlin, C., Reyes-Lamothe, R., Ball, G., and Sherratt, D. J. (2011). Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proc. Natl. Acad. Sci. U.S.A.* 108, E243–E250. doi: 10.1073/pnas.1100874108

Wendel, B. M., Cole, J. M., Courcelle, C. T., and Courcelle, J. (2018). SbcC-SbcD and ExoI process convergent forks to complete chromosome replication. *Proc. Natl. Acad. Sci. U.S.A.* 115, 349–354. doi: 10.1073/pnas.1715960114

Wendel, B. M., Courcelle, C. T., and Courcelle, J. (2014). Completion of DNA replication in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16454–16459. doi: 10.1073/pnas.1415025111

Wiktor, J., van der Does, M., Büller, L., Sherratt, D. J., and Dekker, C. (2018). Direct observation of end resection by RecBCD during double-stranded DNA break repair in vivo. *Nucleic Acids Res.* 46, 1821–1833. doi: 10.1093/nar/gkx1290

Windgassen, T. A., Wessel, S. R., Bhattacharyya, B., and Keck, J. L. (2018). Mechanisms of bacterial DNA replication restart. *Nucleic Acids Res.* 46, 504–519. doi: 10.1093/nar/gkx1203

Winterstein, C., and Ludwig, B. (1998). Genes coding for respiratory complexes map on all three chromosomes of the Paracoccus denitrificans genome. *Arch. Microbiol.* 169, 275–281. doi: 10.1007/s002030050572

Wu, C.-I., and Maeda, N. (1987). Inequality in mutation rates of the two strands of DNA. *Nature* 327, 169–170. doi: 10.1038/327169a0

Wu, Z., Liu, J., Yang, H., and Xiang, H. (2014). DNA replication origins in archaea. *Front. Microbiol.* 5:179. doi: 10.3389/fmicb.2014.00179

Zaritsky, A., and Woldringh, C. L. (2015). Chromosome replication, cell growth, division and shape: a personal perspective. *Front. Microbiol* 6:756. doi: 10.3389/fmicb.2015.00756

Zhang, J., Mahdi, A. A., Briggs, G. S., and Lloyd, R. G. (2010). Promoting and avoiding recombination: contrasting activities of the *Escherichia coli* RuvABC Holliday junction resolvase and RecG DNA translocase. *Genetics* 185, 23–37. doi: 10.1534/genetics.110.114413

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 April 2020
doi: 10.3389/fmicb.2020.00786

Evolutionary Changes in DnaA-Dependent Chromosomal Replication in Cyanobacteria

Ryudo Ohbayashi[1†], Shunsuke Hirooka[1], Ryo Onuma[1], Yu Kanesaki[2], Yuu Hirose[3], Yusuke Kobayashi[1†], Takayuki Fujiwara[1,4], Chikara Furusawa[5,6] and Shin-ya Miyagishima[1,4*]*

[1] Department of Gene Function and Phenomics, National Institute of Genetics, Shizuoka, Japan, [2] Research Institute of Green Science and Technology, Shizuoka University, Shizuoka, Japan, [3] Department of Applied Chemistry and Life Science, Toyohashi University of Technology, Toyohashi, Japan, [4] Department of Genetics, The Graduate University for Advanced Studies (SOKENDAI), Shizuoka, Japan, [5] Center for Biosystems Dynamics Research, RIKEN, Osaka, Japan, [6] Universal Biology Institute, Graduate School of Science, The University of Tokyo, Tokyo, Japan

OPEN ACCESS

Edited by:
Torsten Waldminghaus, Philipps University of Marburg, Germany

Reviewed by:
Ole Skovgaard, Roskilde University, Denmark
Annegret Wilde, University of Freiburg, Germany

***Correspondence:**
Ryudo Ohbayashi
ryudo.ohbayashi@riken.jp
Shin-ya Miyagishima
smiyagis@nig.ac.jp

†Present address:
Ryudo Ohbayashi, Center for Biosystems Dynamics Research, RIKEN, Tokyo, Japan
Yusuke Kobayashi, College of Science, Ibaraki University, Ibaraki, Japan

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: 17 January 2020
Accepted: 02 April 2020
Published: 28 April 2020

Citation:
Ohbayashi R, Hirooka S, Onuma R, Kanesaki Y, Hirose Y, Kobayashi Y, Fujiwara T, Furusawa C and Miyagishima S (2020) Evolutionary Changes in DnaA-Dependent Chromosomal Replication in Cyanobacteria. Front. Microbiol. 11:786. doi: 10.3389/fmicb.2020.00786

Replication of the circular bacterial chromosome is initiated at a unique origin (*oriC*) in a DnaA-dependent manner in which replication proceeds bidirectionally from *oriC* to *ter*. The nucleotide compositions of most bacteria differ between the leading and lagging DNA strands. Thus, the chromosomal DNA sequence typically exhibits an asymmetric GC skew profile. Further, free-living bacteria without genomes encoding *dnaA* were unknown. Thus, a DnaA-*oriC*-dependent replication initiation mechanism may be essential for most bacteria. However, most cyanobacterial genomes exhibit irregular GC skew profiles. We previously found that the *Synechococcus elongatus* chromosome, which exhibits a regular GC skew profile, is replicated in a DnaA-*oriC*-dependent manner, whereas chromosomes of *Synechocystis* sp. PCC 6803 and *Nostoc* sp. PCC 7120, which exhibit an irregular GC skew profile, are replicated from multiple origins in a DnaA-independent manner. Here we investigate the variation in the mechanisms of cyanobacterial chromosome replication. We found that the genomes of certain free-living species do not encode *dnaA* and such species, including *Cyanobacterium aponinum* PCC 10605 and *Geminocystis* sp. NIES-3708, replicate their chromosomes from multiple origins. *Synechococcus* sp. PCC 7002, which is phylogenetically closely related to *dnaA*-lacking free-living species as well as to *dnaA*-encoding but DnaA-*oriC*-independent *Synechocystis* sp. PCC 6803, possesses *dnaA*. In *Synechococcus* sp. PCC 7002, *dnaA* was not essential and its chromosomes were replicated from a unique origin in a DnaA-*oriC* independent manner. Our results also suggest that loss of DnaA-*oriC*-dependency independently occurred multiple times during cyanobacterial evolution and raises a possibility that the loss of *dnaA* or loss of DnaA-*oriC* dependency correlated with an increase in ploidy level.

Keywords: CDS skew, chromosome replication, cyanobacteria, DnaA, GC skew, polyploidy

INTRODUCTION

Precise chromosomal DNA replication is required for the inheritance of genetic information during cellular proliferation. Chromosome replication is tightly controlled, mainly at the initiation stage of DNA replication (Wang and Levin, 2009; Katayama et al., 2010; Skarstad and Katayama, 2013). In eukaryotes and certain archaea, chromosome replication is initiated at multiple origins scattered throughout chromosomes (Bell, 2017; Marks et al., 2017). In contrast, in most bacteria with a single circular chromosome, chromosome replication is initiated at a unique origin named *oriC* (Katayama et al., 2010; Skarstad and Katayama, 2013). Two replication forks assembled at *oriC* proceed bidirectionally around the circular chromosome, simultaneously synthesizing the nascent leading and lagging DNA strands. Replication of circular chromosomal DNA terminates in the *ter* region, located at a site opposite to that of *oriC* (Duggin et al., 2008; Beattie and Reyes-Lamothe, 2015; Dewar and Walter, 2017). *oriC* comprises several copies of the DnaA-box sequence (TTATNCACA) that is bound by DnaA. DnaA unwinds the duplex DNA to form single-stranded DNA templates. Subsequently, the replisome is recruited to the unwound DNA and then initiates DNA synthesis (Katayama et al., 2010). Free-living bacteria, which do not possess *dnaA*, were not reported, and evidence suggested that the DnaA-*oriC*-dependent mechanism of chromosomal DNA replication initiation is conserved among bacteria, except for certain symbionts and parasites, which do not harbor *dnaA* (Akman et al., 2002; Gil et al., 2003; Klasson and Andersson, 2004; Ran et al., 2010; Nakayama et al., 2014).

Nucleotide compositional bias and gene orientation bias between the leading and lagging DNA strands occur in most bacterial chromosomes (Lobry, 1996; Freeman et al., 1998; Bentley and Parkhill, 2004; Nikolaou and Almirantis, 2005; Necşulea and Lobry, 2007). GC skew, defined as (G – C)/(G+C), switches near *oriC* and *ter* (Lobry, 1996; Grigoriev, 1998). Further, coding-sequence orientation bias (CDS skew) switches near *oriC* and *ter* (Freeman et al., 1998; Nikolaou and Almirantis, 2005; Necşulea and Lobry, 2007), because numerous genes, particularly those abundantly expressed or those essential for viability, are encoded on the leading, rather than the lagging strand (Rocha and Danchin, 2003). Moreover, the replication-associated *dnaA-dnaN* operon resides near *oriC* (Mackiewicz et al., 2004). These conserved footprints on the bacterial chromosome and conservation of *dnaA* indicate that the DnaA-*oriC*-dependent mechanism for chromosome replication is highly conserved in bacteria.

The GC and CDS skews and location of the *dnaA-dnaN* operon are used to computationally predict the position of *oriC* in numerous bacterial chromosomal genomes (Luo and Gao, 2018). However, the availability of an increasing number of bacterial genomes, accelerated through the rapid development of nucleotide sequencing technologies, show that the replication origin of certain bacterial genomes cannot be predicted from the genomic sequence. Examples include cyanobacterial genomes with irregular GC and CDS skews (McLean et al., 1998; Nikolaou and Almirantis, 2005; Worning et al., 2006; Arakawa and Tomita, 2007). Further, *dnaA* and *dnaN* are encoded separately in these genomes (Zhou et al., 2011; Huang et al., 2015). In addition, most cyanobacterial species possess multiple copies of the same chromosomes (Simon, 1977; Binder and Chisholm, 1990, 1995; Hu et al., 2007; Zerulla et al., 2016), in contrast to those present in single copies in model bacteria such as *Escherichia coli*, *Bacillus subtilis*, and *Caulobacter crescentus*.

The replication origin was experimentally identified in the chromosome of the model cyanobacterium *Synechococcus elongat*us PCC 7942 (*S. elongatus*), which exhibits a regular GC skew (Watanabe et al., 2012). In this genome, *oriC* resides near *dnaN,* which corresponds to one of the two shift points of the GC skew (Watanabe et al., 2012) (**Figure 1**). Further, DnaA binds *oriC*, and *dnaA* is essential in *S. elongatus* (Ohbayashi et al., 2016). In contrast, the chromosomes of *Synechocystis* sp. PCC 6803 and *Nostoc* sp. PCC 7120 exhibit an irregular GC skew profile (**Figure 1**), and their chromosomes are replicated asynchronously from multiple origins (Ohbayashi et al., 2016). Deletion of *dnaA* does not affect the growth and chromosomal replication activity of these species (Ohbayashi et al., 2016), suggesting that *Synechocystis* sp. PCC 6803 and *Nostoc* sp. PCC 7120 may replicate their chromosomes through an unknown DnaA-*oriC*-independent mechanism (Ohbayashi et al., 2016).

DnaA binds DnaA-boxes in chromosomal regions other than *oriC* to regulate transcription (Messer and Weigel, 1997; Burkholder et al., 2001; Hottes et al., 2005; Ishikawa et al., 2007). Thus, DnaA protein in these cyanobacterial species likely involves the transcriptional regulation of specific genes, but likely is not involved in chromosome replication. Moreover, *dnaA* is not encoded by the genomes of symbiotic cyanobacterial species and chloroplasts, which evolved from a cyanobacterial endosymbiont (Ran et al., 2010; Nakayama et al., 2014; Ohbayashi et al., 2016).

Thus, several species of cyanobacteria, the majority of which possess multiple copies of the same chromosome (Griese et al., 2011; Sargent et al., 2016), likely have developed DnaA-*oriC*-independent mechanism of chromosome replication during evolution, and *dnaA* was subsequently lost from symbionts and chloroplasts. However, the dependence of chromosome replication on DnaA-*oriC* in cyanobacteria other than *S. elongatus, Synechocystis* sp. PCC 6803, and *Nostoc* sp. PCC 7120 and the evolutionary relationships among DnaA-*oriC*-dependent and independent species have not been examined.

To address this gap in our knowledge, here we conducted an up-to-date review of cyanobacterial genome sequences and found that certain free-living species do not encode *dnaA*. *Synechococcus* sp. PCC 7002 is closely related to free-living species without *dnaA* as well as to DnaA-*oriC*-independent *Synechocystis* sp. PCC 6803. We found that the chromosome of *Synechococcus* sp. PCC 7002 is replicated in a DnaA-*oriC* independent manner that occurs in *Synechocystis* sp. PCC 6803, although replication of the chromosome apparently initiated at a specific position. Our results further reveal variations in the mechanism of the initiation of chromosome replication associated with DnaA-*oriC*-dependence and suggest that mechanisms of DnaA-*oriC*-independent

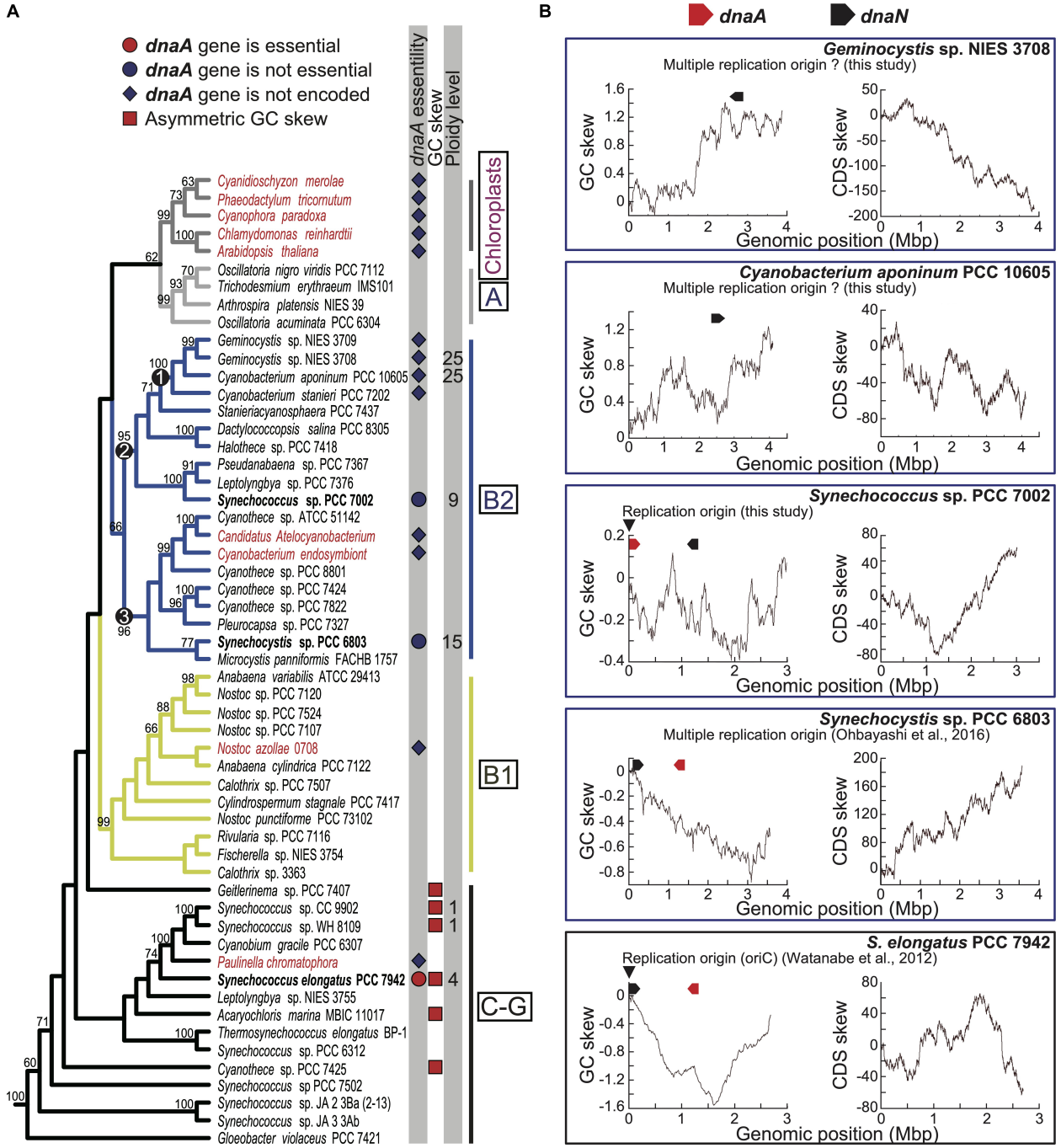


FIGURE 1 | Phylogenetic distributions of *dnaA* among cyanobacterial and chloroplast genomes and GC and CDS skew profiles of chromosomes. **(A)** Phylogenetic relationships of cyanobacteria and chloroplasts and presence or absence of genomic *dnaA*. The tree was constructed using a maximum likelihood method based on 58 concatenated rDNA sequences (16S + 23S + 5S rDNA). Bootstrap values > 50% are shown above the selected branches. The full tree with the outgroup and accession numbers of respective nucleotide sequences are indicated in **Supplementary Figure S1**. Clades A, B1, B2, and C-G are defined according to a previous study (Shih et al., 2013). Chloroplasts and symbiotic species are shown in red, and free-living species are shown in black. The red or blue circle next to the species name indicates that *dnaA* was experimentally shown as essential or not essential for chromosome replication in the species (Ohbayashi et al., 2016; present study). The blue diamond next to the species name indicates the absence of *dnaA*. The red square next to the species name indicates that the chromosomal genome exhibits a clear asymmetric (V-shaped) GC skew profile. **(B)** Cumulative GC and CDS skew profiles of the chromosome of *Geminocystis* sp. NIES-3708, *Cyanobacterium aponinum* PCC 10605, *Synechococcus* sp. PCC 7002, *Synechocystis* sp. PCC 6803, and *Synechococcus elongatus* (profiles of other species are shown in **Supplementary Figures S4**, **S5**). The positions of *dnaA* and *dnaN* are indicated above the profiles. The arrowheads indicate the experimentally determined replication origins of *Synechococcus* sp. PCC 7002 (this study) and *S. elongatus* (Watanabe et al., 2012). The chromosome of *Synechocystis* sp. PCC 6803 is replicated from multiple origins (Ohbayashi et al., 2016).

replication independently evolved multiple times during cyanobacterial evolution.

MATERIALS AND METHODS

Strains and Culture Condition

Synechococcus elongatus, *Synechocystis* sp. PCC 6803, *Cyanobacterium aponinum* PCC 10605, and *Geminocystis* sp. NIES-3708 were cultured in BG-11 liquid medium at 30°C with air bubbling in the light (70 μmol m^{-2} s^{-1} photons). *Synechococcus* sp. PCC 7002 was cultured in modified liquid medium A (Aikawa et al., 2014) with air bubbling at 38°C in the light (70 μmol m^{-2} s^{-1}), unless otherwise indicated.

Construction of GC and CDS Skew Profiles

GC and CDS skew analyses were performed using the G-language Genome Analysis Environment (Arakawa et al., 2003). The cumulative GC and CDS skews were calculated using the "gcskew" and "CDS skew" functions with cumulative parameters, respectively. Accession numbers of the genomic sequences of cyanobacterial species and chloroplasts are indicated in **Supplementary Figure S1**.

Phylogenetic Analysis

The concatenated rDNA sequences (16S + 23S + 5S rDNA) of 58 species were automatically aligned using the L-INS-I method of MAFFT v7.299b (Katoh and Standley, 2013). Poorly aligned regions were eliminated using TrimAl v1.2 (Capella-Gutiérrez et al., 2009) with the "-gappyout" option. The aligned sequences were calculated using RAxML v8.2.9 (Stamatakis, 2014) with GTR + GAMMA model, which was selected using Kakusan4 (Tanabe, 2011), and the corresponding bootstrap support values were calculated through ML analysis of 1,000 pseudoreplicates. Accession numbers of the genomic sequences of cyanobacterial species and chloroplasts are indicated in **Supplementary Figure S1**.

Chromatin Immunoprecipitation (ChIP) and Qualitative PCR (qPCR) Analyses

ChIP and subsequent qPCR analyses were performed according to Hanaoka and Tanaka (2008) with minor modifications. Cells were fixed with 1% formaldehyde for 15 min at room temperature. The cross-linking reaction was stopped by adding glycine (final concentration, 125 mM) and incubating at room temperature for 5 min. Fixed cells were centrifuged, washed twice with ice-cold Tris-buffered saline (TBS; 20 mM Tris-HCl, pH 7.4, 150 mM NaCl), and stored at −80°C. Fixed cells were disrupted using a Beads Crusher (TAITEC) with glass beads (<106 μm, Sigma-Aldrich) in TBS at 4°C, and the genome DNA was sonicated (Covaris Sonication System, MS Instruments) to generate to approximately 500-bp fragments. After centrifugation for 15 min to sediment the insoluble materials, the supernatant fraction containing genomic DNA was subjected to immunoprecipitation with an anti-HA antibody (clone 16B12, BioLegend), diluted 1:250. Precipitated DNAs were quantified using qPCR with the primer sets oriC-F and oriC-R representing the *oriC* region and 1294-F as well as with 1294-R representing the genomic region farthest from *oriC* in the circular genome. Primer sequences are listed in **Supplementary Table S1**.

Flow Cytometric Analysis of DNA and Fluorescence Microscopy of Nucleoids

Exponentially growing cells (**Supplementary Figure S2** and **Figure 4B**) were centrifuged, fixed with 1% glutaraldehyde for 10 min, and washed with PBS. Fixed cells were stained with SYBR Green I (1:3000) and then subjected to flow cytometry (BD, Accuri C6) and observed using fluorescence microscopy (Olympus, BX-52).

Immunoblotting

Total cellular proteins (80 μg of proteins per lane) were separated using a 10% acrylamide gel and then electrophoretically transferred to a PVDF membrane. Membranes were blocked with 5% skim milk in TBS-T (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20) and incubated with an anti-HA antibody (clone 16B12, BioLegend) diluted 1:1,000. A horseradish peroxidase–conjugated goat anti-mouse IgG (Thermo Fisher Scientific), diluted 1:20,000, served as the secondary antibody. The immune complexes were detected using ECL Prime Western Blotting Detection Reagent (GE Healthcare) with an Image Quant LAS 4000 Mini (GE Healthcare).

High-Throughput Nucleotide Sequence Analysis of Chromosome Replication

Genomic DNA was extracted from exponentially growing and stationary phase *S. elongatus*, *C. aponinum* PCC 10605, *Geminocystis* NIES-3708, *Synechococcus* sp. PCC 7002 as well as its *dnaA* disruptant (**Supplementary Figure S2**). A Covaris S2 Sonication System (Covaris, Inc., Woburn, MA, United States) was used to sonicate DNA samples (5 μg each) to generate approximately 500-bp fragments. Sequencing libraries were prepared using the Ultra II DNA Library Prep Kit for Illumina (New England Biolabs). Paired-end sequencing (320 cycles) was conducted using the MiSeq system (Illumina) according to the manufacturer's specifications. The sequence reads (number of reads in all samples was >100 per base) were trimmed using CLC Genomics Workbench ver. 8.5.1 (Qiagen) using the parameters as follows: Phred quality score >30, removal of the terminal 15 nucleotides from the 5′ end and 2 nucleotides from the 3′ end, and removal of truncated reads > 100 nucleotides. Trimmed reads were mapped to the reference genome sequences of the respective species using CLC Genomics Workbench ver. 9.5.1 (Qiagen) with the parameters as follows: Length fraction, 0.7 and Similarity fraction, 0.9. To call SNPs and indels, we used the filter settings as follows: minimum read depth for SNP/indel calling = 10, minimum read depth for SNP calling = 5, and 80% cutoff of percentage aligned-reads-calling the SNP per total mapped reads at the SNP sites. Alternatively, potential nucleotide differences were determined using BRESEQ (Deatherage and Barrick, 2014)

(**Supplementary Data File**). Paired-end reads were assembled *de novo* using Newbler version 2.9 (Roche).

Plasmid Construction and Generation of Stable Transformants

Synechococcus elongatus expressing HA-tagged DnaA under the control of the endogenous *dnaA* promoter were generated as follows: PCR was used to amplify *dnaA orf* with the primers 1 and 2; The pNSHA vector (possessing NS I sequences for double-crossover recombination, *trc* promoter, HA-coding sequence, and spectinomycin resistance gene) (Watanabe et al., 2012) was amplified as a linear DNA using PCR with primers 3 and 4. The amplified *dnaA orf* was subsequently inserted between the *trc* promoter and HA-coding sequence of pNSHA using an In-Fusion Cloning Kit (TAKARA). The *dnaA* promoter (300-bp 5′-upstream sequence flanking the *dnaA* translation start codon), which was amplified using PCR with the primers 5 and 6, was cloned immediately upstream of the HA-coding sequences of the vector and then amplified as a linear DNA using PCR with primers 7 and 8 using the In-Fusion Cloning Kit. The resultant vector was used to transform *S. elongatus*.

To generate an *S. elongatus* strain expressing HA-tagged DnaA of *Synechococcus* sp. PCC 7002 or *Synechocystis* sp, *dnaA orf* of *S. elongatus* in the above vector (expressing HA-tagged DnaA of *S. elongatus* driven by the *dnaA* promoter) was replaced with *dnaA orf* of the respective species as follows: The *dnaA orf* of each respective species was amplified using PCR with the primers 9 and 10 for SYNPCC7002_*A0001* or 11 and 12 for *sll0848*, respectively. The PCR product was cloned just downstream of the *dnaA* promoter of *S. elongatus* and the HA-coding sequence of the above vector (expressing HA-tagged DnaA of *S. elongatus*), which were amplified as a linear DNA using PCR with the primers 3 and 4 with the In-Fusion HD Cloning Kit. The resultant vector was used to transform *S. elongatus*. The primer sequences are listed in **Supplementary Table S1**.

To generate an *Synechococcus* sp. PCC 7002 strain expressing SSB-GFP driven by the endogenous *ssb* promoter, the amplicons were prepared as follows: 1. The *ssb orf* (SYNPCC7002_*A0119*) flanked by the 697-bp upstream sequence of *Synechococcus* sp. PCC 7002 was amplified using primers 7 and 8. 2. The *gfp*-(gentamycin resistance gene) *Gm*^{r} fusion was amplified using the primers 9 and 10 from the genomic DNA of *S. elongatus ssb-gfp* strain (Ohbayashi et al., 2019) as a template. 3. A 1,000-bp 3′-downstream sequence of *ssb orf* of *Synechococcus* sp. PCC 7002 was amplified using primers 11 and 12. The amplicons were mixed and fused using recombinant PCR with the primers 7 and 12, and the fused product was used to transform cells. Insertion of the *gfp-Gm*^{r} into the chromosomal *ssb* locus was confirmed using PCR with primers 7 and 12.

To produce a *dnaA* deletion strain of *Synechococcus* sp. PCC 7002, PCR products were prepared as follows: (1) A 731-bp upstream sequence of *Synechococcus* sp. PCC 7002 *dnaA orf* was amplified using the primers 19 and 20. The kanamycin resistance gene (*Km*^{r}) was amplified using primers 21 and 22. (2) A 700-bp 3′-downstream sequence of *Synechococcus* sp. PCC 7002 *dnaA orf* was amplified using primers 23 and 24. Amplified fragments were mixed and fused using recombinant PCR with primers 19 and 24, and the fused product was used to transform cells. Replacement of chromosomal *dnaA* with *Km*^{r} was confirmed using PCR with the primers 19 and 24.

To generate a *Synechococcus* sp. PCC 7002 strain expressing HA-DnaA driven by the endogenous *dnaA* promoter, PCR products were prepared as follows: (1) A 929-bp 5′-upstream sequence of *Synechococcus* sp. PCC 7002 *dnaA orf* was amplified using the primers 25 and 26. (2) An *HA-dnaA* fusion of *Synechococcus* sp. PCC 7002 was amplified using primers 27 and 28 from the vector described above (for expression in *S. elongatus*). (3) *Sp*^{r} was amplified using primers 29 and 30, and a 900-bp 3′-downstream sequence of *Synechococcus* sp. PCC 7002 *dnaA orf* was amplified using primers 31 and 32. Amplified fragments were mixed and fused using recombinant PCR with the primers 25 and 32, and the fused product was used to transform cells. Insertion of sequences encoding *HA* and *Sp*^{r} into the chromosomal *dnaA* locus was confirmed using PCR with primers 25, 32, and 33. The primer sequences are listed in **Supplementary Table S1**.

To generate a *Synechocystis* sp. PCC 6803 strain expressing HA-DnaA driven by the endogenous *dnaA* promoter, PCR products were prepared as follows: (1) A 700-bp upstream sequence of *Synechocystis* sp. PCC 6803 *dnaA orf* (*sll0848*) was amplified using primers 34 and 45. (2) *HA-dnaA* of *Synechocystis* sp. PCC 6803 fusion was amplified with the primers 36 and 37 from the vector described above (for expression in *S. elongatus*). (3) *Sp*^{r} was amplified using primers 38 and 39. 4. A 700-bp 3′-downstream sequence of *Synechocystis* sp. PCC 6803 *dnaA orf* was amplified using primers 40 and 41. The amplified fragments were mixed and fused using recombinant PCR with the primers 34 and 41, and the fused product was used to transform cells. Insertion of the sequence encoding *HA* and *Sp*^{r} into the chromosomal *dnaA* locus was confirmed using PCR with the primers 33, 34, and 41. The primer sequences are listed in **Supplementary Table S1**.

RESULTS

Distribution of *dnaA* in Cyanobacterial and Chloroplast Genomes

Cyanobacteria emerged on earth more than two billion years ago and globally diversified in numerous, diverse environments (Herrero and Flores, 2008). The genome sequences of 54 cyanobacterial species deposited in the Pasteur Culture Collection were published in Shih et al. (2013), and the complete genome sequences of > 500 cyanobacterial species are available in public databases. We first re-examined the distribution of *dnaA* sequences in these completely sequenced cyanobacterial and chloroplast genomes (**Figure 1A** and **Supplementary Figure S1**). Consistent with our previous search (Ohbayashi et al., 2016), we did not detect *dnaA* in the genomes of symbiotic cyanobacteria (**Figure 1A**). The phylogenetic distributions of these symbionts and chloroplasts suggest that loss of *dnaA* independently occurred in their respective ancestors (**Figure 1A**).

Although it has been believed that *dnaA* is conserved among free-living bacteria (Yoshikawa and Ogasawara, 1991; Messer, 2002; Gao and Zhang, 2007; Katayama et al., 2010), our search revealed that the genomes of the free-living cyanobacterial species *Cyanobacterium stanier*i PCC 7202, *Cyanobacterium* sp. PCC 10605, *Geminocystis* sp. PCC 6308, and *Geminocystis* NIES-3708 and 3709 do not encode *dnaA*. These cyanobacteria formed a monophyletic group (**Figure 1A**, clade 1). Further, the topology of the phylogenetic tree suggests that the common ancestor of this group (**Figure 1A**, clade 1) lost *dnaA* independently from the ancestors of the above-mentioned symbiotic species and chloroplasts.

The clade comprising the *dnaA*-negative and *dnaA*-positive free-living species (**Figure 1A**, clade 2) was further grouped with the clade (**Figure 1A**, clade 3) that contained *dnaA*-negative endosymbionts and *dnaA*-positive free-living species, grouped as the clade B2 (**Figure 1A**). Although the data supporting this grouping were not definitive (boot strap value = 66) this conclusion is strongly supported by previous analyses (Shih et al., 2013). Although clade B2 comprised several *dnaA*-positive species as well as *dnaA*-negative species, our previous and present studies showed that, in *Synechocystis* sp. PCC 6803 (Ohbayashi et al., 2016) and *Synechococcus* sp. PCC 7002 (this study, as described later) of this clade, *dnaA* is not essential for genome replication and cell growth. Thus, even in *dnaA*-positive species, chromosomes are replicated in a DnaA-*oriC* independent manner in at least these two species in clade B2.

Consistent with these results, chromosomes in *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 exhibited irregular GC skew profiles (**Figure 1B**). In addition, all sequenced species in the clade B2 exhibited irregular GC skew profiles regardless of the presence or absence of *dnaA* (**Figure 1B** and **Supplementary Figure S4**) (Typically, the position around the putative DnaA-box sequences near *dnaA* or *dnaN* is defined as a putative replication origin, and the circular genome sequence is deposited as a linear sequence from that position.). This contrasts with clades C-G in which the chromosomes of most species exhibited a clear V-shape (profile with two shift points) except for those of *Synechococcus* sp. PCC 6321 and the cyanobacterium-derived organelle (chromatophore) in *Paulinella chromatophora* (**Figure 1B** and **Supplementary Figure S4**). The clades C-G include the model species *S. elongatus* in which the chromosome is replicated from a unique origin (*oriC*), and DnaA is essential for chromosome replication (Watanabe et al., 2012; Ohbayashi et al., 2016).

GC skew profiles in the clade B1 exhibited an intermediate feature between the clades C-G and B2, in which two of the four sequenced genomes exhibited an V-shape profile (**Supplementary Figure S4**, *Nostoc* sp. PCC 7120 and *Anabaena cylindrica* PCC 7122). GC skew profiles in the clade A and chloroplast did not exhibit V-shape except for that of the chloroplast in *Chlamydomonas reinhardtii*. Overall, the V-shaped GC skew profile apparently tends to be collapsed in the order of clade C-G, B1, B2, to A and chloroplasts (**Figure 1** and **Supplementary Figure S4**). In addition, the ploidy level also apparently tends to increase from the clade C-G to A and chloroplasts although no information on ploidy level was available in the clade B1 (**Figure 1A**).

Similar to the GC skew, in clade B2, the CDS skew profiles of most species did not exhibit a regular pattern with two shift points (**Supplementary Figure S5**). However, the exceptions in clade B2 were the chromosomes of *Synechococcus* sp. 7002 and *Stanieria cyanosphaera* PCC 7437, which exhibited CDS skew profiles with a weak V-shape (**Figure 1B** and **Supplementary Figure S5**). In *Synechococcus* sp. 7002, the shift points are near positions 0 and 1.2 Mbp, respectively. This observation raised the possibility that chromosomes in certain *dnaA*-positive species in clade B2 are replicated in a DnaA-*oriC*-dependent manner, unlike *Synechocystis* sp. PCC 6803.

To acquire insights into the variation and evolution of the mechanisms of chromosome replication in clade B2, we next examined the manner of chromosome replication in free-living *dnaA*-negative *C. aponinum* PCC 10605 and *Geminocystis* sp. NIES-3708 as well as *dnaA*-positive *Synechococcus* sp. PCC 7002.

Chromosome Replication in the *dnaA*-Negative Free-Living Species *C. aponinum* PCC 10605 and *Geminocystis* sp. NIES-3708

To determine the ploidy of *C. aponinum* and *Geminocystis* sp., exponentially growing cells cultured in an inorganic medium with illumination (**Supplementary Figure S2**) were stained with the DNA-specific dye SYBR Green I, and the DNA levels and cell sizes were examined using flow cytometry (**Figures 2A,B**). For comparison, exponentially growing *S. elongatus* cells (3–6 copies of chromosomes per cell) (Chen et al., 2012; Jain et al., 2012; Zheng and O'Shea, 2017) in an inorganic medium with illumination (**Supplementary Figure S2**) were simultaneously examined (**Figures 2A,B**). The fluorescence intensity of SYBR Green per *C. aponinum* was approximately 8-times higher compared with that of *S. elongatus* (**Figure 2B**). The chromosome sizes of *C. aponinum* and *S. elongatus* are approximately 4.1 and 2.7 Mbp, respectively. Thus, the result indicates that *C. aponinum* possesses approximately 16–32 copies of its chromosomes. Similarly, the fluorescence intensity of SYBR Green per *Geminocystis* sp. was approximately 8-times higher compared with that of *S. elongatus* (**Figure 2B**). The chromosome of *Geminocystis* sp. is approximately 3.9 Mbp, indicating that *Geminocystis* sp. possesses approximately 17–34 copies of its chromosomes. Further, the volume of *C. aponinum* and *Geminocystis* sp. cells were approximately 8-times higher compared with that of *S. elongatus* (**Figure 2B**), showing that an increase in chromosomal ploidy correlated with an increase in cell size.

When bacteria with a single chromosome, such as *Escherichia coli*, grow rapidly in nutrient-rich media, DNA is replicated in a multifork mode, and the *oriC*/*ter* ratio becomes > 2 yielding a V-shaped profile in the depth of high-throughput sequencing reads (lowest at *ter* and highest at *oriC*) (Watanabe et al., 2012). For cyanobacteria possessing multiple copies of the same

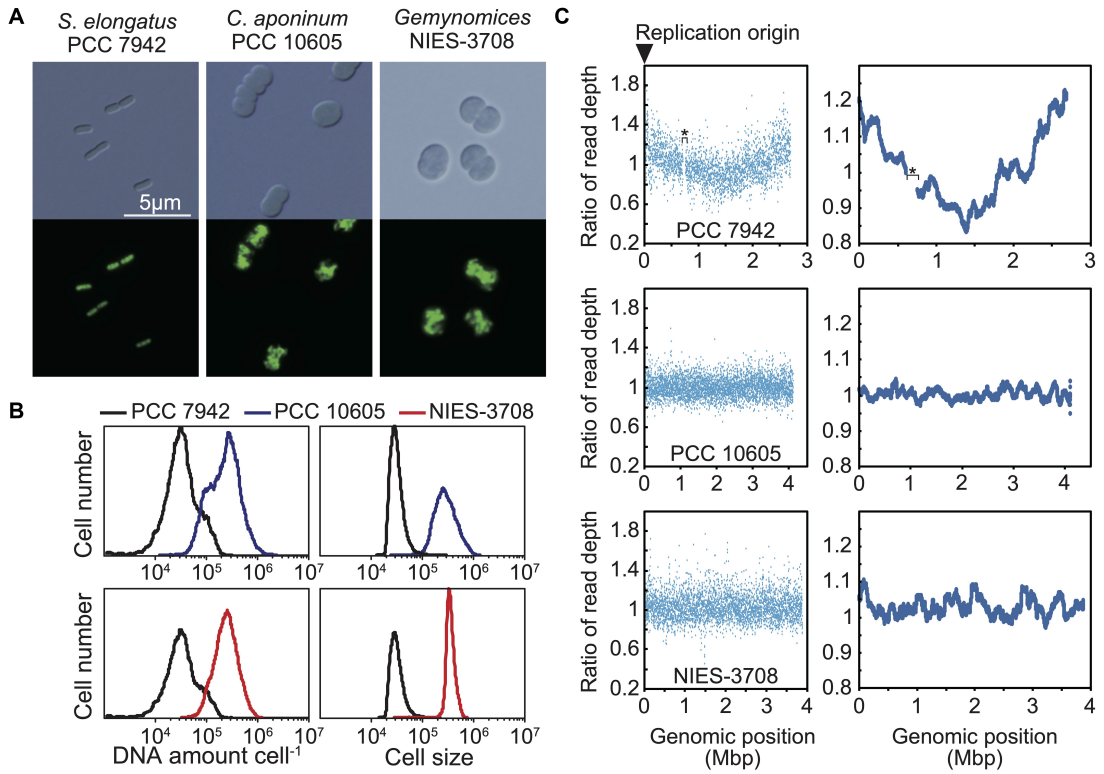


FIGURE 2 | Ploidy and replication manner of the chromosomes of *C. aponinum* PCC 10605 and *Geminocystis* sp. NIES-3708. Exponentially growing *C. aponinum* (chromosome approximately 4.1 Mbp), *Geminocystis* sp. (chromosome approximately 3.9 Mbp), and *S. elongatus* (chromosome approximately 2.7 Mbp; 3–6 copies per cell; Zheng and O'Shea, 2017) (**Supplementary Figure S2**) were fixed and stained with SYBR Green and then examined using flow cytometry. **(A)** Images of SYBR Green-stained cells. Images were acquired using differential interference contrast microscopy (top) and fluorescence microscopy (bottom). **(B)** Distribution of DNA levels per cell and cell volumes of exponentially growing cultures of *C. aponinum* (blue), *Geminocystis* sp. (red), and *S. elongatus* (black). **(C)** Depths of the high-throughput genomic DNA sequence reads at their respective chromosomal regions. Genomic DNA was extracted from the exponentially growing cells (**Supplementary Figure S2**) and analyzed using an Illumina MiSeq System. Plots of 1-kb (left) and 100-kb windows (right). The number of reads (divided by the number of total reads) of the growing (replicating) cells normalized by that of the stationary phase (non-replicating) cells at each genomic position is shown. The asterisk in the profile of *S. elongatus* indicates a ∼50-kb genomic deletion in our wild type strain which has little effect on replication and cellular growth (Watanabe et al., 2012).

chromosomes, in which only one or a few copies are replicated from *oriC*, the *oriC*/*ter* ratio approximates 1.0, but still exhibits a V-shaped profile [Watanabe et al., 2012; **Figure 2C**; The read depth of the exponentially growing (replicating) cells at each genomic potion was normalized by that of stationary phase (non-replicating) cells]. In cyanobacterial species that asynchronously initiate chromosome replication from multiple sites, the ratio of DNA abundance through the genomic position becomes almost constant, as represented by *Synechocystis* sp. PCC 6803 (Ohbayashi et al., 2016). When high-throughput sequence reads of exponentially growing (replicating) *C. aponinum* and *Geminocystis* sp. (**Supplementary Figure S2**) were mapped to the reference genome, the read depth normalized by that of stationary phase (non-replicating) cells was approximately constant throughout the chromosome (**Figure 2C**). These results suggest that in the *dnaA*-negative cyanobacteria *C. aponinum* and *Geminocystis* sp., replication of multiple copies of the same chromosomes is asynchronously initiated from multiple sites rather than a unique point, as for *Synechocystis* sp. PCC 6803.

Dependence of Chromosome Replication on DnaA in *Synechococcus* sp. PCC 7002

To determine the ploidy of *Synechococcus* sp. PCC 7002, exponentially growing cells in an inorganic medium with illumination (**Supplementary Figure S2**) were stained with the SYBR Green I, and the DNA level and cell size were examined using flow cytometry (**Figure 3**). Exponentially growing *S. elongatus* cells (3–6 copies of chromosomes per cell, genome approximately 2.7 Mbp) (Chen et al., 2012; Jain et al., 2012; Zheng and O'Shea, 2017) in an inorganic medium with illumination (**Supplementary Figure S2**) was simultaneously compared (**Figure 3**). The fluorescence intensity of SYBR Green I per *Synechococcus* sp. PCC 7002 (genome approximately 3.0 Mbp) was approximately 2-times higher compared with that of *S. elongatus* (**Figure 3B**). These results indicate that *Synechococcus* sp. PCC 7002 possesses approximately 5–11 copies of its chromosomes, consistent with

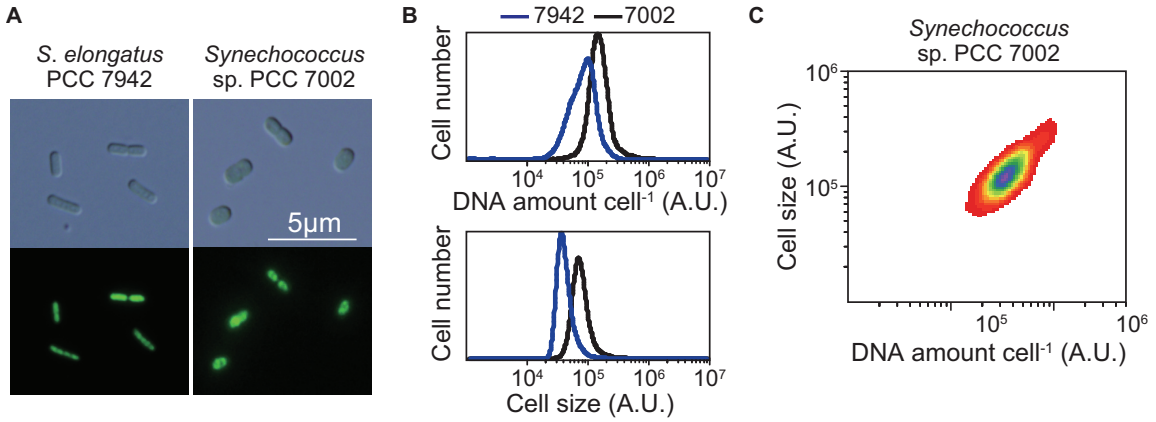


FIGURE 3 | Ploidy of the chromosome of *Synechococcus* sp. PCC 7002. Exponentially growing *Synechococcus* sp. PCC 7002 (chromosome, approximately 3.0 Mbp) and for comparison, *S. elongatus* (chromosome, approximately 2.7 Mbp; 3–6 copies per cell; Zheng and O'Shea, 2017) (**Supplementary Figure S2**) were fixed and stained with SYBR Green and then examined using flow cytometry. **(A)** SYBR Green-stained cells. Images were acquired using differential interference contrast microscopy (top) and fluorescence microscopy (bottom). **(B)** Distribution of the DNA levels per cell and cell volumes of exponentially growing cultures of *Synechococcus* sp. PCC 7002 (black) and *S. elongatus* (blue). **(C)** The relationship between the DNA level and size of *Synechococcus* sp. PCC 7002.

the findings of others (Moore et al., 2019; this article is preprint). The cell volume of *Synechococcus* sp. PCC 7002 was approximately 2-times higher compare with that of *S. elongatus* (**Figure 3B**). Further, the amount of DNA (chromosome copy number) exhibited a linear, positive correlation with cell volume (**Figure 3C**).

To examine the dependence of DNA replication on DnaA in *Synechococcus* sp. PCC 7002, we constructed a *dnaA* deletion mutant in which *dnaA* was replaced with a kanamycin resistance gene (Km^r) through homologous recombination. The insertion of Km^r into the *dnaA* locus was confirmed using PCR and we obtained some transformed clones, in which *dnaA* was completely deleted (**Figure 4A**). The growth rates of Δ*dnaA* clones were similar to that of wild type (**Figure 4B** and **Supplementary Figure S6B**). There was no significant difference in the frequency of chromosome replication between wild type and the Δ*dnaA* mutant (**Figure 4C** and **Supplementary Figure S6C**), as indicated by the number of SSB-GFP foci in an exponentially growing cell (**Figure 4B** and **Supplementary Figure S3A**). Note that SSB localizes to replication forks (Chen et al., 2012; Mangiameli et al., 2017). Further, the chromosome copy number and cell size of completely segregated Δ*dnaA* clones of *Synechococcus* sp. PCC 7002 were similar to those of the wild type (**Figures 4D,E**). Thus, there were no significant differences in the chromosome replication and proliferation rates between the wild type and the completely segregated Δ*dnaA*.

We previously isolated mutants of *Synechocystis* sp. PCC 6803 and *Nostoc* sp. PCC 7120 with completely segregated Δ*dnaA* (Ohbayashi et al., 2016), in which chromosomes were replicated from multiple origins in wild type and Δ*dnaA* cells. Further, complete deletion of *dnaA* does not significantly affect the chromosome replication and proliferation rates of these species (Ohbayashi et al., 2016). In contrast, we isolated completely and incompletely segregated Δ*dnaA* clones of *S. elongatus* (Ohbayashi et al., 2016). In these mutants, Δ*dnaA* completely segregated via an episomal plasmid that integrated into the chromosome (Ohbayashi et al., 2016). Further, the replication initiation site of the chromosome shifted from *oriC* to the origin of the integrated plasmid (Ohbayashi et al., 2016). Accordingly, we next determined whether the Δ*dnaA* of *Synechococcus* sp. PCC 7002 harbored an additional suppressor mutation such as that found in *S. elongatus* Δ*dnaA*. We therefore analyzed the complete genome sequence and profile of genome replication of this mutant (**Figure 5** and **Supplementary Figure S6**).

When the high-throughput sequence reads of exponentially growing wild type *Synechococcus* 7002 (**Supplementary Figure S2**) were mapped to the reference genome, a V-shaped profile was observed (**Figure 5**; the read depth of growing cells at each genomic position was normalized that of stationary phase cells). The peak was detected around *dnaA* at position 1 (the leftmost base) (**Figure 5**), suggesting that the wild type possessed a unique replication origin near *dnaA*. When the sequence reads of the Δ*dnaA* clones harvested during log phase were mapped to the reference genome, all clones exhibited a V-shaped profile similar to that of the wild type (**Figure 5**), suggesting that the replication of Δ*dnaA* started from a nearby site. Further, mutations such as in-del or point mutation, plasmid integration and transposition of *dnaA* gene were not detected in the Δ*dnaA* clones by high-throughput genomic DNA sequencing (**Supplementary Figure S6A** and **Supplementary Data Sheet**). We concluded therefore that *dnaA* was not essential for the proliferation of *Synechococcus* sp. PCC 7002 and that the wild type chromosome replicated in a DnaA-*oriC* independent manner starting from a specific position, unlike that of *Synechocystis* sp. PCC 6803.

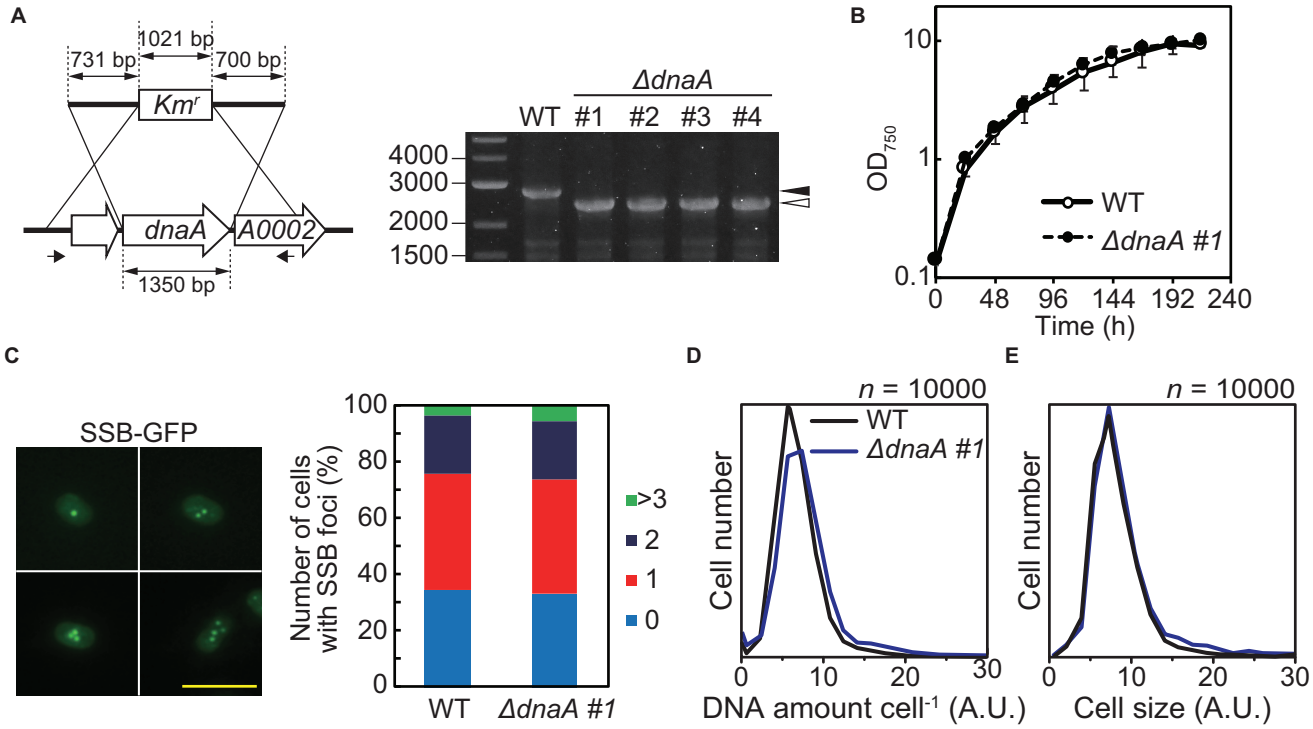


FIGURE 4 | Effect of deleting *dnaA* on the growth and DNA replication of *Synechococcus* sp. PCC 7002. **(A)** Chromosomal *dnaA* was replaced with the gene encoding kanamycin resistance (*Km*^{r}) through homologous recombination. Insertion of *Km*^{r} into the chromosomal *dnaA* locus was confirmed using PCR with the primers indicated by the arrows. The wild type (WT) was used as a negative control. The black and white arrowheads indicate the bands amplified from the WT and mutated chromosomes, respectively. The four independent colonies of the *dnaA*-deficient mutant were analyzed. **(B)** Growth of WT and Δ*dnaA* clone #1 in an inorganic medium with illumination (70 μmol m^{-2} s^{-1}). Other Δ*dnaA* clones (#2–#4) are shown in **Supplementary Figure S6B**. **(C)** Frequencies of cells exhibiting zero (blue), one (red), two (deep blue), or >3 (green) SSB-GFP foci in WT and Δ*dnaA* clone #1 cultures 12 h after inoculation (n > 300 cells, each strain). For reference, a representative fluorescence image of WT expressing SSB-GFP (green fluorescence) is shown (one to four SSB-GFP foci are shown). Scale bar = 5 μm. The construction of SSB-GFP expresser is described in **Supplementary Figure S3**. **(D,E)** Flow cytometric analysis showing the distributions of the DNA level per cell **(D)** and the cell volumes **(E)** of cultures of the WT (black line) and Δ*dnaA* clone #1 (blue line) 12 h after inoculation, respectively. DNA was stained with SYTOX Green, and the levels were determined according to the intensity of the green fluorescence.

Comparison of the Function and Expression Levels of DnaA in *S. elongatus*, *Synechocystis* sp. PCC 6803, and *Synechococcus* sp. PCC 7002

Synechococcus sp. PCC 7002 is phylogenetically closely related to *Synechocystis* sp. PCC 6803 (**Figure 1A**), in which the chromosome is replicated from multiple origins and *dnaA* is dispensable without additional suppressor mutations (Ohbayashi et al., 2016). Further, the amino acid sequence of DnaA of *S. elongatus* is similar to that of *Synechococcus* sp. PCC 7002 (61% identical) and *Synechocystis* sp. PCC 6803 (59% identical). However, previous and the present results suggest that *dnaA* is non-essential for *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 but is essential for chromosome replication of *S. elongatus*.

To determine the basis for this difference, we asked whether DnaA molecules in the respective species bind DnaA-box sequences. We therefore expressed HA-tagged DnaAs of *Synechocystis* sp. PCC 6803, *Synechococcus* sp. PCC 7002, and *S. elongatus* (positive control) under the control of the *S. elongatus dnaA* promoter in *S. elongatus* (**Figures 6A,B**). ChIP-qPCR analysis using an anti-HA antibody revealed that DnaA of each of the three species specifically bound *oriC*, but not *orf*1294, which is farthest from *oriC* in the circular genome and does not possess a DnaA-box sequence (**Figure 6C**).

We next expressed HA-DnaA in each of the above species under the control of their respective endogenous *dnaA* promoters (**Supplementary Figure S3B,C**). Immunoblotting using an anti-HA antibody showed that the level of HA-DnaA in *Synechococcus* sp. PCC 7002 (deduced size including the HA-tag = 54.4 kDa) was lower compared with that of *S. elongatus* (deduced size including HA-tag = 55.9 kDa) (**Figure 7**). Further, we did not detect HA-DnaA in *Synechocystis* sp. PCC 6803 (deduced size including HA-tag = 53.9 kDa) (**Figure 7**). Thus, in contrast to *S. elongatus*, in which *dnaA* is essential for chromosome replication, the level of DnaA was below the detection limit in *Synechocystis* sp. PCC 6803 or very low in *Synechococcus* sp.

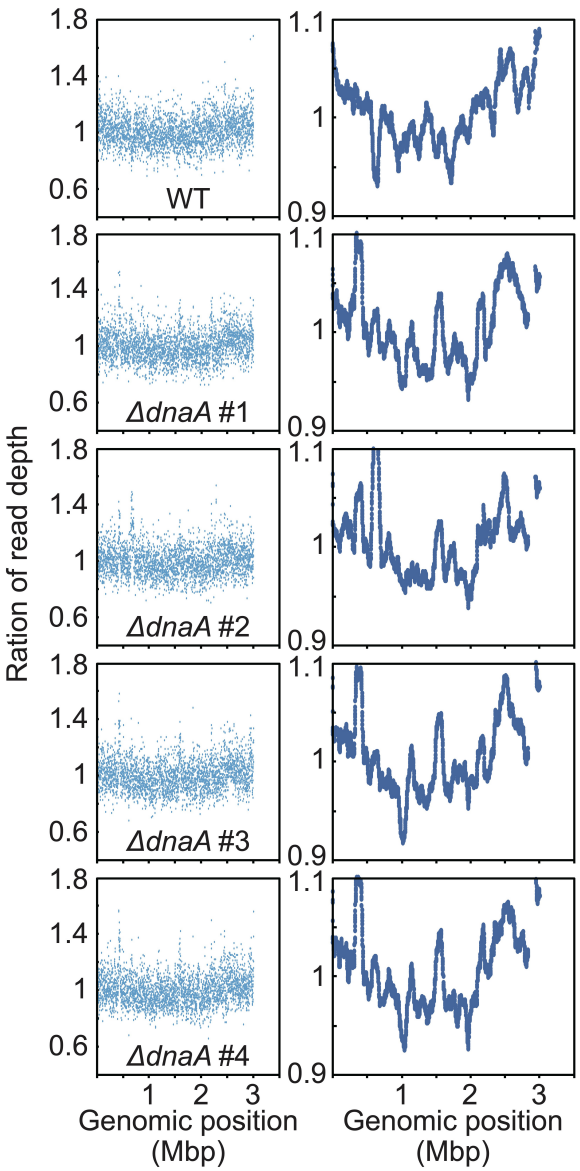


FIGURE 5 | Chromosomal replication origins of *Synechococcus* sp. PCC 7002 WT and *dnaA* disruptants. Depth of the high-throughput genomic DNA reads at respective chromosomal regions in exponentially growing WT and Δ*dnaA* cells. Genomic DNA was extracted from cells 24 h after inoculation (**Figure 4B**) and analyzed using an Illumina MiSeq System. The number of reads (divided by the number of total reads) of the growing (replicating) cells was normalized by that of the stationary phase (non-replicating) cells at each genomic position. 1-kb window (left) and 100-kb window (right) of WT and Δ*dnaA* clones.

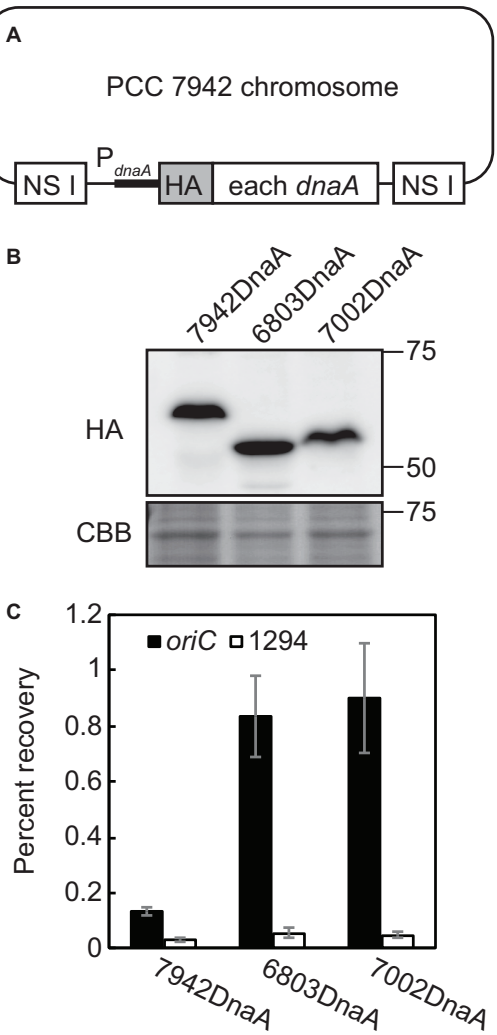


FIGURE 6 | Binding of DnaA to the DnaA-boxes of *Synechococcus elongatus*, *Synechocystis* sp. PCC 6803, and *Synechococcus* sp. PCC 7002. **(A)** Diagram of the *S. elongatus* chromosomes expressing HA-tagged DnaA of *S. elongatus* (7942 DnaA), *Synechocystis* (6803 DnaA), and *Synechococcus* 7002 (7002 DnaA). DNA encoding the respective DnaA was integrated into the chromosomal neutral site I (NS I) of *Synechococcus elongatus*. HA-DnaA expression was driven by the promoter of *S. elongatus dnaA*. **(B)** Immunoblot analysis of the expression of HA-DnaA in *S. elongatus*. Total proteins were subjected to immunoblotting using an anti-HA antibody. Coomassie Brilliant Blue (CBB)-stained proteins are shown as a loading control. **(C)** ChIP-qPCR analysis of the affinity of the binding of DnaA to the *oriC* region (DnaA-boxes) of the *S. elongatus* chromosome. The DnaA-chromatin complex was immunoprecipitated using an anti-HA antibody. The samples were quantified using qPCR with the primers representing the *oriC* region (*oriC*: solid bar) and *Syf1294*, which is farthest from *oriC* in the circular chromosome and lacks a DnaA-box sequence (1294: open bar). The percent recoveries of input DNA are indicated with the standard deviation (n = 3 biological replicates).

PCC 7002, in which complete deletion of *dnaA* had no effect on chromosome replication.

DISCUSSION

The chromosome of *S. elongatus* is replicated from a unique origin (*oriC*) in a DnaA-dependent manner, similar to the mechanism employed by most bacterial species (Ohbayashi et al., 2016). In contrast, in *Synechocystis* sp. PCC 6803 and *Nostoc* sp. 7120, DnaA is not required for chromosome replication, which is initiated from multiple sites (Ohbayashi et al., 2016). Here we extended these findings

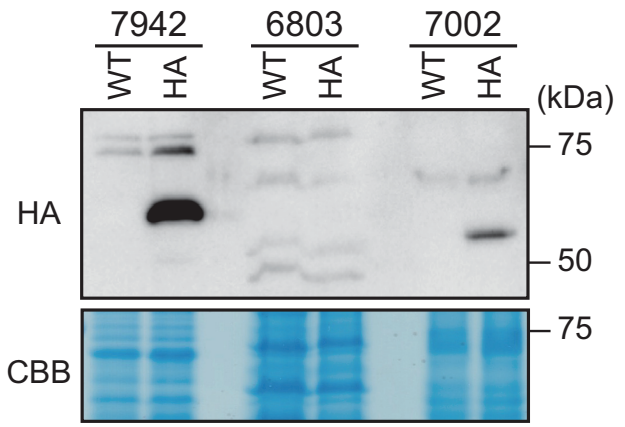


FIGURE 7 | Comparison of DnaA expression levels in *Synechococcus elongatus*, *Synechocystis* sp. PCC 6803, and *Synechococcus* sp. PCC 7002. DNAs encoding HA-tagged DnaA of *S. elongatus* (7942), *Synechocystis* sp. PCC 6803 (6803), or *Synechococcus* sp. PCC 7002 (7002) were integrated into the respective chromosomal *dnaA* locus (Ohbayashi et al., 2019; **Supplementary Figure S3B,C**). HA-DnaA expression was driven by the endogenous *dnaA* promoter. Exponentially growing transformants were inoculated into fresh inorganic medium and cultured for 6 h with illumination (70 μmol m^{-2} s^{-1}). The same amounts (80 μg) of proteins extracted from the respective cultures were subjected to immunoblotting using an anti-HA antibody. The respective WTs served as negative controls. CBB-stained protein samples are shown as a loading control.

to show that chromosome replication in *Synechococcus* sp. PCC 7002, which is evolutionarily related to *Synechocystis* sp. PCC 6803, is initiated from a unique origin in a DnaA-independent manner, unlike *Synechocystis* sp. PCC 6803. We further found that certain free-living cyanobacterial species do not possess *dnaA*. The DnaA-*oriC*-independent species *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. 6803, four *dnaA*-negative free-living species, and two *dnaA*-negative endosymbiotic species are phylogenetically closely related (clade B2 in **Figure 1A**). Their phylogenetic positions suggest that (1) loss of *dnaA* from the common ancestor of the four *dnaA*-negative free-living species (**Figure 1A**, clade 1), (2) the loss of *dnaA* from the symbiont *Epithemia turgida*, and (3) loss of *dnaA* in the symbiont *Atelocyanobacterium* were independent events. Further, the loss of *dnaA* from other endosymbiotic species (*Nostoc azollae* 0708 and *Paulinella chromatophora*), as well as from the common ancestor of chloroplasts, occurred independently. Similarly, an ancestor of *Nostoc* sp. 7120 (clade B1 in **Figure 1A**) independently lost DnaA-*oriC*-dependence of an ancestor(s) of *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 (**Figure 1A**). Thus, the DnaA-*oriC*-dependent chromosome replication mechanism was lost multiple times during cyanobacterial evolution.

In bacterial groups other than cyanobacteria, there has been no report on free-living species that does not possess *dnaA* gene. However, the genomes of certain bacterial symbionts of insects such as *Wigglesworthia glossinidia* (Akman et al., 2002), *Blochmannia floridanus* (Gil et al., 2003), and Candidatus *Carsonella ruddii* (Nakabachi et al., 2006) do not encode *dnaA*, although they encode genes required for replication (e.g., DNA helicase, DNA polymerase, and DNA primase) (Klasson and Andersson, 2004). Further, mitochondrial and their eukaryotic host genomes do not encode *dnaA* (Ohbayashi et al., 2016). Moreover, certain *dnaA*-negative symbionts possess multiple copies of the same chromosomes as do mitochondria, chloroplasts, and most cyanobacterial species (Bendich, 1987; Griese et al., 2011; Sargent et al., 2016). Thus, loss of DnaA-dependent chromosome replication and loss of *dnaA* are associated with an increase in chromosomal copy number per cell/organelle. At this point, it is unclear how the loss of *dnaA* and endosymbioses are related. One possibility is that loss of a DnaA-*oriC*-regulated mechanism of chromosome replication in endosymbionts was presumably advantageous for host cells to regulate the proliferation of endosymbionts and the replication of their chromosomes.

Another important question is the nature of the function of DnaA in *dnaA*-positive but DnaA-*oriC*-independent species such as *Synechococcus* sp. PCC 7002, *Synechocystis* sp. PCC 6803, and *Nostoc* sp. PCC 7120. Another function of DnaA is to regulate the transcription of a discrete set of genes. For example, *B. subtilis* DnaA binds DnaA-boxes of eight intergenic chromosomal regions and positively or negatively regulates transcription of specific genes (Ishikawa et al., 2007). However, complete disruption of *dnaA* of *Synechocystis* sp. PCC 6803, *Nostoc* sp. PCC 7120 (Ohbayashi et al., 2016), and *Synechococcus* sp. PCC 7002 (this study) did not affect growth under optimal growth conditions. Further, we show here that when *Synechocystis* sp. PCC 6803 is grown under optimal growth conditions, the level of DnaA was below the detection limit. Thus, even if DnaA is involved in transcriptional regulation, the function is not essential in these three DnaA-*oriC*-independent species. At this point, the function of DnaA in these species remains unclear.

Although the GC and CDS skews of most bacteria exhibit clear asymmetric profiles with shift points at *ori* and *ter*, the chromosomes of most cyanobacterial species exhibit irregular patterns. Exceptions such as *S. elongatus*, which exhibits regular GC and CDS skew profiles, were found only in the clades C-G in the phylogenetic tree (**Figure 1** and **Supplementary Figures S4**, **S5**). The common feature of this group is its relatively lower ploidy level compared with other clades (Griese et al., 2011; Sargent et al., 2016). Further, species with relatively reduced genome sizes only occur in the clades C-G (Shih et al., 2013). Thus, loss of regular GC and CDS skews during cyanobacterial evolution presumably correlated with an increase in the chromosomal copy number, although further characterization of genome copy number in many more cyanobacterial species (species in which ploidy level has not been determined in **Figure 1A**) is required. It is assumed that in most bacteria the asymmetrical replication machinery between that of the leading strand and that of the discontinuous replication in the lagging strand contributes to differential mutational bias (Bhagwat et al., 2016). Further, evidence indicates that

transcriptional mutations create strand-specific nucleotide compositional skew (asymmetric GC skew) (Francino et al., 1996; Rocha et al., 2006; Chen et al., 2016). In most bacteria, genes are preferentially encoded by the leading strand (asymmetric CDS skew), which may be advantageous to avoid the head-on collision of DNA and RNA polymerases (Merrikh et al., 2012; Hamperl and Cimprich, 2016). Further, a preference in the third codon position for G vs. C and T vs. A in bacterial genes may have contributed to the creation of strand-specific nucleotide compositions (Kerr et al., 1997; McLean et al., 1998). In most cyanobacteria, multiple copies of the same chromosomes are replicated asynchronously while transcription occurs in all copies (Ohbayashi et al., 2019). This characteristic of multiple copies of the same chromosomes theoretically reduces the frequency of the head-on collision of DNA and RNA polymerases, leading to loss of regular GC and CDS skews during cyanobacterial evolution.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

ROh and SM designed the study. ROh, SH, YKa, and YH performed the nucleotide sequence analyses. ROh and ROn performed the phylogenetic analysis. ROh performed the all other experiments. YKo, TF, and CF provided new reagents and analytical tools. ROh and SM wrote the manuscript.

FUNDING

This work was supported by grants-in-aid for Scientific Research from the Japan Society for the Promotion of Science (no. 16H07418 to ROh, no. 18H04827 to ROh, and no. 17H01446 to SM), the JST-MIRAI Program of Japan Science and Technology Agency (JST) (to ROh), an NIG postdoctoral fellowship (to ROh), and by RIKEN SPDR research funding (to ROh).

ACKNOWLEDGMENTS

We thank Dr. Masato Kanemaki and Dr. Toyoaki Natsume of the National Institute of Genetics for technical advice and Kumi Tanabe and Hazuki Kotani of the Furusawa Laboratory for technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00786/full#supplementary-material

FIGURE S1 | The full phylogenetic tree (partially shown in **Figure 1A**) with the outgroups and accession numbers of sequences. Branch lengths are proportional to the number of nucleotide substitutions indicated below the tree. Sequences of *Escherichia coli* K-12, *Rhodobacter sphaeroides*, and *Bacillus subtilis* 168 served as an outgroup. The accession numbers are indicated next to the species name. Other details are described in **Figure 1**.

FIGURE S2 | Growth curves of *S. elongatus*, *C. aponinum* PCC 10605, and *Geminocystis* sp. NIES-3708. **(A)** An exponentially growing culture of each species was inoculated into fresh inorganic medium and cultured with air bubbling in the light (70 μmol m^{-2} s^{-1}) at 30°C. The arrow indicates the sampling time for the other analyses. The black and white arrowheads indicate sampling points of exponential and stationary phases, respectively. **(B)** Growth rate at the exponential phase.

FIGURE S3 | Generation of transformants. **(A)** To express SSB-GFP in *Synechococcus* sp. PCC 7002, the *gfp orf* was integrated into the chromosome immediately before the stop codon of the *ssb* loci of the wild type (WT) and Δ*dnaA* genomes. Gm^r was used as a selectable marker. Insertion of *gfp* and Gm^r into the chromosomal *ssb* locus was confirmed using PCR with the primers indicated by the arrows below the illustration. The WT served as a negative control. **(B,C)** To express HA-DnaA under the control of the *dnaA* promoters of *Synechococcus* sp. PCC 7002 **(B)** and *Synechocystis* sp. PCC 6803 **(C)**, DNA encoding HA-DnaA was integrated into the chromosomal *dnaA* locus of each species. Sp^r was used as a selectable marker. Insertion of the gene encoding *HA* and Sp^r into the chromosomal *dnaA* locus of each species was confirmed using the primers indicated by the arrows below the illustrations.

FIGURE S4 | Cumulative GC skew profiles of cyanobacterial species not shown in **Figure 1B**. *Nostoc azollae* possesses a pseudo-*dnaA* gene, which is disrupted by insertion of a transposon. Other details are described in **Figure 1**.

FIGURE S5 | Cumulative CDS skew profiles of cyanobacterial species not shown in **Figure 1B**. The details are described in **Figure 1**.

FIGURE S6 | Growth and chromosomal replication of *Synechococcus* sp. PCC 7002 Δ*dnaA* strains. **(A)** High-throughput genomic DNA reads of WT and Δ*dnaA* strains analyzed using IGV software. Genomic positions (1–3000 bases) including *dnaA* are shown. **(B)** Growth of Δ*dnaA* clones not shown in **Figure 4**. **(C)** Frequency of cells exhibiting 0 (blue), 1 (red), 2 (deep blue), or ≥3 (green) SSB-GFP foci in Δ*dnaA* clones not shown in **Figure 4**.

TABLE S1 | Primers used in this study.

DATA SHEET S1 | BRESEQ analysis of WT *Synechococcus* sp. PCC 7002 and Δ*dnaA* strains.

REFERENCES

Aikawa, S., Nishida, A., Ho, S.-H., Chang, J.-S., Hasunuma, T., and Kondo, A. (2014). Glycogen production for biofuels by the euryhaline cyanobacteria *Synechococcus* sp. *strain PCC* 7002 from an oceanic environment. *Biotechnol. Biofuels* 7:88. doi: 10.1186/1754-6834-7-88

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., et al. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies. *Wigglesworthia glossinidia*. Nat. Genet. 32, 402–407. doi: 10.1038/ng986

Arakawa, K., Mori, K., Ikeda, K., Matsuzaki, T., Kobayashi, Y., and Tomita, M. (2003). G-language genome analysis environment: a workbench for nucleotide sequence data mining. *Bioinformatics* 19, 305–306. doi: 10.1093/bioinformatics/19.2.305

Arakawa, K., and Tomita, M. (2007). The GC skew index: a measure of genomic compositional asymmetry and the degree of replicational selection. *Evol. Bioinform.* 3:117693430700300006.

Beattie, T. R., and Reyes-Lamothe, R. (2015). A replisome's journey through the bacterial chromosome. *Front. Microbiol.* 6:562. doi: 10.3389/fmicb.2015.00562

Bell, S. D. (2017). Initiation of DNA replication in the archaea. *Adv. Exp. Med. Biol.* 1042, 99–115. doi: 10.1007/978-981-10-6955-0_5

Bendich, A. J. (1987). Why do chloroplasts and mitochondria contain so many copies of their genome? *BioEssays* 6, 279–282. doi: 10.1002/bies.950060608

Bentley, S. D., and Parkhill, J. (2004). Comparative genomic structure of prokaryotes. *Annu. Rev. Genet.* 38, 771–791. doi: 10.1146/annurev.genet.38.072902.094318

Bhagwat, A. S., Hao, W., Townes, J. P., Lee, H., Tang, H., and Foster, P. L. (2016). Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2176–2181. doi: 10.1073/pnas.1522325113

Binder, B. J., and Chisholm, S. W. (1990). Relationship between DNA cycle and growth rate in *Synechococcus* sp. strain *PCC* 6301. *J. Bacteriol.* 172, 2313–2319. doi: 10.1128/jb.172.5.2313-2319.1990

Binder, B. J., and Chisholm, S. W. (1995). Cell cycle regulation in marine *Synechococcus* sp. strains. *Appl. Enviro. Microbiol.* 61, 708–717. doi: 10.1128/aem.61.2.708-717.1995

Burkholder, W. F., Kurtser, I., and Grossman, A. D. (2001). Replication initiation proteins regulate a developmental checkpoint in bacillus subtilis. *Cell* 104, 269–279. doi: 10.1016/s0092-8674(01)00211-2

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Chen, A. H., Afonso, B., Silver, P. A., and Savage, D. F. (2012). Spatial and temporal organization of chromosome duplication and segregation in the cyanobacterium *Synechococcus elongatus* PCC 7942. *PloS One* 7:e47837. doi: 10.1371/journal.pone.0047837

Chen, W.-H., Lu, G., Bork, P., Hu, S., and Lercher, M. J. (2016). Energy efficiency trade-offs drive nucleotide usage in transcribed regions. *Nat. Commun.* 7:11334. doi: 10.1038/ncomms11334

Deatherage, D.E., and Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* 1151, 165–188. doi: 10.1007/978-1-4939-0554-6_12

Dewar, J. M., and Walter, J. C. (2017). Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* 18, 507–516. doi: 10.1038/nrm.2017.42

Duggin, I. G., Wake, R. G., Bell, S. D., and Hill, T. M. (2008). The replication fork trap and termination of chromosome replication. *Mol. Microbiol.* 70, 1323–1333. doi: 10.1111/j.1365-2958.2008.06500.x

Francino, M. P., Chao, L., Riley, M. A., and Ochman, H. (1996). Asymmetries generated by transcription-coupled repair in enterobacterial genes. *Science* 272, 107–109. doi: 10.1126/science.272.5258.107

Freeman, J. M., Plasterer, T. N., Smith, T. F., and Mohr, S. C. (1998). Patterns of genome organization in bacteria. *Science* 279, 1827–1827.

Gao, F., and Zhang, C. T. (2007). DoriC: a database of oriC regions in bacterial genomes. *Bioinformatics* 23, 1866–1867. doi: 10.1093/bioinformatics/btm255

Gil, R., Silva, F. J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., et al. (2003). The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci. U.S,A.* 100, 9388–9393. doi: 10.1073/pnas.1533499100

Griese, M., Lange, C., and Soppa, J. (2011). Ploidy in cyanobacteria. *FEMS Microbiol. Lett.* 323, 124–131. doi: 10.1111/j.1574-6968.2011.02368.x

Grigoriev, A. (1998). Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* 26, 2286–2290. doi: 10.1093/nar/26.10.2286

Hamperl, S., and Cimprich, K. A. (2016). Conflict resolution in the genome: how transcription and replication make it work. *Cell* 167, 1455–1467. doi: 10.1016/j.cell.2016.09.053

Hanaoka, M., and Tanaka, K. (2008). Dynamics of RpaB–promoter interaction during high light stress, revealed by chromatin immunoprecipitation (ChIP) analysis in *Synechococcus elongatus* PCC 7942. *Plant J.* 56, 327–335. doi: 10.1111/j.1365-313X.2008.03600.x

Herrero, A., and Flores, E. (2008). *The Cyanobacteria: Molecular Biology, Genomics, and Evolution*. Poole: Horizon Scientific Press.

Hottes, A. K., Shapiro, L., and Mcadams, H. H. (2005). DnaA coordinates replication initiation and cell cycle transcription in *Caulobacter crescentus*. *Mol. Microbiol.* 58, 1340–1353. doi: 10.1111/j.1365-2958.2005.04912.x

Hu, B., Yang, G., Zhao, W., Zhang, Y., and Zhao, J. (2007). MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Mol. Microbiol.* 63, 1640–1652. doi: 10.1111/j.1365-2958.2007.05618.x

Huang, H., Song, C. C., Yang, Z. L., Dong, Y., Hu, Y. Z., and Gao, F. (2015). Identification of the replication origins from *Cyanothece* ATCC 51142 and their interactions with the dnaa protein: from in silico to in vitro studies. *Front. Microbiol.* 6:1370. doi: 10.3389/fmicb.2015.01370

Ishikawa, S., Ogura, Y., Yoshimura, M., Okumura, H., Cho, E., Kawai, Y., et al. (2007). Distribution of stable DnaA-binding sites on the *Bacillus subtilis* genome detected using a modified ChIP-chip method. *DNA Res.* 14, 155–168. doi: 10.1093/dnares/dsm017

Jain, I. H., Vijayan, V., and O'shea, E. K. (2012). Spatial ordering of chromosomes enhances the fidelity of chromosome partitioning in cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13638–13643. doi: 10.1073/pnas.1211144109

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kerr, A. R., Peden, J. F., and Sharp, P. M. (1997). Systematic base composition variation around the genome of *Mycoplasma genitalium*, but not *Mycoplasma pneumoniae*. *Mol. Microbiol.* 25, 1177–1179. doi: 10.1046/j.1365-2958.1997.5461902.x

Klasson, L., and Andersson, S. G. (2004). Evolution of minimal-gene-sets in host-dependent bacteria. *Trends Microbiol.* 12, 37–43. doi: 10.1016/j.tim.2003.11.006

Lobry, J. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* 13, 660–665. doi: 10.1093/oxfordjournals.molbev.a025626

Luo, H., and Gao, F. (2018). DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids Res.* 47, D74–D77. doi: 10.1093/nar/gky1014

Mackiewicz, P., Zakrzewska-Czerwińska, J., Zawilak, A., Dudek, M. R., and Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Res.* 32, 3781–3791. doi: 10.1093/nar/gkh699

Mangiameli, S. M., Veit, B. T., Merrikh, H., and Wiggins, P. A. (2017). The replisomes remain spatially proximal throughout the cell cycle in bacteria. *PLoS Genet.* 13:e1006582. doi: 10.1371/journal.pgen.1006582

Marks, A. B., Fu, H., and Aladjem, M. I. (2017). Regulation of replication origins. *Adv. Exp. Med. Biol.* 1042, 43–59. doi: 10.1007/978-981-10-6955-0_2

McLean, M. J., Wolfe, K. H., and Devine, K. M. (1998). Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. *J. Mol. Evol.* 47, 691–696. doi: 10.1007/pl00006428

Merrikh, H., Zhang, Y., Grossman, A. D., and Wang, J. D. (2012). Replication-transcription conflicts in bacteria. *Nat. Rev. Microbiol.* 10, 449–458. doi: 10.1038/nrmicro2800

Messer, W. (2002). The bacterial replication initiator *DnaA*. *DnaA* and *oriC*, the bacterial mode to initiate DNA replication. *FEMS Microbiol. Rev.* 26, 355–374. doi: 10.1016/s0168-6445(02)00127-4

Messer, W., and Weigel, C. (1997). DnaA initiator—also a transcription factor. *Mol. Microbiol.* 24, 1–6. doi: 10.1046/j.1365-2958.1997.3171678.x

Moore, K. A., Tay, J. W., and Cameron, J. C. (2019). Multi-generational analysis and manipulation of chromosomes in a polyploid cyanobacterium. *bioRxiv* [Preprint]. doi: 10.1101/661256

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., et al. (2006). The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314:267. doi: 10.1126/science.1134196

Nakayama, T., Kamikawa, R., Tanifuji, G., Kashiyama, Y., Ohkouchi, N., Archibald, J. M., et al. (2014). Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc. Natl. Acad. Sci. U.S.A.* 111, 11407–11412. doi: 10.1073/pnas.1405222111

Necşulea, A., and Lobry, J. R. (2007). A new method for assessing the effect of replication on DNA base composition asymmetry. *Mol. Biol. Evol.* 24, 2169–2179. doi: 10.1093/molbev/msm148

Nikolaou, C., and Almirantis, Y. (2005). A study on the correlation of nucleotide skews and the positioning of the origin of replication: different modes of replication in bacterial species. *Nucleic Acids Res.* 33, 6816–6822. doi: 10.1093/nar/gki988

Ohbayashi, R., Nakamachi, A., Hatakeyama, T. S., Watanabe, S., Kanesaki, Y., Chibazakura, T., et al. (2019). Coordination of polyploid chromosome replication with cell size and growth in a Cyanobacterium. *mBio* 10:e0510-19. doi: 10.1128/mBio.00510-19

Ohbayashi, R., Watanabe, S., Ehira, S., Kanesaki, Y., Chibazakura, T., and Yoshikawa, H. (2016). Diversification of DnaA dependency for DNA replication in cyanobacterial evolution. *ISME J.* 10, 1113–1121. doi: 10.1038/ismej.2015.194

Ran, L., Larsson, J., Vigil-Stenman, T., Nylander, J. A., Ininbergs, K., Zheng, W. W., et al. (2010). Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. *PLoS One* 5:e11486. doi: 10.1371/journal.pone.0011486

Rocha, E. P., and Danchin, A. (2003). Essentiality, not expressiveness, drives gene-strand bias in bacteria. *Nat. Genet.* 34, 377–338.

Rocha, E. P., Touchon, M., and Feil, E. J. (2006). Similar compositional biases are caused by very different mutational effects. *Genome Res.* 16, 1537–1547. doi: 10.1101/gr.5525106

Sargent, E. C., Hitchcock, A., Johansson, S. A., Langlois, R., Moore, C. M., Laroche, J., et al. (2016). Evidence for polyploidy in the globally important diazotroph *Trichodesmium*. *FEMS Microbiol. Lett.* 363:fnw244. doi: 10.1093/femsle/fnw244

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., et al. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad,. Sci. U.S.A.* 110, 1053–1058. doi: 10.1073/pnas.1217107110

Simon, R. D. (1977). Macromolecular composition of spores from the filamentous cyanobacterium *Anabaena cylindrica*. *J. Bacteriol.* 129:1154. doi: 10.1128/jb.129.2.1154-1155.1977

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Tanabe, A. S. (2011). Kakusan4 and aminosan: two programs for comparing nonpartitioned, proportional and separate models for combined molecular phylogenetic analyses of multilocus sequence data. *Mol. Ecol. Resour.* 11, 914–921. doi: 10.1111/j.1755-0998.2011.03021.x

Wang, J. D., and Levin, P. A. (2009). Metabolism, cell growth and the bacterial cell cycle. *Nat. Rev. Microbiol.* 7, 822–827. doi: 10.1038/nrmicro2202

Watanabe, S., Ohbayashi, R., Shiwa, Y., Noda, A., Kanesaki, Y., Chibazakura, T., et al. (2012). Light-dependent and asynchronous replication of cyanobacterial multi-copy chromosomes. *Mol. Microbiol.* 83, 856–865. doi: 10.1111/j.1365-2958.2012.07971.x

Worning, P., Jensen, L. J., Hallin, P. F., Stærfeldt, H. H., and Ussery, D. W. (2006). Origin of replication in circular prokaryotic chromosomes. *Environ. Microbiol.* 8, 353–361. doi: 10.1111/j.1462-2920.2005.00917.x

Yoshikawa, H., and Ogasawara, N. (1991). Structure and function of *DnaA* and the *DnaA*-box in eubacteria: evolutionary relationships of bacterial replication origins. *Mol. Microbiol.* 5, 2589–2597. doi: 10.1111/j.1365-2958.1991.tb01967.x

Zerulla, K., Ludt, K., and Soppa, J. (2016). The ploidy level of *Synechocystis* sp. *PCC* 6803 is highly variable and is influenced by growth phase and by chemical and physical external parameters. *Microbiology* 162, 730–739. doi: 10.1099/mic.0.000264

Zheng, X.-Y., and O'Shea, E. K. (2017). Cyanobacteria maintain constant protein concentration despite genome copy-number variation. *Cell Rep.* 19, 497–504. doi: 10.1016/j.celrep.2017.03.067

Zhou, Y., Chen, W.-L., Wang, L., and Zhang, C.-C. (2011). Identification of the oriC region and its influence on heterocyst development in the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120. *Microbiology* 157, 1910–1919. doi: 10.1099/mic.0.047241-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 28 August 2020
doi: 10.3389/fmicb.2020.582113

Bacterial Chromosome Replication and DNA Repair During the Stringent Response

Anurag Kumar Sinha†‡, Anders Løbner-Olesen and Leise Riber*†‡*

Department of Biology, University of Copenhagen, Copenhagen, Denmark

OPEN ACCESS

Edited by:
Feng Gao,
Tianjin University, China

Reviewed by:
Gregory Marczynski,
McGill University, Canada
Gert Bange,
University of Marburg, Germany

****Correspondence:***
Anurag Kumar Sinha
akusi@food.dtu.dk
Leise Riber
lriber@plen.ku.dk

†Present address:
Anurag Kumar Sinha,
National Food Institute, Technical University of Denmark, Lyngby, Denmark
Leise Riber,
Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *10 July 2020*
Accepted: *13 August 2020*
Published: *28 August 2020*

Citation:
Sinha AK, Løbner-Olesen A and Riber L (2020) Bacterial Chromosome Replication and DNA Repair During the Stringent Response. Front. Microbiol. 11:582113. doi: 10.3389/fmicb.2020.582113

The stringent response regulates bacterial growth rate and is important for cell survival under changing environmental conditions. The effect of the stringent response is pleiotropic, affecting almost all biological processes in the cell including transcriptional downregulation of genes involved in stable RNA synthesis, DNA replication, and metabolic pathways, as well as the upregulation of stress-related genes. In this Review, we discuss how the stringent response affects chromosome replication and DNA repair activities in bacteria. Importantly, we address how accumulation of (p)ppGpp during the stringent response shuts down chromosome replication using highly different strategies in the evolutionary distant Gram-negative *Escherichia coli* and Gram-positive *Bacillus subtilis.* Interestingly, (p)ppGpp-mediated replication inhibition occurs downstream of the origin in *B. subtilis*, whereas replication inhibition in *E. coli* takes place at the initiation level, suggesting that stringent cell cycle arrest acts at different phases of the replication cycle between *E. coli* and *B. subtilis*. Furthermore, we address the role of (p)ppGpp in facilitating DNA repair activities and cell survival during exposure to UV and other DNA damaging agents. In particular, (p)ppGpp seems to stimulate the efficiency of nucleotide excision repair (NER)-dependent repair of DNA lesions. Finally, we discuss whether (p)ppGpp-mediated cell survival during DNA damage is related to the ability of (p)ppGpp accumulation to inhibit chromosome replication.

Keywords: (p)ppGpp, DNA replication, DNA repair, stringent response, genome stability, *Escherichia coli*, *Bacillus subtilis*

INTRODUCTION

Bacteria respond to a variety of changing environmental conditions by inducing the stringent response. Known inducers of the stringent response include nutrient limitations such as amino acids, fatty acids, carbon and nitrogen starvation, and other stresses such as high temperature and low pH (Gallant et al., 1977; Gentry and Cashel, 1996; Wells and Gaynor, 2006; Winther et al., 2018; Sinha et al., 2019; Schafer et al., 2020). The hallmark of stringent response is the accumulation of guanosine tetra- or pentaphosphate, ppGpp and pppGpp, respectively [collectively called (p)ppGpp or alarmone], which leads to reprogramming of cell physiology facilitating cell survival under stress (Potrykus and Cashel, 2008; Hauryliuk et al., 2015). Importantly, (p)ppGpp plays a role in antibiotic tolerance and is essential for virulence in pathogenic bacteria (Dalebroux et al., 2010; Hauryliuk et al., 2015). Additionally, (p)ppGpp regulates bacterial growth rates even in the absence of external environmental stress (Potrykus et al., 2011).

Alarmones are synthesized and hydrolyzed by the long RelA/SpoT Homolog (RSH) protein superfamily. In the Gram-negative γ–proteobacterium *Escherichia coli*, two paralogous enzymes modulate (p)ppGpp levels; monofunctional RelA, which has only synthetase activity, and bifunctional SpoT, which has both synthetase and hydrolase activities. In the spore-forming Gram-positive bacterium, *Bacillus subtilis*, (p)ppGpp levels are metabolized by one long RSH superfamily protein Rel and two small alarmone synthetases (SASs) called RelP and RelQ (Liu et al., 2015; reviewed in Ronneau and Hallez, 2019). Accumulation of (p)ppGpp rapidly alters the levels of a wide range of gene transcripts and metabolites to allow cell survival and adaptation to new growth conditions (Eymann et al., 2002; Traxler et al., 2008). The major changes involve transcriptional down-regulation of genes involved in stable RNA (rRNA and tRNA) synthesis, DNA replication, and metabolic pathways, whereas genes engaged in stress and amino-acid biosynthesis are activated (Sanchez-Vazquez et al., 2019; Gummesson et al., 2020). In *E. coli*, (p)ppGpp directly binds two sites on RNA polymerase (RNAP) to allosterically alter its binding to- and efficiency at different gene promoters, which results in genome-wide transcriptional reprogramming. (p)ppGpp binding to RNAP and the consequent RNAP-driven transcriptional response is potentiated by another small RNAP binding protein, DksA (reviewed in Gourse et al., 2018). In *B. subtilis,* RNAP lacks critical (p)ppGpp binding sites and no DskA homologs have been identified. As a consequence, (p)ppGpp does not directly target *B. subtilis* RNAP. Instead (p)ppGpp synthesis strongly depletes the pool of available GTP, which leads to an indirect inhibition of stable RNA promoter activity since GTP is used as start nucleotide for most of the stable RNAs (Krasny and Gourse, 2004; Gourse et al., 2018; Sanchez-Vazquez et al., 2019). Importantly, apart from transcriptional responses, (p)ppGpp directly targets many other proteins to affect metabolic processes such as nucleotide metabolism and biosynthetic pathways (Zhang et al., 2018, 2019; Wang et al., 2019).

Here, we discuss how the stringent response affects chromosome replication, DNA damage and repair activities, focusing mainly on recent studies done in the evolutionarily distant *E. coli* and *B. subtilis*.

ROLE OF THE STRINGENT RESPONSE IN CHROMOSOME REPLICATION

In *E. coli,* chromosome replication initiates at a single origin of replication, *oriC*, which contains an AT-rich region and multiple binding-sites for the initiator protein, DnaA (Leonard and Mechali, 2013). DnaA belongs to the family of AAA + proteins and binds ATP and ADP with similar affinity (Sekimizu et al., 1987), of which only the ATP-bound form, $DnaA^{ATP}$, is required for oligomerization at *oriC,* and hence active for initiation (reviewed in Skarstad and Katayama, 2013; Riber et al., 2016). Origin unwinding leads to loading of DNA helicase, DnaB, onto single-stranded DNA (ssDNA) by the helicase loader, DnaC, followed by recruitment of primase, DnaG, as well as assembly of two replisomes to direct replication bidirectionally, until the replication forks meet and terminate at the terminus region, opposite to *oriC* (Kornberg and Baker, 1992). In *B. subtilis*, chromosome replication is mediated by the same overall steps, but the bipartite replication origin, containing two DnaA-box clusters separated by the *dnaA* gene (Moriya et al., 1992), is structurally different as compared to the continuous replication origin of *E. coli*. Also, assembly of the helicase, DnaC, onto ssDNA by the helicase loader, DnaI, occurs via a different mechanism known as "ring assembly" (Soultanas, 2012), but the following recruitment of DnaG primase and assembly of the replication elongation machinery is largely similar to that of *E. coli* (reviewed by Jameson and Wilkinson, 2017).

Highly different strategies have been adopted for (p)ppGpp-mediated chromosome replication inhibition in *E. coli* and *B. subtilis*. It is widely accepted that replication arrest in *B. subtilis* occurs downstream from the origin (i.e., on the elongation level), whereas replication inhibition in *E. coli* occurs at the initiation level, suggesting that stringent cell cycle arrest points differ between *E. coli* and *B. subtilis* (Levine et al., 1991).

(p)ppGpp-Mediated Inhibition of Initiation of Chromosome Replication

High levels of (p)ppGpp inhibit chromosome replication initiation in *E. coli* (Levine et al., 1991; Schreiber et al., 1995; Ferullo and Lovett, 2008; Riber and Lobner-Olesen, 2020), but the exact mechanism responsible for this inhibition has been somewhat unclear. However, several recent papers have made crucial discoveries adding valuable insight into this area of research.

Previously, the transcriptional activity of both *dnaA* operon promoters was reported to be stringently controlled (Chiaramello and Zyskind, 1990; Zyskind and Smith, 1992), suggesting that reduced *dnaA* gene transcription, and hence lowered *de novo* DnaA protein synthesis, could explain the initiation arrest observed in the presence of elevated (p)ppGpp levels. This was supported by a recent study, reporting that continued DnaA synthesis, expressed from a (p)ppGpp-insensitive T7 RNAP-dependent promoter, allowed for replication initiation during (p)ppGpp accumulation (Riber and Lobner-Olesen, 2020). Additionally, it was reported that polyphosphate during the stringent response activates Lon protease to degrade $DnaA^{ADP}$. As several regulatory systems work in concert to convert $DnaA^{ATP}$ into $DnaA^{ADP}$ (Katayama et al., 1998; Kato and Katayama, 2001; Kasho and Katayama, 2013), this indirectly lowers the amount of active $DnaA^{ATP}$, causing replication initiation to cease (Gross and Konieczny, 2020). However, degradation of DnaA has been reported only for *Caulobacter crescentus*, and not for *E. coli* (Gorbatyuk and Marczynski, 2005; Katayama et al., 2010). Also, recent data give no indication of DnaA degradation during (p)ppGpp accumulation (Riber and Lobner-Olesen, 2020).

Interestingly, several studies address the importance of DnaA activity, i.e., the $DnaA^{ATP}$-to-$DnaA^{ADP}$ ratio, during (p)ppGpp accumulation. Continuous *de novo* DnaA synthesis was found to allow for new rounds of replication initiation during (p)ppGpp

accumulation (Riber and Lobner-Olesen, 2020). As the level of ATP is more abundant than ADP in the cell (Petersen and Møller, 2000), and because DnaA binds these nucleotides with similar affinity (Sekimizu et al., 1987), *de novo* synthesized DnaA will be mainly ATP-bound, which ensures that the pool of $DnaA^{ATP}$ is continuously being replenished. Thus, while overall cell growth ceases due to (p)ppGpp accumulation $DnaA^{ATP}$ continues to increase due to *de novo* synthesis, which in turn allows for continued replication initiation during high levels of (p)ppGpp. In contrast, overproduction of DnaA during otherwise normal cell growth does not notably increase the $DnaA^{ATP}$ level (Flatten et al., 2015). Following induction of (p)ppGpp in such cells, transcription of *dnaA* will be repressed, which results in insufficient accumulation of active $DnaA^{ATP}$ to sustain further initiations (Kraemer et al., 2019).

Altogether, these observations suggest that (p)ppGpp-mediated replication initiation inhibition occurs through prevention of *de novo* DnaA synthesis, which lowers both the amount and activity (i.e., ATP-bound status) of DnaA. In agreement with this, (p)ppGpp fails to arrest replication initiation in cells where a hyperactive DnaA protein, mimicking ATP-bound DnaA, is overproduced (Kraemer et al., 2019).

Limitation of DnaA does, however, not seem to be the sole mechanism responsible of (p)ppGpp-mediated replication initiation inhibition. Recent studies emphasize lack of transcriptional activation of *oriC* to explain the negative effect of (p)ppGpp on initiation. Here, (p)ppGpp-driven reduction in transcriptional activity of promoters located close to *oriC*, presumably preventing introduction of negative supercoils in the wake of the migrating RNA polymerase complex, was suggested to cause less transcriptional activation of the origin, hence inhibiting initiation (Kraemer et al., 2019). Also, DNA gyrase (*gyrA*) and topoisomerase IV (*parC*) expression was found to be inhibited by high levels of (p)ppGpp, and the negative superhelicity of *oriC* was suggested to be lowered, despite not actually being measured (Fernandez-Coll et al., 2020).

Both *mioC* and *gidA* promoters, located adjacent to *oriC*, can be deleted without measurable effects (Lobner-Olesen and Boye, 1992; Bates et al., 1997; Lies et al., 2015), showing that they are dispensable for replication initiation during normal growth. However, when *oriC* becomes sufficiently impaired for initiation, such as when DnaA box R4 is deleted, transcription from these promoters becomes important (Bates et al., 1997). This is supported by the initiation kinetics of rifampicin and chloramphenicol. As rifampicin inhibits transcription initiation (Hartmann et al., 1967) rifampicin-treated cells will gradually stop to accumulate DnaA, but translation will continue as long as intact *dnaA* mRNA is present. On the other hand, chloramphenicol treatment will immediately block DnaA translation (Vazquez, 1979). Yet, chloramphenicol did not inhibit initiation as fast as rifampicin (Lark, 1972; Messer, 1972; Riber and Lobner-Olesen, 2020). As transcription is still on-going in chloramphenicol treated cells, this supports the ability of transcriptional activation of *oriC* to allow for extra initiations during suboptimal, e.g., DnaA limiting, conditions.

In conclusion, failure to *de novo* synthesize DnaA (i.e., reduced *dnaA* transcription) and to replenish the $DnaA^{ATP}$ pool along with lowered transcriptional activation of *oriC* (i.e., reduced *gidA/mioC* and/or *gyrA/parC* transcription) contribute in arresting replication initiation during (p)ppGpp accumulation in *E. coli* (**Figure 1A**; left). However, it is difficult to quantitate the exact contribution from each of those mechanisms.

(p)ppGpp-Mediated Inhibition of Elongation of Chromosome Replication

In contrast to *E. coli*, substantial replication occurs at the *B. subtilis* origin following induction of the stringent response. Also, regulation of chromosome replication initiation was shown to be independent of (p)ppGpp accumulation in *B. subtilis* (Levine et al., 1991; Murray and Koh, 2014). This indicates that (p)ppGpp might not regulate the synthesis of replication initiation proteins and/or transcriptional activation of *oriC* in *B. subtilis*. The lack of RNAP-driven transcriptional reprogramming due to *B. subtilis* RNAP not being a direct target of (p)ppGpp partly supports the latter (**Figure 1B**; left). Replication was instead shown to be arrested at distinct termination sites located approximately 200 kb downstream on either side of *oriC* (Levine et al., 1991), suggesting (p)ppGpp-mediated inhibition of chromosome replication in *B. subtilis* to be regulated at the post-initiation level.

By using genomic microarrays to monitor the progression of replication forks in synchronized cell cultures of *B. subtilis*, it was later revealed that starvation-induced replication arrest occurred throughout the chromosome, irrespective of the location of the replication forks. A direct (p)ppGpp-mediated inhibition of DNA primase (DnaG) activity, known to affect replication fork progression (Wu et al., 1992; Lee et al., 2006), was proposed to underlie the observed replication elongation arrest (Wang et al., 2007). This inhibition was found to be dose-dependent, suggesting that the severity of stress (i.e., concentration of (p)ppGpp) is tightly coupled to an equivalent reduction in replication progression rate, thus providing a tunable stress response (Wang et al., 2007; Denapoli et al., 2013). Interestingly, replication forks arrested in the presence of high levels of (p)ppGpp did not recruit the SOS response protein RecA, indicating that stalled forks were not disrupted, but reversibly halted with the ability to restart replication upon nutrient availability (Wang et al., 2007). These observations support that (p)ppGpp-mediated primase inhibition serves to maintain genome integrity during periods of stress.

Another factor that might contribute to the strong (p)ppGpp inhibition of progressing replication forks in *B. subtilis* is the equivalent decrease in the cellular pool of GTP available for continued DNA strand extension. This decrease is caused by increased consumption of GTP during (p)ppGpp biosynthesis, and by a direct inhibition of the activity of inosine monophosphate (IMP) dehydrogenase that catalyzes an early step in GTP biosynthesis (Lopez et al., 1981; **Figure 1B**; right).

(p)ppGpp binds and inhibits the *E. coli* DnaG primase *in vitro* (Maciag et al., 2010; Rymer et al., 2012). To date, no other replication proteins in *E. coli*, including DnaA, have been reported as direct targets for (p)ppGpp (Zhang et al., 2018; Wang et al., 2019). Obviously, this finding contradicts decades of

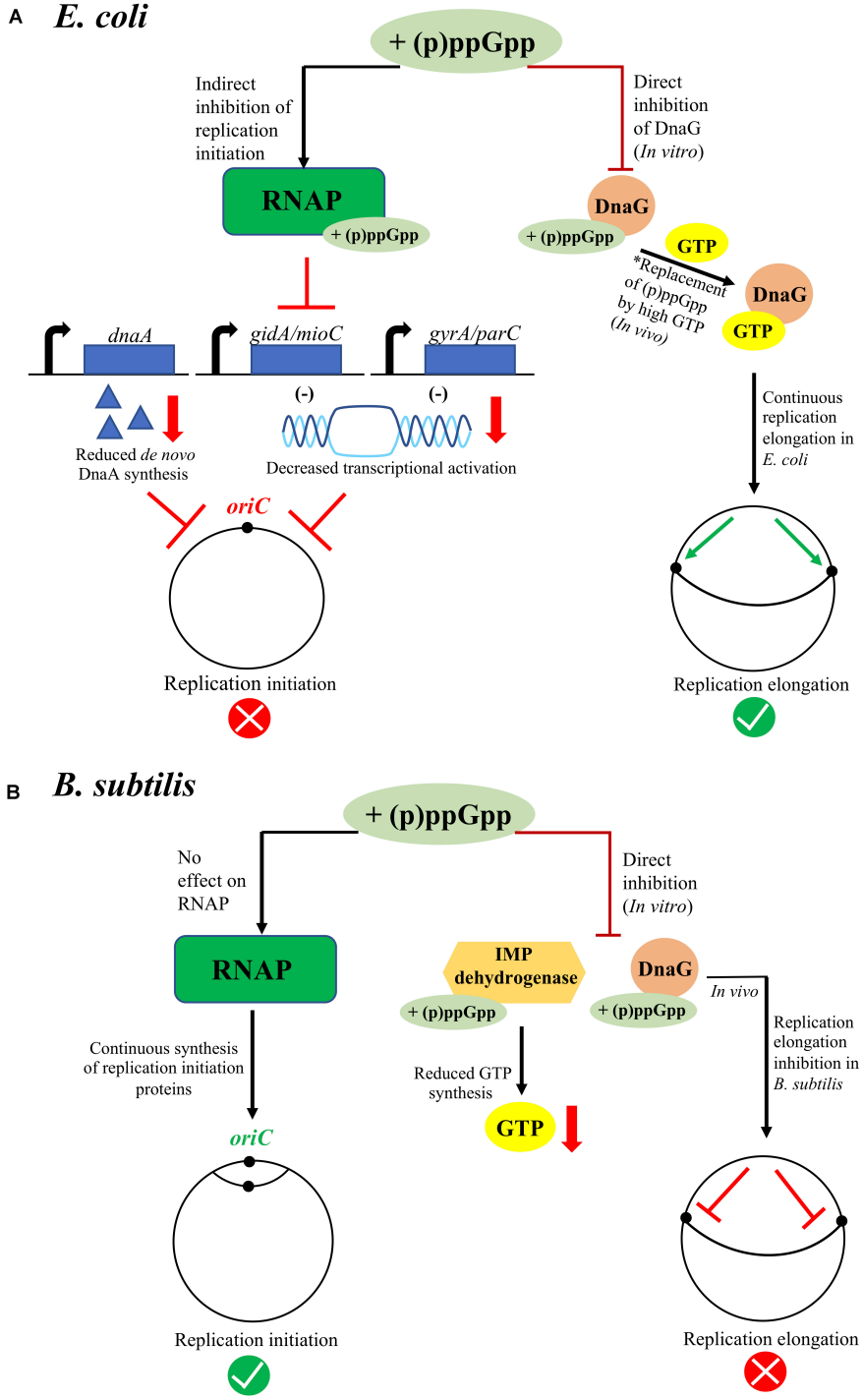


FIGURE 1 | Overview of (p)ppGpp-meditated inhibition of chromosome replication in *E. coli* **(A)** and *B. subtilis* **(B)**. In *E. coli* **(A)** replication inhibition occurs at the initiation level during (p)ppGpp accumulation. Here, (p)ppGpp binds the RNA Polymerase (RNAP), which indirectly affects the global gene expression profile through RNAP-driven transcriptional reprogramming. Downregulated gene transcripts include *dnaA, gidA, mioC, gyrA,* and *parC*, leading to lack of *de novo* DnaA synthesis and possibly lowered transcriptional activation of *oriC,* which all together contribute in arresting replication initiation during (p)ppGpp accumulation. Also, (p)ppGpp binds DnaG primase *in vitro*, but replication elongation remains unaffected *in vivo*. As GTP levels are not significantly reduced in *E. coli* during (p)ppGpp accumulation, and since GTP also binds DnaG, we hypothesize that GTP might outcompete (p)ppGpp in binding DnaG *in vivo* (this hypothesis is marked as *). In *B. subtilis* **(B)** replication inhibition occurs at the elongation level during (p)ppGpp accumulation. Here, (p)ppGpp binds IMP dehydrogenase, lowering the pool of available GTP, as well as DnaG. The significantly reduced level of GTP leads to DnaG being susceptible to strongly binding (p)ppGpp *in vivo*. Substantial replication occurs at the *B. subtilis* origin during (p)ppGpp accumulation, possibly because (p)ppGpp does not directly bind RNAP, excluding any RNAP-driven transcriptional reprogramming, or any replication initiation proteins.

research stating that ongoing rounds of replication are continued until completion following induction of the stringent response in *E. coli*, proposing that DNA replication elongation is not arrested during (p)ppGpp accumulation *in vivo* (Schreiber et al., 1995; Ferullo and Lovett, 2008; Kraemer et al., 2019; Riber and Lobner-Olesen, 2020). DeNapoli et al. did quantify genome-wide replication fork progression in *E. coli* and revealed that the replication elongation rate was modestly reduced by (p)ppGpp induction, but possibly the response was restricted to acute stress conditions (Denapoli et al., 2013).

Factors preventing binding of (p)ppGpp to DnaG, or the competing action between RNAP and DnaG in binding (p)ppGpp, were suggested to explain the lack of effect on DnaG activity *in vivo* (Maciag et al., 2010). Indeed, (p)ppGpp was found to bind DnaG at partially overlapping sites with nucleotides and inhibit primase activity in a GTP-concentration dependent manner (Rymer et al., 2012). As GTP levels are not reduced by more than 50% in *E. coli* during the stringent response (Varik et al., 2017), whereas *B. subtilis* experiences a significant drop in GTP concomitant with (p)ppGpp accumulation (Ochi et al., 1982), this supports a stronger (p)ppGpp-mediated binding to- and inhibition of DnaG in *B. subtilis*, hence leading to a more potent inhibition of replication elongation as compared to *E. coli* (**Figures 1A,B**; right).

ROLE OF THE STRINGENT RESPONSE IN DNA DAMAGE AND REPAIR

Bacterial genomic integrity is often threatened by DNA damage induced either by natural fork breakage, fork stalling, replication-transcription collision, or by external threats such as radiation and DNA modifying drugs (Kuzminov, 1999). Faithful damage repair orchestrated by DNA repair proteins is essential to maintain genomic integrity, chromosomal replication and cell viability. Accordingly, mutants lacking repair proteins are sensitive to DNA damaging agents and are less viable (Van Houten, 1990; Kuzminov, 1999; Sinha et al., 2020). Since (p)ppGpp binding to RNAP in *E. coli* destabilizes the open promoter complexes, it is expected to modulate replication-transcription collision and to play a role in maintaining genomic integrity.

The observation that loss of both RelA and SpoT ($ppGpp^0$ strain), i.e., inability to synthesize (p)ppGpp, enhanced UV sensitivity of an *E. coli ruvAB* mutant, suggested a possible role of (p)ppGpp in facilitating DNA repair (McGlynn and Lloyd, 2000). RuvAB along with RuvC play a role in branch migration and resolution of Holliday junctions, formed during RecBCD-RecA-mediated DNA double-strand break (DSB) repair and RecFOR-RecA-mediated gap repair (Kuzminov, 1999; Sinha et al., 2020). Interestingly, a slight increase in the basal level of (p)ppGpp by using the *spoT1* allele, having reduced (p)ppGpp hydrolytic activity, improved UV survival of the *ruvAB* mutant (McGlynn and Lloyd, 2000). Thus, high (p)ppGpp increases/promotes viability, whereas no (p)ppGpp increases UV sensitivity of the *ruvAB* mutant. The $ppGpp^0$ strain alone was also found to be UV sensitive (McGlynn and Lloyd, 2000).

The $ppGpp^0$ strain displays an amino acid auxotrophy phenotype and accumulates suppressor mutations (known as "stringent mutants") that allow cells to grow in minimal medium lacking amino acids. These suppressor mutations occur in RNAP subunits encoded by *rpoB* and *rpoC* (Zhou and Jin, 1998; McGlynn and Lloyd, 2000), and were shown to destabilize the transcriptional complex in a manner similar to (p)ppGpp binding to RNAP (Trautinger et al., 2005). Remarkably, some of these suppressor mutations (denoted *rpo**) significantly improved survival of the Δ*relA* Δ*spoT* Δ*ruvAB* strain after UV treatment (McGlynn and Lloyd, 2000).

Thus, it was proposed that (p)ppGpp/*rpo**-mediated destabilization of transcriptional complexes reduces the occurrence of stalled RNAP on DNA, hence allowing free space for efficient excision repair of UV-induced DNA lesions and for simultaneous facilitation of replication fork progression by avoiding replication-transcription conflicts (McGlynn and Lloyd, 2000; Trautinger and Lloyd, 2002; Trautinger et al., 2005). Additionally, it was shown that (p)ppGpp-mediated suppression of *ruvAB* mutant UV sensitivity is complex and requires RecA, RecG, and PriA, but not RecBCD, and was proposed to involve replication fork stalling, regression and restart (McGlynn and Lloyd, 2000). Since replication fork stalling, regression and restart are the major reactions following UV irradiation in *E. coli* cells (Khan and Kuzminov, 2012), the most plausible explanation for the UV resistance phenotype of *spoT1 ruvAB* (or *rpo* ruvAB*) cells would be destabilization of the RNAP array allowing replication forks to directly encounter DNA lesions followed by an active fork regression and lesion bypass, instead of fork breakage, to facilitate replication restart (Trautinger et al., 2005; **Figure 2**).

In contrast to UV, high (p)ppGpp (or *rpo**) cannot suppress sensitivity of the Δ*ruvAB* strain against exposure to mitomycin C (MMC) or γ rays (McGlynn and Lloyd, 2000). It should be noted that DNA lesions generated by both UV and MMC are removed/repaired by nucleotide excision repair (NER) (Van Houten, 1990). However, MMC treatment generates inter-strand crosslinks that most often get converted into DSBs, whereas UV treatment induces intra-strand pyrimidine dimers with generation of DSBs being primarily dependent on replication fork stalling at the lesion site (Khan and Kuzminov, 2012). These observations exclude a direct role of (p)ppGpp in DSBs repair.

Transcription-Coupled DNA Repair (TCR)

Another study, corroborating the above finding, confirmed that *E. coli* $ppGpp^0$ cells were highly sensitive to UV radiation, 4-nitroquinoline-1-oxide (4NQO), and nitrofurazone (NFZ) (Kamarthapu et al., 2016). These agents induce formation of DNA adducts, which are mainly removed and repaired by NER pathways (Ikenaga et al., 1975; Ona et al., 2009). Remarkably, wild-type cells rapidly accumulated a 20-fold increase in (p)ppGpp when treated with 4NQO or NFZ, suggesting that DNA lesions induce (p)ppGpp synthesis. However, the mechanism of (p)ppGpp synthesis during these treatments remains to be determined (Kamarthapu et al., 2016).

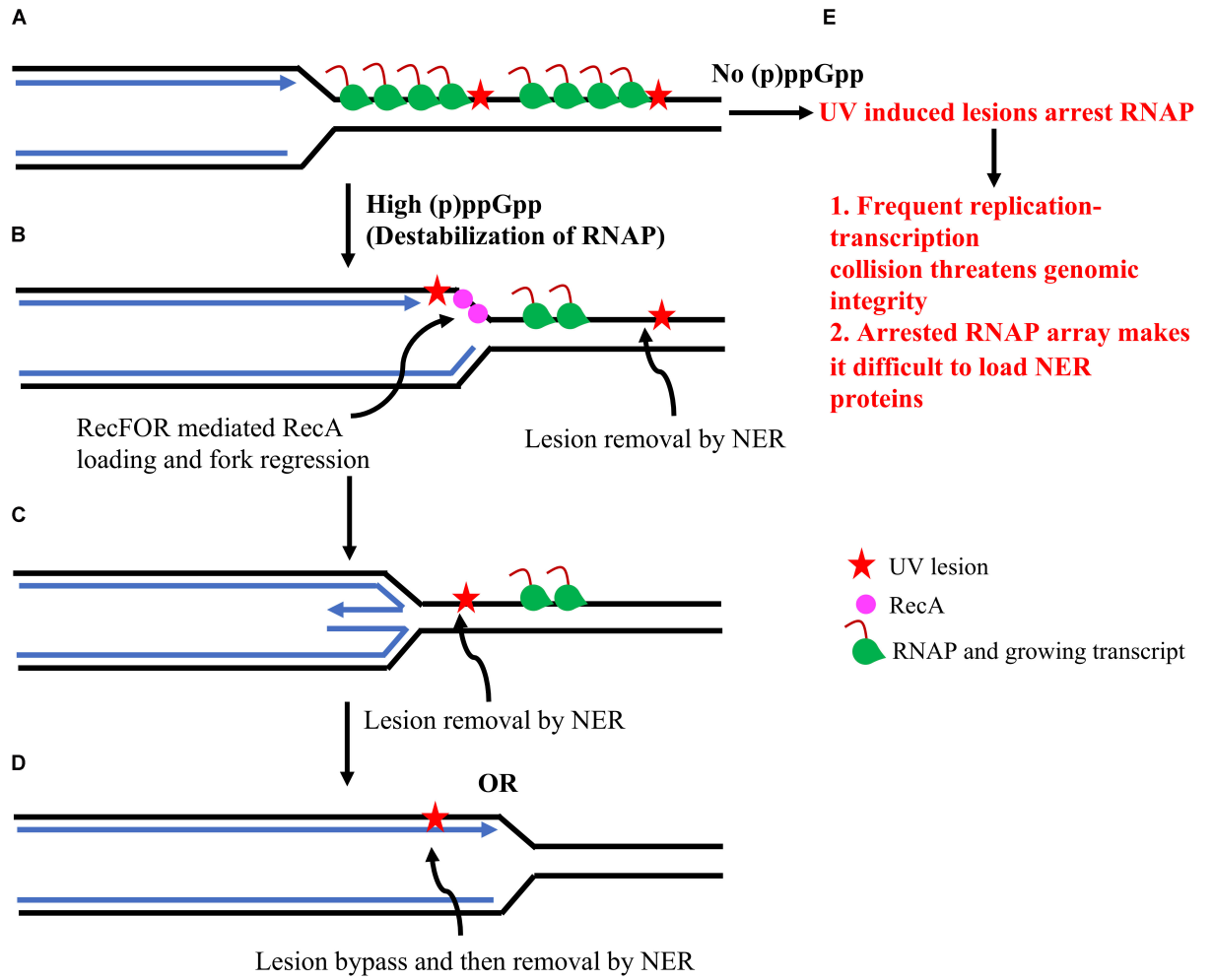


FIGURE 2 | Overview of (p)ppGpp-meditated repair of UV induced DNA damage in *E. coli*. The UV induced DNA lesions arrest RNAP and halt transcription progression **(A)**. This can lead to frequent replication-transcription collision. In **(A)**, only co-directional collision has been shown but there is an equal possibility for head-on collision and both threaten genomic integrity. This scenario will probably be escalated in absence of (p)ppGpp since the RNAP array will be stably arrested for a long time in absence of (p)ppGpp **(A–E)**. Whereas, (p)ppGpp binding to RNAP will destabilize it and remove it from the DNA template. Removal of RNAP will help in two ways: 1. It will create space to load NER proteins and remove/repair DNA lesions. 2. It will help the replication fork to progress toward DNA lesions **(B)**. Arrested replication forks can get reversed with the help of RecFOR mediated RecA loading and fork regression **(C)**. DNA synthesis and resetting of the replication fork will help in lesion bypass **(D)**. DNA lesions can be removed and repaired by NER pathways either at the **(C,D)** step. This model is adapted from Trautinger et al. (2005).

TCR is defined by an active transcription-dependent increase in excision repair of lesions on the transcribed DNA strand in comparison to the non-transcribed strand (Mellon and Hanawalt, 1989). Two factors, Mfd and UvrD, promote TCR by two different pathways: by pushing RNAP forward of the DNA lesion and by promoting RNAP backtracking, respectively, followed by recruitment of NER proteins, such as UvrAB at the lesion site (Mellon and Hanawalt, 1989; Kamarthapu and Nudler, 2015). Interestingly, the preference for repairing the transcribed strand rather than the non-transcribed strand was abolished in $ppGpp^0$ cells suggesting that (p)ppGpp is crucial for TCR. Since the sensitivity of $ppGpp^0$ cells to UV, 4NQO or NFZ was epistatic to *uvrD* mutant sensitivity, it was proposed that (p)ppGpp potentiates the pro-backtracking activity of UvrD (Kamarthapu et al., 2016). The role of (p)ppGpp in facilitating TCR can also occur independent of UvrD either by promoting RNAP backtracking by destabilizing and removing RNAP complexes from tightly packed arrays at the highly transcribed ribosomal genes, thus creating space for backtracking, or by reducing the number of ribosomes trailing RNAP to make space for backtracking (Rasouly et al., 2017). However, extensive backtracked RNAP might increase the risk of replication-transcription collision and has the capacity to induce DSBs and genomic instability (Dutta et al., 2011). The conundrum is perhaps resolved by (p)ppGpp-mediated inhibition of replication initiation, thus minimizing the frequency of replication-transcription collisions when RNAP backtracking is needed to repair genotoxic lesions on DNA.

In *B. subtilis,* the SMC-ScpAB complex is important for chromosome condensation and segregation, and Δ*smc* mutants exhibit pleiotropic phenotypes including defects in chromosome condensation, segregation, DNA repair and viability at high temperature. Upregulation of the stringent response has been shown to suppress chromosome segregation defects, hypersensitivity to gyrase inhibitors and restore viability of Δ*smc* mutants (Benoist et al., 2015). Since the stringent response slows down replication elongation in *B. subtilis*, it might be possible that slow replication allows chromosome segregation to occur even in the absence of the SMC-ScpAB complex. This hypothesis finds support as Δ*smc* mutant cells grow well in minimal medium (i.e., slow growth conditions) as compared to no growth in rich medium (i.e., fast growth conditions) at 37°C (Benoist et al., 2015). Similar studies for the role of the stringent response in chromosome segregation mutant cells of *E. coli* have not been reported.

CONCLUDING REMARKS

Based on the highlights presented throughout this review, the stringent response has clearly proven to affect both bacterial chromosome replication and DNA repair activities. However, whereas (p)ppGpp accumulation negatively affects replication initiation and replication elongation in *E. coli* and *B. subtilis*, respectively, the effect of (p)ppGpp-mediated modulation of DNA repair activities seems positive. Indeed, the absence of (p)ppGpp makes *E. coli* cells sensitive to UV and other DNA damaging agents, and studies suggest a role of (p)ppGpp in enhancing the efficiency of NER-dependent repair of DNA lesions, most likely by destabilizing RNAP complexes and making space for recruitment of NER proteins. Interestingly, these observations might be coupled to (p)ppGpp-mediated replication inhibition, which prevents replication-transcription collisions and/or reduces the frequency of replication forks meeting the UV lesions, thus assisting efficient NER-mediated repair. This intriguing hypothesis, connecting the negative effect of (p)ppGpp on replication to (p)ppGpp-driven stimulation of DNA repair activity, can easily be tested by using a system where (p)ppGpp-dependent replication inhibition is abrogated as recently described (Riber and Lobner-Olesen, 2020).

AUTHOR CONTRIBUTIONS

AS, AL-O, and LR wrote the manuscript. AS and LR designed and prepared the figures. All authors contributed to the article and approved the submitted version.

FUNDING

For this research, AS was funded by the Center for Bacterial Stress Response and Persistence (BASP) supported by grants from the Novo Nordisk Foundation and the Danish National Research Foundation (DNRF120). AL-O was funded by grants from the Danish National Research Foundation (DNRF120) and from the Novo Nordisk Foundation through the Challenge Center for Peptide-Based Antibiotics NNF16OC0021700 (Cepan). LR was funded by a Lundbeck Foundation Experiment 2019 Grant (R324-2019-2001) from the Lundbeck Foundation.

REFERENCES

Bates, D. B., Boye, E., Asai, T., and Kogoma, T. (1997). The absence of effect of gid or mioC transcription on the initiation of chromosomal replication in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 12497–12502. doi: 10.1073/pnas.94.23.12497

Benoist, C., Guérin, C., Noirot, P., and Dervyn, E. (2015). Constitutive stringent response restores viability of *Bacillus subtilis* lacking structural maintenance of chromosome protein. *PLoS One* 10:e0142308. doi: 10.1371/journal.pone.0142308

Chiaramello, A. E., and Zyskind, J. W. (1990). Coupling of DNA replication to growth rate in *Escherichia coli*: a possible role for guanosine tetraphosphate. *J. Bacteriol.* 172, 2013–2019. doi: 10.1128/jb.172.4.2013-2019.1990

Dalebroux, Z. D., Svensson, S. L., Gaynor, E. C., and Swanson, M. S. (2010). ppGpp conjures bacterial virulence. *Microbiol. Mol. Biol. Rev.* 74, 171–199. doi: 10.1128/mmbr.00046-09

Denapoli, J., Tehranchi, A. K., and Wang, J. D. (2013). Dose-dependent reduction of replication elongation rate by (p)ppGpp in *Escherichia coli* and *Bacillus subtilis*. *Mol. Microbiol.* 88, 93–104. doi: 10.1111/mmi.12172

Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M. E., and Nudler, E. (2011). Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell* 146, 533–543. doi: 10.1016/j.cell.2011.07.034

Eymann, C., Homuth, G., Scharf, C., and Hecker, M. (2002). *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. *J. Bacteriol.* 184, 2500–2520. doi: 10.1128/jb.184.9.2500-2520.2002

Fernandez-Coll, L., Maciag-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., et al. (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth Rates. *mBio* 11:e03223-19.

Ferullo, D. J., and Lovett, S. T. (2008). The stringent response and cell cycle arrest in *Escherichia coli*. *PLoS. Genet.* 4:e1000300. doi: 10.1371/journal.pgen.1000300

Flatten, I., Fossum-Raunehaug, S., Taipale, R., Martinsen, S., and Skarstad, K. (2015). The DnaA Protein Is Not the Limiting Factor for Initiation of Replication in *Escherichia coli*. *PLoS.Genet.* 11:e1005276. doi: 10.1371/journal.pgen.1005276

Gallant, J., Palmer, L., and Pao, C. C. (1977). Anomalous synthesis of ppGpp in growing cells. *Cell* 11, 181–185. doi: 10.1016/0092-8674(77)90329-4

Gentry, D. R., and Cashel, M. (1996). Mutational analysis of the *Escherichia coli* spoT gene identifies distinct but overlapping regions involved in ppGpp synthesis and degradation. *Mol. Microbiol.* 19, 1373–1384. doi: 10.1111/j.1365-2958.1996.tb02480.x

Gorbatyuk, B., and Marczynski, G. T. (2005). Regulated degradation of chromosome replication proteins DnaA and CtrA in *Caulobacter crescentus*. *Mol. Microbiol.* 55, 1233–1245. doi: 10.1111/j.1365-2958.2004.04459.x

Gourse, R. L., Chen, A. Y., Gopalkrishnan, S., Sanchez-Vazquez, P., Myers, A., and Ross, W. (2018). Transcriptional Responses to ppGpp and DksA. *Annu. Rev. Microbiol.* 72, 163–184. doi: 10.1146/annurev-micro-090817-062444

Gross, M. H., and Konieczny, I. (2020). Polyphosphate induces the proteolysis of ADP-bound fraction of initiator to inhibit DNA replication initiation upon stress in *Escherichia coli*. *Nucleic Acids Res.* 48, 5457–5466. doi: 10.1093/nar/gkaa217

Gummesson, B., Shah, S. A., Borum, A. S., Fessler, M., Mitarai, N., Sorensen, M. A., et al. (2020). Valine-induced isoleucine starvation in *Escherichia coli* K-12 studied by spike-in normalized RNA sequencing. *Front. Genet.* 11:144. doi: 10.3389/fgene.2020.00144

Hartmann, G., Honikel, K. O., Knusel, F., and Nuesch, J. (1967). The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochim. Biophys. Acta* 145, 843–844. doi: 10.1016/0005-2787(67)90147-5

Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* 13, 298–309. doi: 10.1038/nrmicro3448

Ikenaga, M., Ishii, Y., Tada, M., Kakunaga, T., and Takebe, H. (1975). Excision-repair of 4-nitroquinolin-1-oxide damage responsible for killing, mutation, and cancer. *Basic Life Sci.* 5b, 763–771. doi: 10.1007/978-1-4684-2898-8_54

Jameson, K. H., and Wilkinson, A. J. (2017). Control of Initiation of DNA Replication in *Bacillus subtilis* and *Escherichia coli*. *Genes* 8:22. doi: 10.3390/genes8010022

Kamarthapu, V., Epshtein, V., Benjamin, B., Proshkin, S., Mironov, A., Cashel, M., et al. (2016). ppGpp couples transcription to DNA repair in *E. coli*. *Science* 352, 993–996. doi: 10.1126/science.aad6945

Kamarthapu, V., and Nudler, E. (2015). Rethinking transcription coupled DNA repair. *Curr. Opin. Microbiol.* 24, 15–20. doi: 10.1016/j.mib.2014.12.005

Kasho, K., and 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. doi: 10.1073/pnas.1212070110

Katayama, T., Kubota, T., Kurokawa, K., Crooke, E., and 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. doi: 10.1016/s0092-8674(00)81222-2

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Kato, J., and Katayama, T. (2001). Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*. *EMBO J.* 20, 4253–4262. doi: 10.1093/emboj/20.15.4253

Khan, S. R., and Kuzminov, A. (2012). Replication forks stalled at ultraviolet lesions are rescued via RecA and RuvABC protein-catalyzed disintegration in *Escherichia coli*. *J. Biol. Chem.* 287, 6250–6265. doi: 10.1074/jbc.m111.322990

Kornberg, A., and Baker, T. A. (1992). *DNA Replication*. New York, NY: W.H.Freeman and company.

Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The stringent response inhibits DNA replication initiation in *E. coli* by modulating supercoiling of oriC. *mBio* 10:e01330-19.

Krasny, L., and Gourse, R. L. (2004). An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. *EMBO J.* 23, 4473–4483. doi: 10.1038/sj.emboj.7600423

Kuzminov, A. (1999). Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* 63, 751–813. doi: 10.1128/mmbr.63.4.751-813.1999

Lark, K. G. (1972). Evidence for direct involvement of RNA in the initiation of DNA replication in *E. coli* 15T. *J. Mol. Biol.* 64, 47–60. doi: 10.1016/0022-2836(72)90320-8

Lee, J. B., Hite, R. K., Hamdan, S. M., Xie, X. S., Richardson, C. C., and Van Oijen, A. M. (2006). DNA primase acts as a molecular brake in DNA replication. *Nature* 439, 621–624. doi: 10.1038/nature04317

Leonard, A. C., and Mechali, M. (2013). DNA replication origins. *Cold Spring Harb. Perspect. Biol.* 5:a010116.

Levine, A., Vannier, F., Dehbi, M., Henckes, G., and Séror, S. J. (1991). The stringent response blocks DNA replication outside the ori region in *Bacillus subtilis* and at the origin in *Escherichia coli*. *J. Mol. Biol.* 219, 605–613. doi: 10.1016/0022-2836(91)90657-r

Lies, M., Visser, B. J., Joshi, M. C., Magnan, D., and Bates, D. (2015). MioC and GidA proteins promote cell division in *E. coli*. *Front. Microbiol.* 6:516. doi: 10.3389/fmicb.2015.00516

Liu, K., Bittner, A. N., and Wang, J. D. (2015). Diversity in (p)ppGpp metabolism and effectors. *Curr. Opin. Microbiol.* 24, 72–79. doi: 10.1016/j.mib.2015.01.012

Lobner-Olesen, A., and Boye, E. (1992). Different effects of *mioC* transcription on initiation of chromosomal and minichromosomal replication in *Escherichia coli*. *Nucleic Acids Res.* 20, 3029–3036. doi: 10.1093/nar/20.12.3029

Lopez, J. M., Dromerick, A., and Freese, E. (1981). Response of guanosine 5'-triphosphate concentration to nutritional changes and its significance for *Bacillus subtilis* sporulation. *J. Bacteriol.* 146, 605–613. doi: 10.1128/jb.146.2.605-613.1981

Maciag, M., Kochanowska, M., Lyzen, R., Wegrzyn, G., and Szalewska-Palasz, A. (2010). ppGpp inhibits the activity of *Escherichia coli* DnaG primase. *Plasmid* 63, 61–67. doi: 10.1016/j.plasmid.2009.11.002

McGlynn, P., and Lloyd, R. G. (2000). Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. *Cell* 101, 35–45. doi: 10.1016/s0092-8674(00)80621-2

Mellon, I., and Hanawalt, P. C. (1989). Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 342, 95–98. doi: 10.1038/342095a0

Messer, W. (1972). Initiation of deoxyribonucleic acid replication in *Escherichia coli* B-r: chronology of events and transcriptional control of initiation. *J. Bacteriol.* 112, 7–12. doi: 10.1128/jb.112.1.7-12.1972

Moriya, S., Atlung, T., Hansen, F. G., Yoshikawa, H., and Ogasawara, N. (1992). Cloning of an autonomously replicating sequence (*ars*) from the *Bacillus subtilis* chromosome. *Mol. Microbiol.* 6, 309–315. doi: 10.1111/j.1365-2958.1992.tb01473.x

Murray, H., and Koh, A. (2014). Multiple regulatory systems coordinate DNA replication with cell growth in *Bacillus subtilis*. *PLoS Genet.* 10:e1004731. doi: 10.1371/journal.pgen.1004731

Ochi, K., Kandala, J., and Freese, E. (1982). Evidence that *Bacillus subtilis* sporulation induced by the stringent response is caused by the decrease in GTP or GDP. *J. Bacteriol.* 151, 1062–1065. doi: 10.1128/jb.151.2.1062-1065.1982

Ona, K. R., Courcelle, C. T., and Courcelle, J. (2009). Nucleotide excision repair is a predominant mechanism for processing nitrofurazone-induced DNA damage in *Escherichia coli*. *J. Bacteriol.* 191, 4959–4965. doi: 10.1128/jb.00495-09

Petersen, C., and Møller, L. B. (2000). Invariance of the nucleoside triphosphate pools of *Escherichia coli* with growth rate. *J. Biol. Chem.* 275, 3931–3935. doi: 10.1074/jbc.275.6.3931

Potrykus, K., and Cashel, M. (2008). (p)ppGpp: still magical? *Annu. Rev. Microbiol.* 62, 35–51. doi: 10.1146/annurev.micro.62.081307.162903

Potrykus, K., Murphy, H., Philippe, N., and Cashel, M. (2011). ppGpp is the major source of growth rate control in *E. coli*. *Environ. Microbiol.* 13, 563–575. doi: 10.1111/j.1462-2920.2010.02357.x

Rasouly, A., Pani, B., and Nudler, E. (2017). A magic spot in genome maintenance. *Trends Genet.* 33, 58–67. doi: 10.1016/j.tig.2016.11.002

Riber, L., Frimodt-Moller, J., Charbon, G., and Lobner-Olesen, A. (2016). Multiple DNA Binding Proteins Contribute to Timing of Chromosome Replication in *E. coli*. *Front. Mol. Biosci.* 3:29. doi: 10.3389/fmolb.2016.00029

Riber, L., and Lobner-Olesen, A. (2020). Inhibition of *Escherichia coli* chromosome replication by rifampicin treatment or during the stringent response is overcome by de novo DnaA protein synthesis. *Mol. Microbiol.* doi: 10.1111/mmi.14531 [Epub ahead of print].

Ronneau, S., and Hallez, R. (2019). Make and break the alarmone: regulation of (p)ppGpp synthetase/hydrolase enzymes in bacteria. *FEMS Microbiol. Rev.* 43, 389–400. doi: 10.1093/femsre/fuz009

Rymer, R. U., Solorio, F. A., Tehranchi, A. K., Chu, C., Corn, J. E., Keck, J. L., et al. (2012). Binding mechanism of metalNTP substrates and stringent-response alarmones to bacterial DnaG-type primases. *Structure* 20, 1478–1489. doi: 10.1016/j.str.2012.05.017

Sanchez-Vazquez, P., Dewey, C. N., Kitten, N., Ross, W., and Gourse, R. L. (2019). Genome-wide effects on *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8310–8319. doi: 10.1073/pnas.1819682116

Schafer, H., Beckert, B., Frese, C. K., Steinchen, W., Nuss, A. M., Beckstette, M., et al. (2020). The alarmones (p)ppGpp are part of the heat shock response of *Bacillus subtilis*. *PLoS Genet.* 16:e1008275. doi: 10.1371/journal.pgen.1008275

Schreiber, G., Ron, E. Z., and Glaser, G. (1995). ppGpp-mediated regulation of DNA replication and cell division in *Escherichia coli*. *Curr. Microbiol.* 30, 27–32. doi: 10.1007/bf00294520

Sekimizu, K., Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome. *Cell* 50, 259–265. doi: 10.1016/0092-8674(87)90221-2

Sinha, A. K., Possoz, C., and Leach, D. R. F. (2020). The roles of bacterial DNA double-strand break repair proteins in chromosomal DNA replication. *FEMS Microbiol. Rev.* 44, 351–368. doi: 10.1093/femsre/fuaa009

Sinha, A. K., Winther, K. S., Roghanian, M., and Gerdes, K. (2019). Fatty acid starvation activates RelA by depleting lysine precursor pyruvate. *Mol. Microbiol.* 112, 1339–1349. doi: 10.1111/mmi.14366

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Soultanas, P. (2012). Loading mechanisms of ring helicases at replication origins. *Mol. Microbiol.* 84, 6–16. doi: 10.1111/j.1365-2958.2012.08012.x

Trautinger, B. W., Jaktaji, R. P., Rusakova, E., and Lloyd, R. G. (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol. Cell* 19, 247–258. doi: 10.1016/j.molcel.2005.06.004

Trautinger, B. W., and Lloyd, R. G. (2002). Modulation of DNA repair by mutations flanking the DNA channel through RNA polymerase. *EMBO J.* 21, 6944–6953. doi: 10.1093/emboj/cdf654

Traxler, M. F., Summers, S. M., Nguyen, H. T., Zacharia, V. M., Hightower, G. A., Smith, J. T., et al. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol. Microbiol.* 68, 1128–1148. doi: 10.1111/j.1365-2958.2008.06229.x

Van Houten, B. (1990). Nucleotide excision repair in *Escherichia coli*. *Microbiol. Rev.* 54, 18–51. doi: 10.1128/mmbr.54.1.18-51.1990

Varik, V., Oliveira, S. R. A., Hauryliuk, V., and Tenson, T. (2017). HPLC-based quantification of bacterial housekeeping nucleotides and alarmone messengers ppGpp and pppGpp. *Sci. Rep.* 7:11022.

Vazquez, D. (1979). "Inhibitors of protein biosynthesis," in *Molecular Biology, Biochemistry and Biophysics*, eds A. Kleinzeller, G. F. Springer, and H. G. W. Wittman (Berlin: Springer Verlag), 108–112.

Wang, B., Dai, P., Ding, D., Del Rosario, A., Grant, R. A., Pentelute, B. L., et al. (2019). Affinity-based capture and identification of protein effectors of the growth regulator ppGpp. *Nat. Chem. Biol.* 15, 141–150. doi: 10.1038/s41589-018-0183-4

Wang, J. D., Sanders, G. M., and Grossman, A. D. (2007). Nutritional control of elongation of DNA replication by (p)ppGpp. *Cell* 128, 865–875. doi: 10.1016/j.cell.2006.12.043

Wells, D. H., and Gaynor, E. C. (2006). *Helicobacter* pylori initiates the stringent response upon nutrient and pH downshift. *J. Bacteriol.* 188, 3726–3729. doi: 10.1128/jb.188.10.3726-3729.2006

Winther, K. S., Roghanian, M., and Gerdes, K. (2018). Activation of the Stringent Response by Loading of RelA-tRNA Complexes at the Ribosomal A-Site. *Mol. Cell.* 70:95-105.e104.

Wu, C. A., Zechner, E. L., Reems, J. A., Mchenry, C. S., and Marians, K. J. (1992). Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. V. Primase action regulates the cycle of Okazaki fragment synthesis. *J. Biol. Chem.* 267, 4074–4083.

Zhang, Y., Zbornikova, E., Rejman, D., and Gerdes, K. (2018). Novel (p)ppGpp Binding and Metabolizing Proteins of *Escherichia coli*. *mBio* 9:e02188-17.

Zhang, Y. E., Baerentsen, R. L., Fuhrer, T., Sauer, U., Gerdes, K., and Brodersen, D. E. (2019). (p)ppGpp Regulates a Bacterial Nucleosidase by an Allosteric Two-Domain Switch. *Mol. Cell* 74, 1239-1249.e1234.

Zhou, Y. N., and Jin, D. J. (1998). The rpoB mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent". RNA polymerases in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 95, 2908–2913. doi: 10.1073/pnas.95.6.2908

Zyskind, J. W., and Smith, D. W. (1992). DNA replication, the bacterial cell cycle, and cell growth. *Cell* 69, 5–8. doi: 10.1016/0092-8674(92)90112-p

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 01 September 2020
doi: 10.3389/fmicb.2020.02002

Elucidating the Influence of Chromosomal Architecture on Transcriptional Regulation in Prokaryotes – Observing Strong Local Effects of Nucleoid Structure on Gene Regulation

Thøger Jensen Krogh[1], Andre Franke[2], Jakob Møller-Jensen[1] and Christoph Kaleta[3]*

[1]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark, [2]Institute of Clinical Molecular Biology (IKMB), Christian-Albrechts-University Kiel, Kiel, Germany, [3]Institute of Experimental Medicine, Christian-Albrechts-University Kiel, Kiel, Germany

OPEN ACCESS

Edited by:
Monika Glinkowska, University of Gdansk, Poland

Reviewed by:
Patrick Sobetzko, University of Marburg, Germany
William Nasser, UMR5240 Microbiologie, Adaptation et Pathogenie (MAP), France

***Correspondence:**
Thøger Jensen Krogh
tjkrogh@bmb.sdu.dk

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *15 April 2020*
Accepted: *29 July 2020*
Published: *01 September 2020*

Citation:
Krogh TJ ,Franke A, Møller-Jensen J and Kaleta C (2020) Elucidating the Influence of Chromosomal Architecture on Transcriptional Regulation in Prokaryotes – Observing Strong Local Effects of Nucleoid Structure on Gene Regulation. Front. Microbiol. 11:2002. doi: 10.3389/fmicb.2020.02002

Both intrinsic and extrinsic mechanisms regulating bacterial expression have been elucidated and described, however, such studies have mainly focused on local effects on the two-dimensional structure of the prokaryote genome while long-range as well as spatial interactions influencing gene expression are still only poorly understood. In this paper, we investigate the association between co-expression and distance between genes, using RNA-seq data at multiple growth phases in order to illuminate whether such conserved patterns are an indication of a gene regulatory mechanism relevant for prokaryotic cell proliferation, adaption, and evolution. We observe recurrent sinusoidal patterns in correlation of pairwise expression as function of genomic distance and rule out that these are caused by transcription-induced supercoiling gradients, gene clustering in operons, or association with regulatory transcription factors (TFs). By comparing spatial proximity for pairs of genomic bins with their correlation of pairwise expression, we further observe a high co-expression proportional with the spatial proximity. Based on these observations, we propose that the observed patterns are related to nucleoid structure as a product of transcriptional spilling, where genes actively influence transcription of spatially proximal genes through increases within shared local pools of RNA polymerases (RNAP), and actively spilling transcription onto neighboring genes.

Keywords: bacterial nucleoid, gene co-expression, chromosomal architecture, transcriptional spilling, predicting supercoils

INTRODUCTION

Coordinating the expression of functionally related genes in relation to environmental cues is pivotal for successful competition between species adapting to changing growth conditions (McAdams et al., 2004; Dillon and Dorman, 2010; Binder et al., 2016). This is evident through mutational experiments, and the optimization of gene expression is considered a fundamental driving force in evolution (McAdams et al., 2004; Seward and Kelly, 2018). Several mechanisms for optimization of co-expression are employed by prokaryotes to align and time expression

of genes and to reduce energy spent on regulation, including co-regulation through association of transcription factors (TFs) and sigma factors (SFs; McAdams et al., 2004; Djordjevic, 2013). Genes related within highly defined functional groups are furthermore often organized into co-transcribed and co-regulated groups termed operons. In addition, operons and genes within less defined functional groups are further organized into larger domains (~20 kbp) of coordinated expression termed supra-operons (Junier et al., 2016, 2018).

The bacterial nucleoid requires extensive but reversible compaction to fit inside the confinements of the cell (up to >1.000 fold in length), while retaining accessibility of the entire chromosome (Krogh et al., 2018). Forces compacting the nucleoid include the association of DNA with specific proteins, transcription-induced DNA supercoiling, macromolecular crowding and entropy-driven depletion attraction, whereas the primary hypothesized expanding force is the coupling of transcription, translation, and membrane-insertion of membrane associated proteins, collectively referred to as transertion (Woldringh et al., 1995; Zimmerman, 2002; Cabrera et al., 2009; Mondal et al., 2011; Bakshi et al., 2015). Nucleoid compaction requires tight control to avoid detrimental effects on the cell during growth and division (Dillon and Dorman, 2010). In *Escherichia coli* the nucleoid is condensed during rapid exponential growth, with RNA-polymerases (RNAP) mainly located at a few highly active spatial loci (Cabrera and Jin, 2003; Cabrera et al., 2009). During entry into stationary growth phase the nucleoid of *E. coli* decondenses to cover most of the cell, with RNAP located all over the nucleoid (Cabrera and Jin, 2003). Finally, during late stationary growth, the nucleoid is highly condensed into a crystal-like structure together with the nucleoid associated protein (NAP) Dps (Frenkiel-Krispin et al., 2004; Kim et al., 2004). Based on DNA-DNA interaction studies, the *E. coli* nucleoid has been found to be divisible into four structured and two non-structured macrodomains (Valens et al., 2004). The number of DNA loops resulting from nucleoid compaction has been observed to be highly variable, between 40 and 500, with varying sizes (Postow et al., 2004; Stein et al., 2005; Dillon and Dorman, 2010). In conclusion, prokaryotes appear generally to have highly structured yet dynamically regulated nucleoids with distinct and relatively stable macrodomain conformations during optimal growth. These conformations are mainly established by supercoiled structures and binding of NAPs (Berger et al., 2010; Dillon and Dorman, 2010).

NAPs are a highly diverse family of proteins that share DNA-binding as the only unifying feature. The family includes a number of highly conserved members such as the HU family of histone-like proteins (Anuchin et al., 2011), and some species-specific members, such as H-NS and MukBEF (Dillon and Dorman, 2010; Krogh et al., 2018). NAPs generally exert their function by bending, winding, lassoing, or bridging DNA, thereby bringing distant genomic positions into close spatial proximity (Dillon and Dorman, 2010; Krogh et al., 2018). Due to the saturation of bacterial chromosomes with coding sequence, NAP binding sites often occur within transcribed regions. NAP binding at or in open reading frames may impact gene-regulatory regions, and indeed, many NAPs are considered important TFs (Dillon and Dorman, 2010). The intercellular levels of the different NAPs are highly diverse, and vary between different growth phases, underlining their roles in organizing the genome (Azam and Ishihama, 1999; Talukder and Ishihama, 2015).

Studies in eukaryotes suggest that inter-chromosomal co-expression of genes may be facilitated in part by nuclear structure, and consistent correlation can be observed across expression datasets from highly diverse conditions (Kustatscher et al., 2017). The organization of the prokaryotic genome suggests that similar relationships between co-expression of genes and nucleoid structure may be highly relevant for bacterial gene regulation (Jeong et al., 2004; Xiao et al., 2011; Weng and Xiao, 2014).

In *E. coli*, evolutionarily conserved gene pairs have been found to be located at conserved distances across multiple strains, namely at periods of 117 and 33 kbps (Wright et al., 2007; Mathelier and Carbone, 2010). Past studies in *E. coli* have investigated co-expression as a function of distance between genes over large microarray dataset collections, observing patterns, and defining periods of co-expression in *E. coli* as short-range (~20 kbps distance), medium-range (100–125 genes), and long-range (600–800 genes; Jeong et al., 2004; Mathelier and Carbone, 2010; Junier and Rivoire, 2016). Such periodic patterns, albeit with several different frequencies, have also been observed in *Buchnera* spp. (Viñuelas et al., 2007). However, such microarray studies, investigating correlations in expression change and distance, might be subject to a technical bias due to the probe design of microarray chips (Balázsi et al., 2003; Kluger et al., 2003; Koren et al., 2007).

In this study, we created and utilized RNA-seq data from defined growth phases of *E. coli*, in order to investigate the association between co-expression of genes and their relative position on the chromosome. We find recurrent sinusoidal patterns with region dependent frequencies, in correlation of pairwise expression as function of genomic distance. We investigate potential sources of the observed periodic increases in co-expression and rule out transcription-induced supercoiling gradients and operons. Furthermore, we find no immediate connection to specific binding profiles of various TFs, SFs, or NAPs. We observe that the identified patterns match existing data on DNA-DNA interaction frequencies and propose a model to explain the observed patterns through nucleoid structure and transcriptional spilling and discuss these findings in an evolutionary perspective.

MATERIALS AND METHODS

Bacterial Growth and RNA Harvest

In order to obtain transcriptional data, RNA samples of *E. coli* BW25113 grown in rich media were purified at OD_{600} 0.2, 0.5, 1.2, 2.0, and 5.0 in duplicates, corresponding to early exponential, mid exponential, early stationary, mid stationary, and late stationary phases.

In short, 5 ml Lysogenic broth (or Luria Bertani broth; LB for short) w/o antibiotics were inoculated with a single colony from an ON LB agar plate with *E. coli* BW25113 and grown

at 37°C ON. For each individual sample; 1 L flask containing 200 ml LB (2 L flask with 400 ml for early exponential sample) preheated to 37°C was inoculated to an OD_{600} of 0.05 and grown in a preheated 37°C water-bath with ~160 rpm aeration. OD_{600} was controlled regularly to confirm that culture was in the growth phase of interest. At the wanted OD_{600} 100 ml culture (200 ml for OD_{600} ~0.2) was flash-frozen using liquid nitrogen, spun down at >15.000 RCF for 10 min in a −4°C precooled centrifuge. Supernatant was removed and pellets stored at −80°C for no more than 5 days, before RNA purification.

RNA purification was done using chloroform-phenol extraction, on all samples individually. In short, cell pellets were resuspended in phenol (pH ~4.5) and chloroform before heated to 80°C for 2 min under vigorous shaking. The aqueous phase was extracted and transferred into 10x volume of −20°C cold 96% ethanol and left at −20°C for 20 min for RNA precipitation. Solution was spun down at 15.000 RCF for 1 h at −4°C. Solution was removed and RNA-pellet dried and left at −80°C.

RNA-library preparation and sequencing were outsourced, following the TruSeq Stranded Total RNA standard protocol from Illumina, using a HiSeq4000. Raw data has been deposited at the national center for biotechnology information (NCBI) gene expression omnibus (GEO) database under the accession number GSE153815.

Computing the Pearson Correlation Between Bins Expressional Profiles and Estimate Periodicities Within

RNA-samples were sequenced using a Hi-seq 4,000, as 2x75bp paired-end reads. Raw reads were mapped to *E. coli* BW25113 (NZ_CP037857.1) or *E. coli* MG1655 (NC_000913.3), respectively, using Bowtie2 with local mapping settings (--local). Sequence alignment map (SAM)-files were converted to bam, sorted, and indexed using Samtools (v.0.1.19).

To avoid gene length bias, 500 bp wide genomic features were created, spanning the entire genome. Genome coverage were calculated for all features, using the Rsubread (v.2.2.4; Liao et al., 2019), for each sample respectively, and normalized to fragments per million. Due to differing ribosomal RNA (rRNA)/transfer RNA (tRNA) depletions between the causing bias in the normalization, all bins associated with rRNA/tRNA/transfer-messenger RNA (tmRNA) expression were nulled (based on genome annotations: ASM584v2 and ASM435510v1). Furthermore, all correlation values related to any bin associated with rRNA/tRNA/tmRNA computed were substituted with an empty value. Biological duplicates averaged into single datasets for each growth phase. Finally, correlation of expressional change between bins were calculated for all possible combinations of bins using the R (v.3.6.3) built-in correlation function cor() with arguments use = "complete.obs" (**Figure 1A**).

To investigate periodicities of Pearson correlation between bins relative to distance, the average Pearson correlation for bins within windows of 400 kbp sliding at 5 kpb across the genome was calculated. Periodicities within these averaged correlation profiles were estimated using the Lomb-Scargle periodogram analysis R-package lomb (v.1.2; Ruf, 1999). Lomb-Scargle periodogram analysis was used due to its ability to detect rhythms in noisy incomplete data and the ability to ascertain the significance of estimated peaks using PNmax (Lomb, 1976; Scargle, 1982; VanderPlas, 2018). PNmax is defined as −log(*p*-value), hence the higher PNmax, the higher the probability of the estimated periodicity being valid. The distribution of PNmax values for all estimated periodicities across the genome observed within the transcriptional data acquired in this paper was compared to the PNmax distribution for estimated periods within 400 randomized datasets (**Supplementary Figure S2**). In short, data was randomized by assigning all genomic bins random unique genomic positions, followed by Lomb-Scargle periodogram analysis as described above.

Transcriptional Factor, NAP Enrichment, and Sigma Factor Enrichment

Data from RegulonDB were used. Only TFs, NAPs, and SFs associated with more than a total of 20 genomic bins (+/− 500 bp of regulatory targets) were analyzed. Regulatory targets of TFs and SFs were obtained from RegulonDB (Release: 10.7 Date: 05/04/2020; Huerta et al., 1998; Santos-Zavaleta et al., 2019).

Chromosomal Conformation Capture Comparison

For the investigation of any relation between DNA-DNA interaction frequencies and correlation of pairwise expression of genomic positions processed data from Lioy et al. (2018) were acquired from GEO database, accession GSE107301. Genomic bins used by Lioy et al. (2018) were 5 kbp wide, hence for this part correlation was calculated between bins of the same size, 5 kpb, created as described in 6.2.

General

Software

All data were aligned to *E. coli* K-12 substrain BW25113 (Version NZ_CP037857.1), using Bowtie2 (v.2.3.5.1; with --local; Langmead and Salzberg, 2012) *via* whole genome sequences retrieved through the NCBI Nucleotide database. The output SAM files were converted to Binary Alignment Map (BAM) files and subsequently sort and indexed using Samtools (v.1.10; Li et al., 2009). Initial visual verification of data was done using SeqMonk (v.1.45.4). Fragments were mapped to 500 bp genomic bins using the R-package Rsubread (Liao et al., 2019). For data analyses, *R* was used. R-Packages: For general data tidying and wrangling the tidyr and reshape2 were used. Plots and data-visualizations were made using ggplot2 (Wickham, 2016). For general file and data manipulation Bash, Python, and Powershell was used. Explanatory figures made in power point.

Other Species

To investigate the presence of periodic patterns in other species, RNA-seq as raw fastq data from the NCBI GEO database was retrieved and analyzed as described in 6.2 top section, see **Figure 1A**. Data used from *Streptococcus sanguinis SK36* (Data from BioProject PRJNA381491 available at GEO database; El-Rami et al., 2018), and *Listeria monocytogenes* (Data from BioProject PRJNA270808 available at GEO database; Tang et al., 2015).

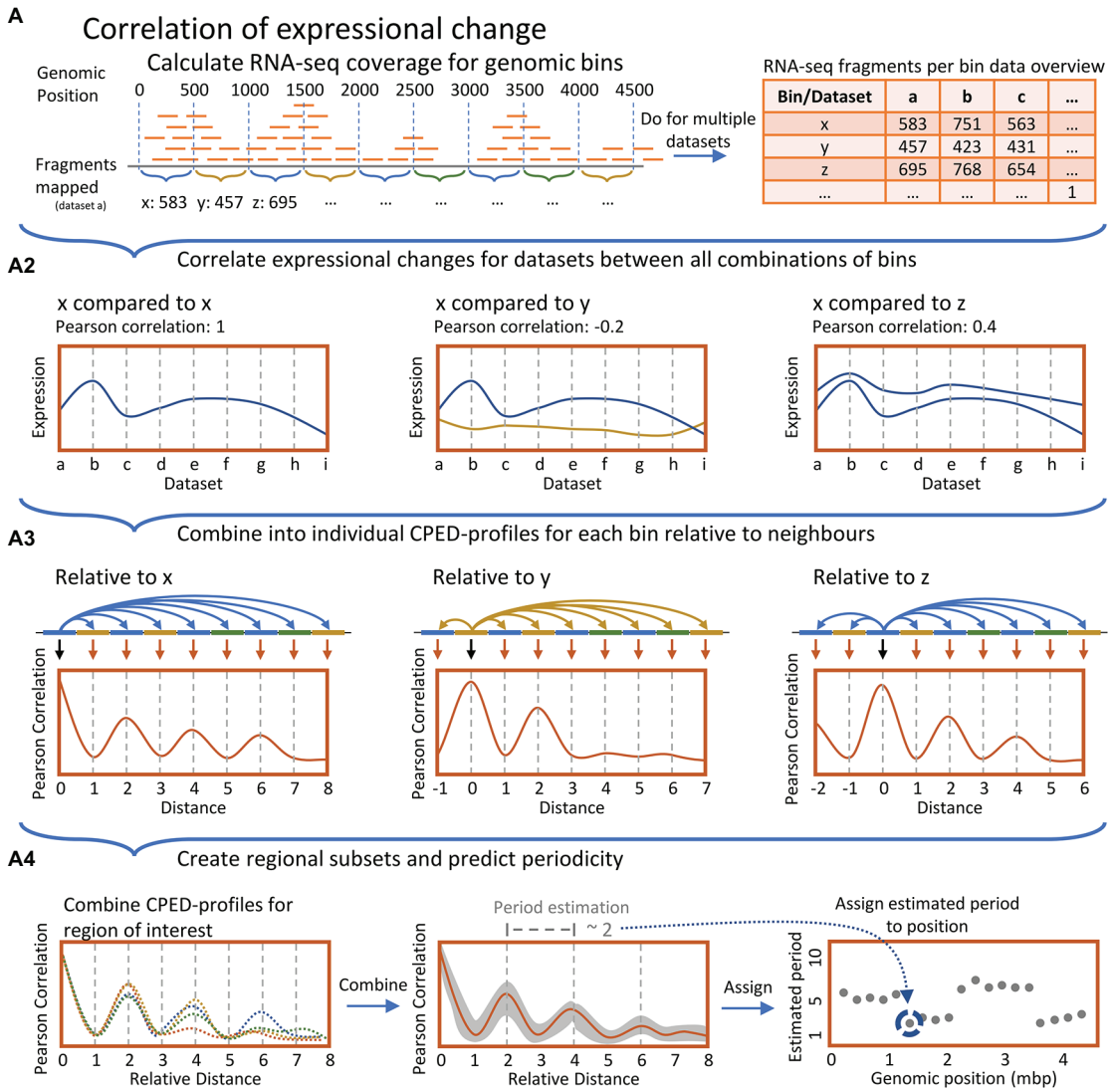


FIGURE 1 | (A) Simplified data processing overview. Top panel: For all datasets, genomic coverage of RNA-seq data were calculated and summed for 1,000 bp genomic bins sliding at 500 bp. **(A2)**: Pearson correlation coefficients were computed relative to pairwise expression over all datasets between all bins. **(A3)**: Correlation coefficients for individual genomic bins relative to all neighbors were combined, yielding individual Expressional Correlation relative to Distance Profiles (CPED-Profiles) for all genomic bins, individually. **(A4)**: Multiple CPED-profiles were combined and averaged for bins within genomic regions of size 400 kbp. Subsequently, recurring periodicities were predicted through a Lomb-Scargle periodogram analysis, and the most statistically significant period for each region was mapped in a plot according to the center of the genomic 400 kbp region investigated.

RESULTS

Genomic Co-expression Correlates in Periodic Patterns

Transcription is highly regulated, with the main component of the transcriptional machinery being the RNAP (Saecker et al., 2011). The RNAP scans for promoters in a three-dimensional Brownian diffusion, where it stochastically binds and releases accessible DNA stretches until a matching promoter region is recognized and transcription is initialized (Wang et al., 2013). The bacterial nucleoid is a highly structured macromolecule that dynamically changes spatial organization in response to changing growth conditions (Valens et al., 2004; Cagliero et al., 2013). Interestingly, the distribution of RNAP across the nucleoid also changes dramatically during different growth conditions, suggesting a possible link between nucleoid structure and transcription regulation (Jin and Cabrera, 2006). The link is further underlined by the impact spatial proximity has on the efficiency of regulation *via* TFs (Pulkkinen and Metzler, 2013). An unbiased comparison of co-expression, of otherwise functionally unrelated genes, across different datasets obtained under different growth conditions, may highlight common co-expression patterns.

Previous studies investigating patterns in co-expression have been based primarily on micro-array data. However, as early

studies have revealed a risk of microarray design bias in this type of study (**Supplementary Figure S1**; Jeong et al., 2004; Xiao et al., 2011; Junier and Rivoire, 2016), we created transcriptomic data based on RNA-seq (materials and methods). In addition, previous studies either investigated entire large databases of micro-array data or only a few seemingly arbitrarily chosen datasets, without taking growth-specific changes to the nucleoid into account, possibly missing important structural features due to noise (Jeong et al., 2004; Xiao et al., 2011; Junier and Rivoire, 2016). Lastly, some of the previous studies erroneously consider distance as ordinal, by gene order, rather than numeric, by genomic position, leading to a misleading interpretation of period in the context of genomic distance (**Supplementary Figure S1**). To determine the potential effect of nucleoid structure in co-expression patterns, we investigated data from few, but well-defined growth phases, with reported changes in nucleoid compaction state (i.e., between exponential and stationary growth). This was done in order to obtain maximum dataset variation in both changes in expression and nucleoid structure (Krogh et al., 2018). This way patterns in co-expression relative to genomic position can be related to changes in nucleoid structure when moving from highly to less condensed nucleoids.

In order to investigate distance determined co-expression patterns, the acquired transcriptomic data was divided into bins of 500 bp in size. Subsequently changes in expression, induced by a change in growth phase, were compared between all possible combinations of these genomic bins and the correlation coefficients of pairwise expression were related to the distance between bins (**Figures 1A**-top panel, **A2**). This yielded a Correlation of Pariwise Expression relative to Distance Profile (CPED-Profile) for all bins, respectively (**Figure 1A3**). Periodic patterns in the data were analyzed by combining CPED-profiles for all bins within a given genomic window followed by an estimation of the observed wavelength of the averaged profile. Wavelength estimation was conducted by use of Lomb-Scargle least-squares frequency periodogram analysis (**Figure 1A4**; Lomb-Scargle for short. See methods; Lomb, 1976; Scargle, 1982; Ruf, 1999; VanderPlas, 2018).

In brief, the Lomb-Scargle method compares modeled sine/cosine frequency functions within a defined window of possible frequencies, to the experimental data and estimates the goodness of fitness by computing the least-squares. Hence, each specific region has unique estimates for all frequencies within the arbitrarily chosen range of 10–70 kbps. In **Figure 2**, the frequency with the highest fitness is reported for each specific region.

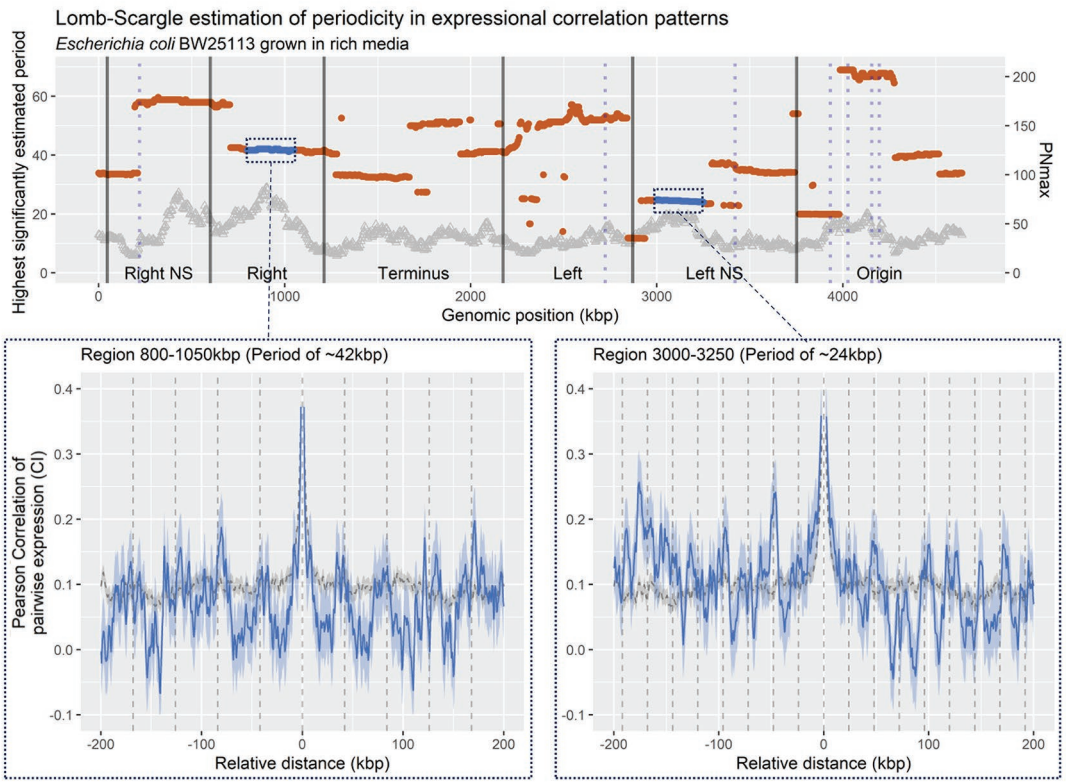


FIGURE 2 | Lomp-Scargle Periodogram frequency estimation. **Top panel**: Estimated frequencies of correlation profiles at a relative distance of +/– 200 kbp, for bins within 400 kbp genomic regions, sliding at 5 kbp. Dots are positioned at the genomic region middle. Black vertical lines indicate structured and unstructured (mixed) macrodomains in *Escherichia coli* as defined by Valens et al. (Valens et al., 2004; corrected for strain BW25113 coordinates). Blue vertical lines indicate the position of ribosomal RNA (rRNA) operons in *E. coli* BW25113. Left y axis indicates PNmax [–log(*p*-value) of the estimated period], PNmax > 3, equal $p < 0.05$. **Bottom panel**: Averaged CPED-profiles of regions (blue line), or for entire genome (gray dashed), with confidence interval (CI; 95%) of true mean. **Left panel**: Region 800–1,050 kbp, within Right structured macrodomain, with a predicted period of ~42 kbp. **Right panel**: Region 3.000–3.250 kbp, within Left unstructured macrodomain, with a predicted period of ~24 kbp.

However, we note that there may be sub or super frequencies in each region with almost as high estimated fits, explaining the abrupt changes in highest estimated period. The wavelength estimations across the entire genome (**Figure 2**-top panel) shows multiple large regions with a conserved estimate of wavelength, somewhat matching those defined prior by Valens et al. (2004). The regions show statistically significant periodicities estimated through PNmax defined as −log(*p*-value), as compared to shuffled genomes (**Supplementary Figure S2**). For example, at position 0–200 kbp, which shows a periodic pattern with an estimated period of 33 kpb, and terminus showing a roughly 35 kbp wavelength, as proposed based on *matS* domains (Dupaigne et al., 2012). This is indicative of recurring increases in co-expression for genomic bins relative to distance between bins. When looking closer at specific regions, it is evident that periodicities in co-expression are present as wavelike patterns in correlation relative to distance (**Figure 2**-bottom panel).

The observed change in averaged CPED-profile depending on genomic position suggests a systematic organization of the genome. Two known structural mechanisms, which could explain such high correlation in expressional change between neighboring genes are; (1) the transcriptional clustering of genes in operons or (2) the impact of supercoiling gradients created by active transcription and/or replication. However, operons can readily be excluded as the source, since only 13 operons larger than 10 kbp exist in the *E. coli* genome with the largest at 17.840 bp (RegulonDB; Operons Release: 10.6 Date: 10/04/2019; Huerta et al., 1998; Santos-Zavaleta et al., 2019). This leaves mechanism (2) to be further explored.

Patterns in Co-expression Are Independent of Transcription-Induced Supercoiling Gradients

Transcription-induced supercoiling gradients are considered a central regulatory mechanism since more than 2.000 genes in *E. coli* are sensitive to DNA supercoiling (Blot et al., 2006; Marr et al., 2008). Such gradients are generated by the accumulation of positive supercoiling in front of the transcribing RNAP, which promote expression of downstream genes, and negative supercoiling trailing the transcribing RNAP, which will inhibit transcription of upstream genes (Liu and Wang, 1987; Ma et al., 2013). Since transcription supercoiling gradients are dependent on the orientation of transcription, the distance to neighboring bins in the CPED-profiles must be analyzed relative to transcription orientation, rather than genomic position in order to investigate the potential contribution of supercoiling gradients. In order to do this, data bins were assigned an orientation according to the transcriptional direction of annotated gene(s) spanning and/or starting/ending within the bin. Bins without any assigned orientation or assigned with both clockwise and counter-clockwise orientations (relative to genomic coordinate) were discarded (~2/3rds of bins). Subsequently, all bins with a counter-clockwise orientation were mirrored to correctly indicate position relative to transcription (**Figure 3A**).

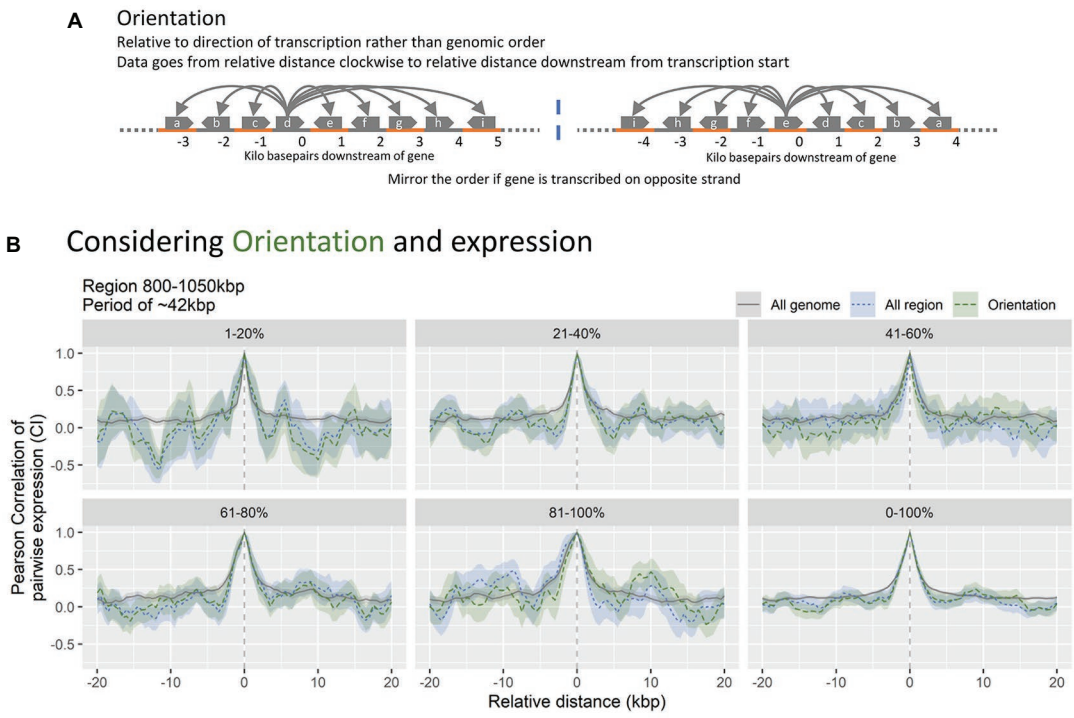


FIGURE 3 | (A) Simplified data mutation overview. The distance variable of the CPED-profiles is converted to distance relative to transcription orientation of the bin, rather than being relative to the genomic position. Notice how the named bins change order upon mirroring, negating the original distance. **(B)** Data subsets relative to expression strength and orientation, for region 800–1,050 kbp, shown in **Figure 2**. Notice no significant changes to the CPED-profiles when accounting for orientation (blue vs. green).

Besides the orientation dependency, supercoiling is further related to the expressional strength. To further account for this, bins were divided into five percentiles of equal size according to their expressional strength (**Figure 3B**).

If the contribution of supercoiling gradients to the observed gene expression patterns is significant, the expected Pearson correlation of bins aligned to transcription orientation must be high downstream, relative to interval 0, whereas negative correlation is expected upstream. We do not observe an increase in co-expression upstream of direction of transcription-oriented bins compared to non-orientated data for the region or the entire genome. The lack of changes to co-expression relative to orientation of transcription suggests that transcription-induced supercoiling gradients are relieved instantly, possibly through DNA gyrase or Topoisomerase IV (Sharma and Mondragón, 1995; Schvartzman et al., 2013). This is in accordance with observations that supercoiling gradients lack during exponential phase where the nucleoid is condensed (Lal et al., 2016). The periodic pattern within the region 800–1,050 kbp is present when accounting for orientation, indicating that transcription-induced supercoiling is not the direct mechanism behind the observed patterns (**Supplementary Figure S3**). The exclusion of transcription-induced supercoiling is further supported by previous observations that the influence of supercoiling generated gradients decrease over distances of roughly 10 kbp, and that it has limited propagation outside domains of 10 kbp size as well (Postow et al., 2004; Sobetzko, 2016). When accounting for expression strength, the pattern amplitudes and frequencies however seem to vary, but due to the high amount of noise it is difficult to conclude any relation to expression strength (**Supplementary Figure S3**).

Having investigated structural mechanisms, which could explain the observed patterns in co-expression, we turn to another well-described mechanism behind co-expression, namely TFs modulating transcription of target-genes in bulk.

Patterns in Co-expression Are Independent of Association With Transcription Modulating Proteins

Co-expressional patterns between distantly positioned genes on the genome have historically been attributed to TFs. These are proteins that modulate transcription through association to specific motifs and/or DNA structures (Djordjevic, 2013). TFs commonly modulate groups of genes associated with distinct environmental cues. An example is fumarate and nitrate reduction regulatory protein (FNR) that mediates transition from aerobic to anaerobic conditions by inducing genes related to anaerobic metabolism, and inhibits genes related to aerobic metabolism (Salmon et al., 2003; Kang et al., 2005; Myers et al., 2013). In this section, we also consider SFs, which are proteins that mediate selective promoter recognition by the RNAP and thereby modulate transcription as well (Saecker et al., 2011; Svetlov and Nudler, 2013; Duzdevich et al., 2014). In addition, the NAPs, which confer genome structuring, are considered gene-regulatory proteins since they often have a major impact on the transcription profile (Dillon and Dorman, 2010). Thus, in the following paragraph, the term TF includes both NAPs and SFs.

To investigate the relation between high co-expression and proximity known regulatory targets of the modulating proteins, we obtained information about TF regulates from RegulonDB (TF – gene interactions and Sigma – gene interactions Release: 10.7 Date: 05/04/2020; Huerta et al., 1998; Santos-Zavaleta et al., 2019).

The co-expression was compared between different bin subsets for each type of TF, respectively (see **Figure 4** panel A for visual explanation). Co-expression of bins within 500 bp on either side of a regulatory target was considered as associated with binding sites (Termed *At*). These bins were compared to bins associated with mutated mock binding sites (Termed *Between*), created by moving real binding sites to positions between real sites in a clockwise manner. These were further compared to 100 randomly generated bin-subsets of equal size to the original binding site data (Termed *Random*), and mutated data where co-expression was computed for all bins without any association to the modulating protein in question (Termed *Without*).

The expected outcome of TF mediated co-expression would be a higher averaged correlation for bins at regulatory targets compared to all other subsets. This is indeed observed for TF's ArcA, Cra, FlhDC, FNR, IscR LexA, Lrp, ModE, NarL, NarP, OxyR, H-NS, Sigma 24, Sigma 32, Sigma 38, and Sigma 54, suggesting that bins associated with binding of these specific TFs are indeed co-regulated during the investigated growth phases, as expected. However, this approach only identifies genome-wide co-expression of bins and does not sufficiently account for the observed periodic patterns in co-expression observed locally.

To investigate the impact of TFs at the local genome scale, bins associated with TF regulation were excluded from the averaged CPED-profile. This way any impact on co-expression conferred by the TF would be eliminated, and a diminished pattern would be expected in the case of TF binding impacting the observed periodic patterns. Six of the 32 TFs investigated were associated with more than 20 bins (of 500 possible) within the specific region shown in **Figure 4B2** and were chosen for analysis. None of the six TFs show major impact on the observed periodic pattern after exclusion of TF-associated bins. The lack of major changes to the periodicity, indicates that TF association is not the primary mechanism behind the observed patterns. Hence the patterns in co-expression might emerge through a less known, but perhaps fundamental mechanism, which depends solely on the spatial structure of the nucleoid. The lack of major changes when excluding NAPs does not exclude the possibility of their involvement in the observations through genome structuring; it does simply indicate that their regulatory impact is not significant for the periodicities in transcription.

Transcriptional Spilling, Expression Strength, and Correlation Pattern of Expression Change

The observed patterns in expressional change cannot be accounted for by expression modulation through protein binding (**Figure 4**), operon structure, or transcription induced supercoiling gradients, suggesting a more fundamental mechanism (**Figure 5**). This notion is further supported by

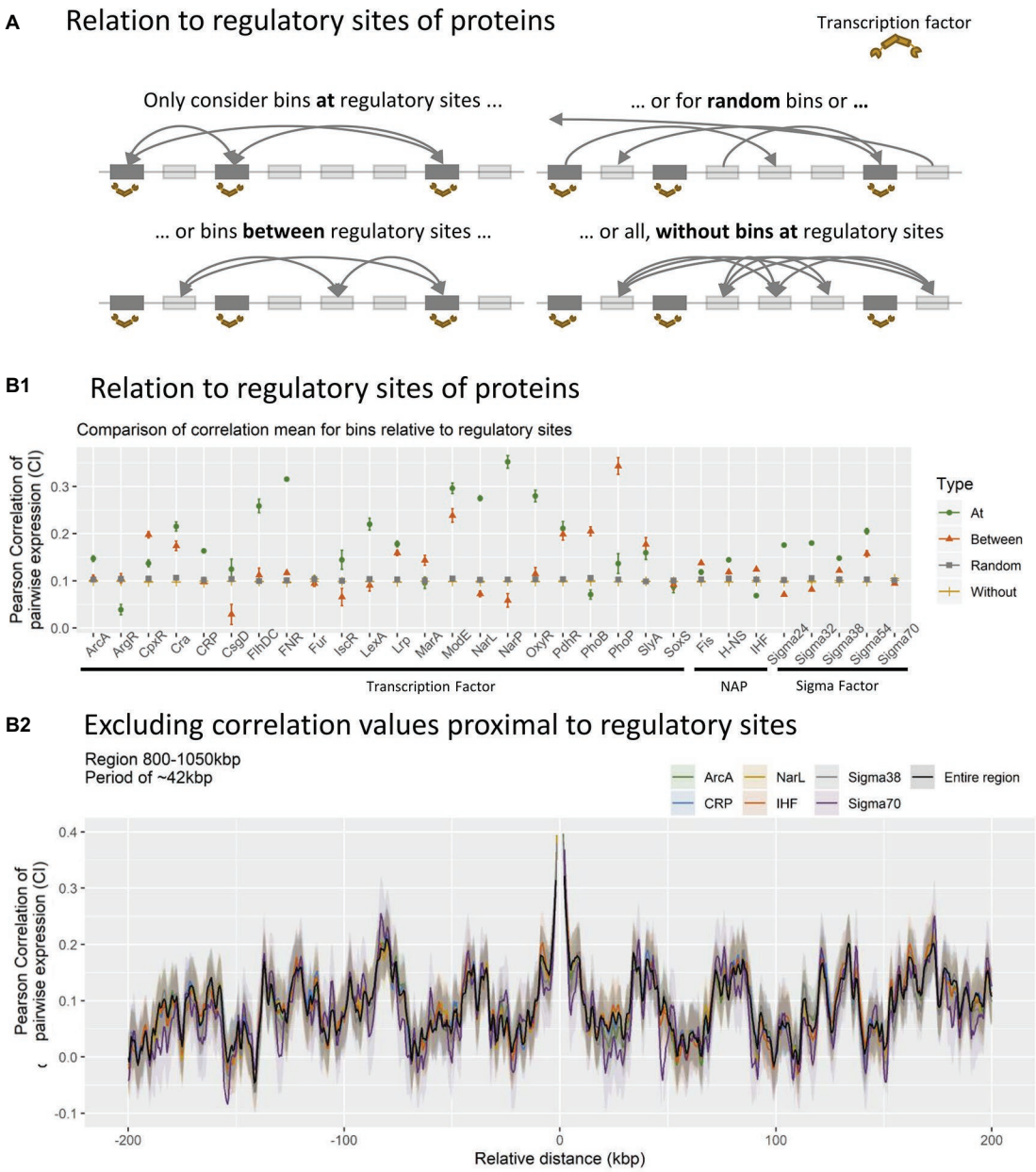


FIGURE 4 | (A) Schematic overview of how different data subsets were created. All bins within 500 bp of nucleoid associated protein (NAP)/transcriptional factor (TF) regulatory site profiles (BS), were identified and the expressional correlation distributions was calculated for all BS-profiles individually (At). This was compared to mutated binding site profiles (between), random binding site profiles (random), and the overall distribution without the coefficients at binding sites (without). **(B)** Relation between protein binding and correlation coefficients. **(B1)**: Distributions of correlation coefficients within subsets. All bins associated with binding (At, green circle), All bins exactly between bins with associated binding (between, red triangle), average distribution of 100 random (with $N = N^{At} = N^{Between}$) sets of bins (random, gray square), or total distribution between all bins NOT associated with binding (without, yellow line). **(B2)**: Impact of excluding bins at binding sites for TFs/sigma factors (SFs)/NAPs with more than 20 binding sites within the 800–1,050 kbp region.

observations of periodicities in correlation of expressional change and distance between positions of the genome in RNA-seq data (**Figure 6C**) from both *S. sanguinis SK36* (Data from BioProject PRJNA381491 available at GEO database; El-Rami et al., 2018), and *L. monocytogenes* (Data from BioProject PRJNA270808 available at GEO database; Tang et al., 2015). We thus speculate that the mechanism is related to the fundamental parts of the transcriptional machinery and nucleoid structure.

RNAP searches for promoters in a three-dimensional Brownian diffusion, and binds stochastically onto accessible DNA stretches until a promoter matching the associated SF is recognized

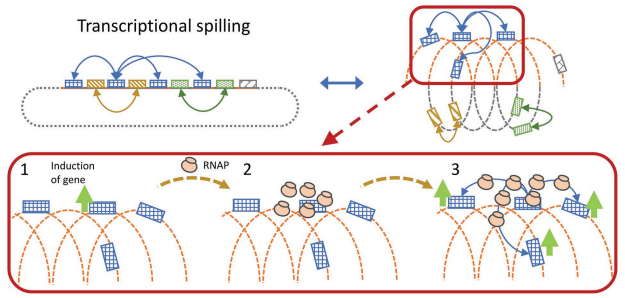


FIGURE 5 | Simplified illustration of transcriptional spilling. Genes far from each other in the linear genome may be spatially close due to compaction of the genome. Gene induction leads to RNA polymerase (RNAP) recruitment, whereby the increase in local concentration of RNAP might spill onto spatially close genes increasing the chance of successful gene transcription initiation.

(Saecker et al., 2011; Wang et al., 2013). In a condensed DNA structure, distant genomic regions may be brought within close spatial proximity, enabling the RNAP to readily diffuse between regions, thus sharing the same local pool of RNAP. Gene induction will recruit RNAP and increase the local concentration, which in turn increases the transcription initiation rate of genes in close spatial proximity. This will effectively create co-expression (**Figure 5**) and is visible as condensed RNAP foci during optimal growth in *E. coli*, where especially the rRNA operons are spatially clustered (Cabrera and Jin, 2003).

According to the transcriptional spilling hypothesis (Krogh et al., 2018), sharing an RNAP pool in an expanding nucleoid will cause the expression of all genes at the specific RNAP pool to drop at a similar rate. Given that the distance between the genes will increase, diluting the RNAP pool and reduce the chance of transcription initiation, and vice versa in a condensing nucleoid. This will cause spatially close genes to exhibit high correlation in pairwise expression over the cause of datasets where the nucleoid structure is modulated. If the change in distance occurs faster than the sampling rate, the pattern would be less obvious due to more noise originating from stress induction. However, if the structural changes are slowly induced within regularly structured DNA-regions, the mechanism would effectively be observable as periodic co-expressional patterns relative to distance.

If the observed sinusoidal patterns in expressional correlation are related to nucleoid structure, we hypothesize that they should match existing observations of the topology of the bacterial nucleoid. Furthermore, periodic patterns should be observable within all biological systems with nucleotide macromolecules. To test this hypothesis, the relation between DNA-DNA interaction frequencies and correlation of pairwise expression of genomic positions was investigated. The spatial proximity of DNA was determined through Chromosomal Conformation Capture (3C) data of *E. coli* during exponential growth phase, since the nucleoid is condensed during this phase, from Lioy et al. (2018). RNA-seq data was binned into 5,000 bp sized genomic bins, to match the structure of the 3C data. The correlation of pairwise expression between all unique bin pairs were computed and grouped according to the spatial proximity of the pair, respectively. **Figure 6B** shows the relation between spatial proximity of bins and their respective level of correlation in pairwise expression. Any relation between bins within +/−20 k bp of each other relative to genomic position was excluded to eliminate co-expression from operons and transcription-induced supercoiling gradients.

The analysis shows a genome-wide increase in correlation of pairwise expression relative to the level of measured DNA-DNA interaction. That is, the more spatially close bins are, the more they correlate in pairwise expression. This indicates that bins in spatial proximity tend to change expression levels in a coordinated manner. The observation of similar patterns in distantly related species underlines the possibility of a fundamental mechanism, related to basic systems present in most if not all living systems. It further reduces the risk of observations being an experimental bias, and to the best of our knowledge no technical bias exists in RNA-seq analysis that could cause periodic patterns in co-expression.

DISCUSSION AND OUTLOOK

In this study, sinusoidal patterns of correlation of pairwise expression, co-expression, with respect to genomic distance were observed in RNA-seq data. The estimated frequencies of the sinusoidal patterns were not explainable by operon structure or transcription-induced supercoiling. Furthermore, no apparent effect of most common TFs, nucleoid associated proteins, or SFs was observed.

This led to the hypothesis that the observed patterns are due to supercoiled macro structures (of more than 20 kbps, excluding immediate transcription-induced supercoiling gradients). These structures lead to spatial clusters of genes, which are distant with respect to the linear genomic sequence, but proximal in space. Such spatial clusters could potentially share RNAP pools, in which induction of genes might increase the local concentration of RNAP and thus induce the expression of other genes in spatial proximity. Underlined by the observed link between correlation of pairwise expression and DNA-DNA interaction frequencies (**Figure 6**). We hypothesize that spatial positioning of genes in the three-dimensional structure of the nucleoid, possible through NAPs, plays a significant role in the regulation of gene expression *via* RNAP, due to transcriptional spilling.

The size of a regular sinusoidal DNA structure with 20 kpb loops would be able to go around an averaged sized *E. coli* cell, hence we do not propose a strict sinusoidal DNA structure of the regions with patterns. But rather local flowerlike plectonemic structures, were bins at specific distances are anchored together, with long stretches of less structured DNA in between (Krogh et al., 2018). Given a length of an *E. coli* cell during exponential growth is ~2–3 μm, with a genome size of ~4.6 mbp. The length of 1 bp DNA is ~3.4 Å (340 pm), making the genome ~1.564 μm long, or 500–800 times the length of the typical *E. coli* cell. A circular stretch of 20 kbp DNA would be ~6.8 μm in circumference, with a diameter of ~2.1 μm.

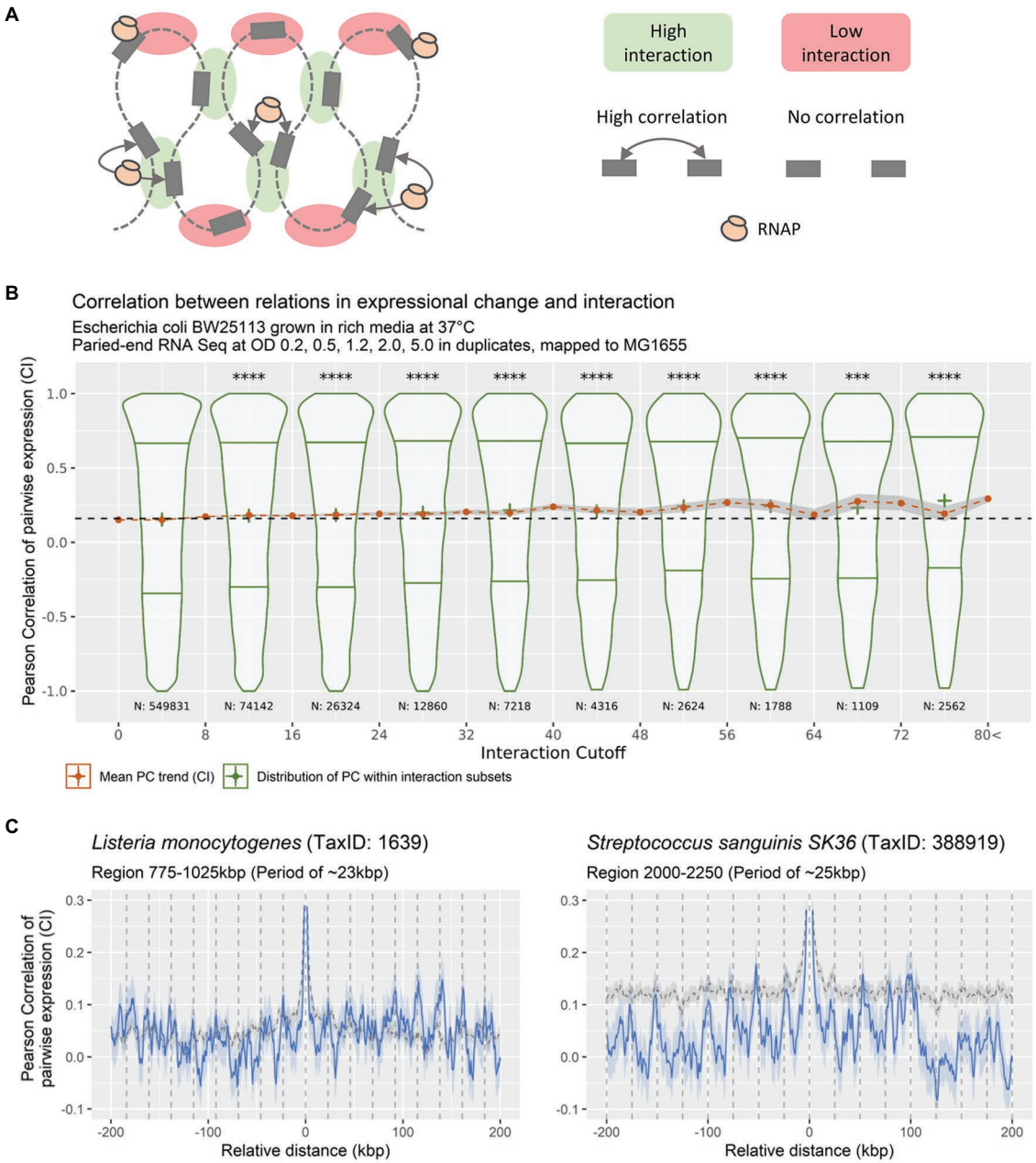


FIGURE 6 | (A) Theoretical relation between DNA-DNA interaction frequencies and correlation of expression change between bins. If a local pool of RNAP is shared between spatially close genes, any changes to this pool will result in a coordinated response of the genes within the pool. **(B)** DNA-DNA Interaction relative to pairwise expression correlation (bin size 5 kbp), for DNA-DNA interactions of bins with a linear distance of more than 20 kbps. Violin-plots showing the distribution of expressional correlation for genomic positions grouped in eight distributions of inequal size according to increasing DNA-DNA interaction, with lower and upper quantiles and cross at mean (green distributions). Red dots at mean +/– confidence interval (95%) for expressional correlation grouped in 21 unequal-sized groups based on DNA-DNA. Stars indicate significance of distribution means being equal to the distribution with the lowest DNA-DNA interaction (****$p < 0.00005$ using a pair-wise Wilcoxon test, adjusting for multiple testing using Bonferroni, most were at $p < 2e^{-16}$). **(C)** Averaged CPED-profiles showing mean and 95% confidence interval for true mean for data retrieved from NCBI GEO, mapped and analyzed locally. Gray dashed line indicates averaged CPED-profile for the entire genome. Blue solid line indicates averaged CPED-profile for the specific local regions as indicated in subtitle. Left panel: Data from BioProject PRJNA270808 (Tang et al., 2015), showing periodic patterns of ~23 kbp, within region 775–1,025 kbp of the Gram-positive *Listeria monocytogenes*. Right panel: Data from BioProject PRJNA381491 (El-Rami et al., 2018), showing periodic patterns of ~25 kbp, within region 2000–2,250 kbp of the Gram-positive *Streptococcus sanguinis SK36*.

In the light of recent findings showing an inconsistency between regulatory TF networks and gene co-expression, there is a gap in the understanding of how cells adapt co-expression in changing environments (Larsen et al., 2019).

A plethora of mechanisms interact and interfere with transcription of genes within complex organisms such as a bacterium. Here, DNA-structure dependent transcriptional spilling might provide an explanation for observations that do not fit in the existing paradigm of transcription regulation. Central to environmental adaption and proliferation are changes to the nucleoid structure and expressional profile of prokaryotes (Dillon and Dorman, 2010). In this perspective a fundamental regulatory mechanism utilizing structure could represent a primordial way of handling stress for DNA based organisms. Perhaps transcriptional spilling is an original simple regulatory mechanism used before more advance regulatory circuits developed, but after structural proteins such as NAPs evolved. The impact of DNA-structure on regulation of transcription is further underlined by the effect spatial distance between genes encoding TFs and the corresponding regulatory gene targets have (Pulkkinen and Metzler, 2013).

By extension, if genes are positioned in an organized manner according to expression change, it potentiates the impact location will have on insertion of synthetic DNA into the chromosome. Indeed, position-specific effects on expression of chromosomal inserted synthetic DNA have been observed (Bryant et al., 2014; Sauer et al., 2016; Scholz et al., 2019). With up to 1.000-fold differences in expression based on genomic position, it is critical to assess the insertion of synthetic DNA, and transcriptional spilling may be an important mechanism to consider. This is further supported by reduced transferability of highly expressed genes and observed proportional relation between DNA-DNA contact frequencies and transcriptional level, which suggest that highly transcribed genes have a high degree of spatial organization (Park and Zhang, 2012; Lioy et al., 2018). Highly expressed genes would exert increased constrain on the conservation of the spatial structure, since high expression yields high spilling, and thus any uncontrolled spilling could cause unwanted gene induction. This is in line with experimental data, which show a good correlation between transcription activity and stability of looped DNA domains (Dillon and Dorman, 2010).

In an evolutionary perspective, the investigation of patterns in co-expression and genomic position may unveil evolutionary forces driving optimal gene insertion and deletion. Forces dependent on the organization and conservation of distances between genes and further explain the conservation of seemingly non-sense DNA such as pseudo- and phantom-genes. Such genes account for ~5% of the total gene pool (225 genes; Rogozin, 2002; Keseler et al., 2013; Goodhead and Darby, 2015). These genes may act as structural elements that help the cell adjust local transcriptional spilling, both reducing unwanted spilling by acting as inactive DNA-element without recruitment of RNAP or vice versa.

It may also explain why deletion of cryptic prophage elements in *E. coli* leads to decrease in resistance to osmotic stress and antibiotics (Wang et al., 2010). If transcriptional spilling significantly influences the expressional profiles of genes it might play an important role in nucleoid structure as an evolutionary driving force. Since integration of foreign DNA, that is not optimal with regards to transcriptional spilling, might influence the ability of quick adaption to changing environments and general proliferation.

Transcriptional spilling would be another constraint that should be considered when investigating chromosomal structural evolution. In addition, this may have implications for expression of recombinant proteins *via* chromosomal insertions and may also apply to large plasmids. Especially in an evolutionary timeframe or productional framework, it could have implications in adaption and productional output. Therefore, the mechanism is highly relevant for understanding basic prokaryotic and archaic genome evolution and for optimization of cell-lines used in synthetic biology.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153815.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

CK acknowledges support by the German Research Foundation (KA 3541/3-2 and FOR 5042/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.02002/full#supplementary-material

SUPPLEMENTARY FIGURE S1 | Correlations between genes associated to either genomic coordinate or to ordinal gene number. Plot made using data from Many Microbe Micro-arrays Database (Faith et al., 2008). Gene expression data either associated to ranked gene number based on genomic position zero, or to genomic bin based on gene coordinates. Only deviates from expression association (A-top panel) in method outlined in **Figure 1**.

SUPPLEMENTARY FIGURE S2 | Distributions of PNmax, –log(p-value) of the estimated period for periods estimated within real data (Real), compared to the distribution within iterations of randomized datasets (Random), PNmax > 3, equal $p < 0.05$.

SUPPLEMENTARY FIGURE S3 | Data subsets relative to expression strength and orientation, for region 800–1,050 kbp. % indicates the ranked expression strength of gene-subset.

REFERENCES

Anuchin, A. M., Goncharenko, A. V., Demidenok, O. I., and Kaprelyants, A. S. (2011). Histone-like proteins of bacteria (review). *Appl. Biochem. Microbiol.* 47, 580–585. doi: 10.1134/S0003683811060020

Azam, T. A., and Ishihama, A. (1999). Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* 274, 33105–33113. doi: 10.1074/jbc.274.46.33105

Bakshi, S., Choi, H., and Weisshaar, J. C. (2015). The spatial biology of transcription and translation in rapidly growing *Escherichia coli*. *Front. Microbiol.* 6:636. doi: 10.3389/fmicb.2015.00636

Balázsi, G., Kay, K. A., Barabási, A. L., and Oltvai, Z. N. (2003). Spurious spatial periodicity of co-expression in microarray data due to printing design. *Nucleic Acids Res.* 31, 4425–4433. doi: 10.1093/nar/gkg485

Berger, M., Farcas, A., Geertz, M., Zhelyazkova, P., Brix, K., Travers, A., et al. (2010). Coordination of genomic structure and transcription by the main bacterial nucleoid-associated protein HU. *EMBO Rep.* 11, 59–64. doi: 10.1038/embor.2009.232

Binder, S. C., Eckweiler, D., Schulz, S., Bielecka, A., Nicolai, T., Franke, R., et al. (2016). Functional modules of sigma factor regulons guarantee adaptability and evolvability. *Sci. Rep.* 6:22212. doi: 10.1038/srep22212

Blot, N., Mavathur, R., Geertz, M., Travers, A., and Muskhelishvili, G. (2006). Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep.* 7, 710–715. doi: 10.1038/sj.embor.7400729

Bryant, J. A., Sellars, L. E., Busby, S. J. W., and Lee, D. J. (2014). Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res.* 42, 11383–11392. doi: 10.1093/nar/gku828

Cabrera, J. E., Cagliero, C., Quan, S., Squires, C. L., and Ding, J. J. (2009). Active transcription of rRNA operons condenses the nucleoid in *Escherichia coli*: examining the effect of transcription on nucleoid structure in the absence of transertion. *J. Bacteriol.* 191, 4180–4185. doi: 10.1128/JB.01707-08

Cabrera, J. E., and Jin, D. J. (2003). The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues. *Mol. Microbiol.* 50, 1493–1505. doi: 10.1046/j.1365-2958.2003.03805.x

Cagliero, C., Grand, R. S., Jones, M. B., Jin, D. J., and O'Sullivan, J. M. (2013). Genome conformation capture reveals that the *Escherichia coli* chromosome is organized by replication and transcription. *Nucleic Acids Res.* 41, 6058–6071. doi: 10.1093/nar/gkt325

Dillon, S. C., and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* 8, 185–195. doi: 10.1038/nrmicro2261

Djordjevic, M. (2013). Efficient transcription initiation in bacteria: an interplay of protein-DNA interaction parameters. *Integr. Biol.* 5, 796–806. doi: 10.1039/c3ib20221f

Dupaigne, P., Tonthat, N. K., Espéli, O., Whitfill, T., Boccard, F., and Schumacher, M. A. (2012). Molecular basis for a protein-mediated DNA-bridging mechanism that functions in condensation of the *E. coli* chromosome. *Mol. Cell* 48, 560–571. doi: 10.1016/j.molcel.2012.09.009

Duzdevich, D., Redding, S., and Greene, E. C. (2014). DNA dynamics and single-molecule biology. *Chem. Rev.* 114, 3072–3086. doi: 10.1021/cr4004117

El-Rami, F., Kong, X., Parikh, H., Zhu, B., Stone, V., Kitten, T., et al. (2018). Analysis of essential gene dynamics under antibiotic stress in streptococcus sanguinis. *Microbiology* 164, 173–185. doi: 10.1099/mic.0.000595

Faith, J. J., Driscoll, M. E., Fusaro, V. A., Cosgrove, E. J., Hayete, B., Juhn, F. S., et al. (2008). Many microbe microarrays database: uniformly normalized Affymetrix compendia with structured experimental metadata. *Nucleic Acids Res.* 36, D866–D870. doi: 10.1093/nar/gkm815

Frenkiel-Krispin, D., Ben-Avraham, I., Englander, J., Shimoni, E., Wolf, S. G., and Minsky, A. (2004). Nucleoid restructuring in stationary-state bacteria. *Mol. Microbiol.* 51, 395–405. doi: 10.1046/j.1365-2958.2003.03855.x

Goodhead, I., and Darby, A. C. (2015). Taking the pseudo out of pseudogenes. *Curr. Opin. Microbiol.* 23, 102–109. doi: 10.1016/j.mib.2014.11.012

Huerta, A. M., Salgado, H., Thieffry, D., and Collado-Vides, J. (1998). RegulonDB: a database on transcriptional regulation in *Escherichia coli*. *Nucleic Acids Res.* 26, 55–59. doi: 10.1093/nar/26.1.55

Jeong, K. S., Ahn, J., and Khodursky, A. B. (2004). Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol.* 5:R86. doi: 10.1186/gb-2004-5-11-r86

Jin, D. J., and Cabrera, J. E. (2006). Coupling the distribution of RNA polymerase to global gene regulation and the dynamic structure of the bacterial nucleoid in *Escherichia coli*. *J. Struct. Biol.* 156, 284–291. doi: 10.1016/j.jsb.2006.07.005

Junier, I., Frémont, P., and Rivoire, O. (2018). Universal and idiosyncratic characteristic lengths in bacterial genomes. *Phys. Biol.* 15:035001. doi: 10.1088/1478-3975/aab4ac

Junier, I., and Rivoire, O. (2016). Conserved units of co-expression in bacterial genomes: an evolutionary insight into transcriptional regulation. *PLoS One* 11:e0155740. doi: 10.1371/journal.pone.0155740

Junier, I., Unal, E. B., Yus, E., Lloréns-Rico, V., and Serrano, L. (2016). Insights into the mechanisms of basal coordination of transcription using a genome-reduced bacterium. *Cell Syst.* 2, 391–401. doi: 10.1016/j.cels.2016.04.015

Kang, Y., Weber, K. D., Qiu, Y., Kiley, P. J., and Blattner, F. R. (2005). Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function. *J. Bacteriol.* 187, 1135–1160. doi: 10.1128/JB.187.3.1135-1160.2005

Keseler, I. M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., et al. (2013). EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* 41, D605–D612. doi: 10.1093/nar/gks1027

Kim, J., Yoshimura, S. H., Hizume, K., Ohniwa, R. L., Ishihama, A., and Takeyasu, K. (2004). Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Res.* 32, 1982–1992. doi: 10.1093/nar/gkh512

Kluger, Y., Yu, H., Qian, J., and Gerstein, M. (2003). Relationship between gene co-expression and probe localization on microarray slides. *BMC Genomics* 4:49. doi: 10.1186/1471-2164-4-49

Koren, A., Tirosh, I., and Barkai, N. (2007). Autocorrelation analysis reveals widespread spatial biases in microarray experiments. *BMC Genomics* 8:164. doi: 10.1186/1471-2164-8-164

Krogh, T. J., Møller-Jensen, J., and Kaleta, C. (2018). Impact of chromosomal architecture on the function and evolution of bacterial genomes. *Front. Microbiol.* 9:2019. doi: 10.3389/fmicb.2018.02019

Kustatscher, G., Grabowski, P., and Rappsilber, J. (2017). Pervasive coexpression of spatially proximal genes is buffered at the protein level. *Mol. Syst. Biol.* 13:937. doi: 10.15252/msb.20177548

Lal, A., Dhar, A., Trostel, A., Kouzine, F., Seshasayee, A. S. N., and Adhya, S. (2016). Genome scale patterns of supercoiling in a bacterial chromosome. *Nat. Commun.* 7:11055. doi: 10.1038/ncomms11055

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Larsen, S. J., Röttger, R., Schmidt, H. H. H. W., and Baumbach, J. (2019). *E. coli* gene regulatory networks are inconsistent with gene expression data. *Nucleic Acids Res.* 47, 85–92. doi: 10.1093/nar/gky1176

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Liao, Y., Smyth, G. K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 47:e47. doi: 10.1093/nar/gkz114

Lioy, V. S., Cournac, A., Marbouty, M., Duigou, S., Mozziconacci, J., Espéli, O., et al. (2018). Multiscale structuring of the *E. coli* chromosome by nucleoid-associated and condensin proteins. *Cell* 172, 771–783.e18. doi: 10.1016/j.cell.2017.12.027

Liu, L. F., and Wang, J. C. (1987). Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U. S. A.* 84, 7024–7027. doi: 10.1073/pnas.84.20.7024

Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci.* 39, 447–462. doi: 10.1007/BF00648343

Ma, J., Bai, L., and Wang, M. D. (2013). Transcription under torsion. *Science* 340, 1580–1583. doi: 10.1126/science.1235441

Marr, C., Geertz, M., Hütt, M. T., and Muskhelishvili, G. (2008). Dissecting the logical types of network control in gene expression profiles. *BMC Syst. Biol.* 2:18. doi: 10.1186/1752-0509-2-18

Mathelier, A., and Carbone, A. (2010). Chromosomal periodicity and positional networks of genes in *Escherichia coli*. *Mol. Syst. Biol.* 6:366. doi: 10.1038/msb.2010.21

McAdams, H. H., Srinivasan, B., and Arkin, A. P. (2004). The evolution of genetic regulatory systems in bacteria. *Nat. Rev. Genet.* 5, 169–178. doi: 10.1038/nrg1292

Mondal, J., Bratton, B. P., Li, Y., Yethiraj, A., and Weisshaar, J. C. (2011). Entropy-based mechanism of ribosome-nucleoid segregation in *E. coli* cells. *Biophys. J.* 100, 2605–2613. doi: 10.1016/j.bpj.2011.04.030

Myers, K. S., Yan, H., Ong, I. M., Chung, D., Liang, K., Tran, F., et al. (2013). Genome-scale analysis of *Escherichia coli* FNR reveals complex features of transcription factor binding. *PLoS Genet.* 9:e1003565. doi: 10.1371/journal.pgen.1003565

Park, C., and Zhang, J. (2012). High expression hampers horizontal gene transfer. *Genome Biol. Evol.* 4, 523–532. doi: 10.1093/gbe/evs030

Postow, L., Hardy, C. D., Arsuaga, J., and Cozzarelli, N. R. (2004). Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev.* 18, 1766–1779. doi: 10.1101/gad.1207504

Pulkkinen, O., and Metzler, R. (2013). Distance matters: the impact of gene proximity in bacterial gene regulation. *Phys. Rev. Lett.* 110:198101. doi: 10.1103/PhysRevLett.110.198101

Rogozin, I. B. (2002). Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. *Nucleic Acids Res.* 30, 4264–4271. doi: 10.1093/nar/gkf549

Ruf, T. (1999). The Lomb-Scargle periodogram in biological rhythm research: analysis of incomplete and unequally spaced time-series. *Biol. Rhythm. Res.* 30, 178–201. doi: 10.1076/brhm.30.2.178.1422

Saecker, R. M., Record, M. T., and Dehaseth, P. L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.* 412, 754–771. doi: 10.1016/j.jmb.2011.01.018

Salmon, K., Hung, S.-P., Mekjian, K., Baldi, P., Hatfield, G. W., and Gunsalus, R. P. (2003). Global gene expression profiling in *Escherichia coli* K12: the effects of oxygen availability and FNR. *J. Biol. Chem.* 278, 29837–29855. doi: 10.1074/jbc.M213060200

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sánchez-Pérez, M., Gómez-Romero, L., Ledezma-Tejeida, D., et al. (2019). RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.* 47, D212–D220. doi: 10.1093/nar/gky1077

Sauer, C., Syvertsson, S., Bohorquez, L. C., Cruz, R., Harwood, C. R., van Rij, T., et al. (2016). Effect of genome position on heterologous gene expression in *Bacillus subtilis*: an unbiased analysis. *ACS Synth. Biol.* 5, 942–947. doi: 10.1021/acssynbio.6b00065

Scargle, J. D. (1982). Studies in astronomical time series analysis. II - statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.* 263:835. doi: 10.1086/160554

Scholz, S. A., Diao, R., Wolfe, M. B., Fivenson, E. M., Lin, X. N., and Freddolino, P. L. (2019). High-resolution mapping of the *Escherichia coli* chromosome reveals positions of high and low transcription. *Cell Syst.* 8, 212–225.e9. doi: 10.1016/j.cels.2019.02.004

Schvartzman, J. B., Martínez-Robles, M. L., Hernández, P., and Krimer, D. B. (2013). The benefit of DNA supercoiling during replication. *Biochem. Soc. Trans.* 41, 646–651. doi: 10.1042/BST20120281

Seward, E. A., and Kelly, S. (2018). Selection-driven cost-efficiency optimization of transcripts modulates gene evolutionary rate in bacteria. *Genome Biol.* 19:102. doi: 10.1186/s13059-018-1480-7

Sharma, A., and Mondragón, A. (1995). DNA topoisomerases. *Curr. Opin. Struct. Biol.* 5, 39–47. doi: 10.1016/0959-440X(95)80007-N

Sobetzko, P. (2016). Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. *Nucleic Acids Res.* 44, 1514–1524. doi: 10.1093/nar/gkw007

Stein, R. A., Deng, S., and Higgins, N. P. (2005). Measuring chromosome dynamics on different time scales using resolvases with varying half-lives. *Mol. Microbiol.* 56, 1049–1061. doi: 10.1111/j.1365-2958.2005.04588.x

Svetlov, V., and Nudler, E. (2013). Looking for a promoter in 3D. *Nat. Struct. Mol. Biol.* 20, 141–142. doi: 10.1038/nsmb.2498

Talukder, A. A., and Ishihama, A. (2015). Growth phase dependent changes in the structure and protein composition of nucleoid in *Escherichia coli*. *Sci. China Life Sci.* 58, 902–911. doi: 10.1007/s11427-015-4898-0

Tang, S., Orsi, R. H., den Bakker, H. C., Wiedmann, M., Boor, K. J., and Bergholz, T. M. (2015). Transcriptomic analysis of the adaptation of *Listeria monocytogenes* to growth on vacuum-packed cold smoked salmon. *Appl. Environ. Microbiol.* 81, 6812–6824. doi: 10.1128/AEM.01752-15

Valens, M., Penaud, S., Rossignol, M., Cornet, F., and Boccard, F. (2004). Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J.* 23, 4330–4341. doi: 10.1038/sj.emboj.7600434

VanderPlas, J. T. (2018). Understanding the Lomb–Scargle periodogram. *Astrophys. J. Suppl. Ser.* 236:16. doi: 10.3847/1538-4365/aab766

Viñuelas, J., Calevro, F., Remond, D., Bernillon, J., Rahbé, Y., Febvay, G., et al. (2007). Conservation of the links between gene transcription and chromosomal organization in the highly reduced genome of Buchnera aphidicola. *BMC Genomics* 8:143. doi: 10.1186/1471-2164-8-143

Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturino, J. M., et al. (2010). Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1, 147–149. doi: 10.1038/ncomms1146

Wang, F., Redding, S., Finkelstein, I. J., Gorman, J., Reichman, D. R., and Greene, E. C. (2013). The promoter-search mechanism of *Escherichia coli* RNA polymerase is dominated by three-dimensional diffusion. *Nat. Struct. Mol. Biol.* 20, 174–181. doi: 10.1038/nsmb.2472

Weng, X., and Xiao, J. (2014). Spatial organization of transcription in bacterial cells. *Trends Genet.* 30, 287–297. doi: 10.1016/j.tig.2014.04.008

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Woldringh, C. L., Jensen, P. R., and Westerhoff, H. V. (1995). Structure and partitioning of bacterial DNA: determined by a balance of compaction and expansion forces? *FEMS Microbiol. Lett.* 131, 235–242. doi: 10.1016/0378-1097(95)00243-X

Wright, M. A., Kharchenko, P., Church, G. M., and Segrè, D. (2007). Chromosomal periodicity of evolutionarily conserved gene pairs. *Proc. Natl. Acad. Sci. U. S. A.* 104, 10559–10564. doi: 10.1073/pnas.0610776104

Xiao, G., Wang, X., and Khodursky, A. B. (2011). Modeling three-dimensional chromosome structures using gene expression data. *J. Am. Stat. Assoc.* 106, 61–72. doi: 10.1198/jasa.2010.ap0950

Zimmerman, S. B. (2002). Toroidal nucleoids in *Escherichia coli* exposed to chloramphenicol. *J. Struct. Biol.* 138, 199–206. doi: 10.1016/S1047-8477(02)00036-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 02 October 2020
doi: 10.3389/fmicb.2020.585717

DNA and Polyphosphate in Directed Proteolysis for DNA Replication Control

*Malgorzata Ropelewska, Marta H. Gross† and Igor Konieczny**

Laboratory of Molecular Biology, Intercollegiate Faculty of Biotechnology of University of Gdańsk and Medical University of Gdańsk, Gdańsk, Poland

The strict control of bacterial cell proliferation by proteolysis is vital to coordinate cell cycle processes and to adapt to environmental changes. ATP-dependent proteases of the AAA + family are molecular machineries that contribute to cellular proteostasis. Their activity is important to control the level of various proteins, including those that are essential for the regulation of DNA replication. Since the process of proteolysis is irreversible, the protease activity must be tightly regulated and directed toward a specific substrate at the exact time and space in a cell. In our mini review, we discuss the impact of phosphate-containing molecules like DNA and inorganic polyphosphate (PolyP), accumulated during stress, on protease activities. We describe how the directed proteolysis of essential replication proteins contributes to the regulation of DNA replication under normal and stress conditions in bacteria.

Keywords: DNA replication, replication initiators, proteolysis, polyphosphate, lon protease

OPEN ACCESS

Edited by:
Torsten Waldminghaus, University of Marburg, Germany

Reviewed by:
Dhruba Chattoraj, National Institutes of Health (NIH), United States
Kristina Jonas, Stockholm University, Sweden
Ulf Gerth, University of Greifswald, Germany

***Correspondence:**
Igor Konieczny
igor.konieczny@ug.edu.pl

†Present address:
Marta H. Gross, Chromosome Replication Laboratory, The Francis Crick Institute, London, United Kingdom

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *21 July 2020*
Accepted: *10 September 2020*
Published: *02 October 2020*

Citation:
Ropelewska M, Gross MH and Konieczny I (2020) DNA and Polyphosphate in Directed Proteolysis for DNA Replication Control. Front. Microbiol. 11:585717. doi: 10.3389/fmicb.2020.585717

INTRODUCTION

Several mechanisms responsible for the control of DNA replication in bacteria were described (Zakrzewska-Czerwińska et al., 2007). Most of those mechanisms aim at decreasing the availability of active replication protein, e.g., by regulating the transcription (Gora et al., 2013), spatial sequestration (Iniesta et al., 2006), or protein inactivation (Kurokawa et al., 1999). It was shown that particular bacterial proteases are involved in the proteolysis of replication proteins and proteins associated with the process of DNA replication (Wickner et al., 1994; Pierechod et al., 2009; Kubik et al., 2012; Karlowicz et al., 2017). The major proteases in bacteria belong to the family of ATPases associated with diverse cellular activities (AAA +). In *Escherichia coli*, there are four cytosolic proteases (i.e., ClpXP, ClpAP, HslUV, and Lon) (Gottesman, 2003). Bacterial AAA + proteases function efficiently under different growth conditions participating in regulation of several cellular processes. For instance, the intracellular levels of the HslUV protease are increased under heat-shock conditions when it has the maximum substrate degradation rate (Burton et al., 2005). In addition to HslUV functions under thermal stress, this protease plays an important role in SOS response caused by DNA damage (Khattar, 1997) and in response to acidic stress (Kannan et al., 2008). ClpXP participates in the response to starvation (Schweder et al., 1996), heat shock, and oxidative stress (Frees et al., 2003). Similarly, ClpAP protease is responsible for the control of regulatory pathways in bacteria and response to proteotoxic stress caused by pH downshift or high temperature (Jenal and Hengge-Aronis, 2003). Lon protease contributes to genome maintenance during stress (e.g., heat shock or nutrient depletion) by regulating DNA replication (Nicoloff et al., 2007; Jonas et al., 2013; Leslie et al., 2015; Gross and Konieczny, 2020). Furthermore, LonA protease is involved in the tolerance of *Actinobacillus pleuropneumoniae* to osmotic or oxidative stress

(Xie et al., 2016). Since proteolysis is irreversible, it must be induced at particular conditions and target specific proteins in a tightly controlled manner. Bacterial AAA + proteases are regulated temporally (Goff and Goldberg, 1987; Jonas et al., 2013), spatially (Simmons et al., 2008), and structurally (Jonas et al., 2013) and by interaction with ligand or adaptors (Goldberg et al., 1980; Wah et al., 2002; Martin et al., 2008; Puri, 2016). Proteases interact with various phosphate-containing molecules including membrane components [e.g., lipopolysaccharide (LPS) (Sugiyama et al., 2013) and cardiolipin (CL) (Minami et al., 2011)], stress-induced factors [e.g., guanosine tetraphosphate ((p)ppGpp) (Osbourne et al., 2014) and inorganic polyphosphate (PolyP) (Kuroda, 2006)], ATP (Charette et al., 1981), and ADP (Waxman and Goldberg, 1985) as well as with DNA (Zehnbauer et al., 1981; Zylicz et al., 1998; Kubik et al., 2012). The protease binding to phosphate-containing molecules may change protease localization, ATPase activity, or substrate specificity, thereby modulating its proteolytic activity (Kubik et al., 2012; Karlowicz et al., 2017; Gross and Konieczny, 2020).

THE IMPACT OF DNA BINDING ON PROTEASE ACTIVITY

In *Escherichia coli*, only Lon and ClpAP, but not ClpXP or HslUV, interact with DNA (Kubik et al., 2012). Interaction of Lon with nucleic acid is a conserved property among species (Zehnbauer et al., 1981; Fu and Markovitz, 1998; Lee et al., 2004; Lu et al., 2003). It was demonstrated that the α subdomain in the AAA + module of *Brevibacillus thermoruber* Lon is involved in DNA binding (Lee et al., 2004, 2014; Lin et al., 2009). In various organisms, Lon has different preference for the type of DNA with which it forms a complex. *E. coli* Lon binds to double-stranded DNA (dsDNA) in a sequence-non-specific manner (Charette et al., 1984; Nomura et al., 2004). On the contrary, eukaryotic proteases bind single-stranded DNA (ssDNA) or RNA (Fu and Markovitz, 1998; Lu et al., 2003; Liu et al., 2004). *Bacillus subtilis* LonA is present in the nucleoid under normal growth conditions, while ClpXP is present in cytosol (Simmons et al., 2008). During spore development, LonA changes its localization to the forespore (Simmons et al., 2008). Under heat shock, LonA remains bound to the nucleoid (Simmons et al., 2008). Yet in *E. coli* when temperature is increased, Lon loses its ability to bind DNA *in vitro*, although ATP-dependent proteolytic activity is retained (Sonezaki et al., 1995). It is proposed that the Lon presence within the nucleoid allows for the degradation of DNA-associated proteins involved in DNA metabolism. The protease dissociation from DNA upon stress-related factors may provide rapid adaptive mechanism to hamper Lon activity toward specific proteins (Sonezaki et al., 1995).

The interaction of DNA with Lon stimulates its ATPase activity (Charette et al., 1984). At the surface of *E. coli* Lon ATPase domain, there are located positively charged residues, which are responsible for direct interaction with DNA (Karlowicz et al., 2017). The presence of DNA in a reaction mixture containing Lon and substrate protein enhances protease activity to hydrolyze ATP (Karlowicz et al., 2017). The ATPase activity of Lon mutant defective in DNA interaction is not increased in the presence of substrate and DNA. Hence, it is the direct DNA–Lon interaction that stimulates protease ATPase activity (Karlowicz et al., 2017). It was also demonstrated that Lon nucleoprotein complex formation is essential for the proteolysis of DNA-interacting substrates, but not other substrates (Karlowicz et al., 2017).

The ClpAP proteolysis of DNA-binding substrates is also stimulated by DNA. For example, ParD protein, the component of toxin–antitoxin system of RK2 plasmid (Kubik et al., 2012; Dubiel et al., 2018) is degraded by ClpAP in a DNA-dependent manner (Dubiel et al., 2018). *In vitro* experiments suggest that it is the protease—DNA interaction, but not substrate–DNA interaction, that contributes to the enhanced proteolysis. Although *E. coli* ClpXP and HslUV do not form nucleoprotein complexes, the addition of DNA to the *in vitro* reaction mixture affects the proteolysis of particular substrates (Kubik et al., 2012). As opposed to Lon and ClpAP, the process of proteolysis is inhibited by DNA. This may be explained by the ability of substrates to interact directly with DNA, thus hampering their proteolysis.

THE IMPACT OF POLYPHOSPHATE BINDING ON PROTEASE ACTIVITY

When bacteria encounter stress such as amino acid starvation or oxidative stress, they accumulate inorganic PolyP, which forms granular superstructures and contributes to cell survival (Kuroda et al., 2001). The production of PolyP was initially correlated with the synthesis of second messenger stress molecule, (p)ppGpp, which was shown to inhibit the activity of exopolyphosphatase (PPX), thereby enabling uncontrolled production of PolyP by PolyP kinase (PPK) (Kuroda et al., 1997; Magnusson et al., 2005; Traxler et al., 2008; Rao et al., 2009). *Ppk* mutants fail to survive in stationary phase and are less resistant to heat or oxidants (Crooke et al., 1994; Rao and Kornberg, 1996). Recent data argue that (p)ppGpp is not required for PolyP synthesis and that transcription factor DksA contributes to the control of PolyP level instead (Gray, 2019). In *Caulobacter crescentus*, PolyP has been shown to be involved in the regulation of DNA replication during carbon starvation (Boutte et al., 2012). During nitrogen starvation in *Pseudomonas aeruginosa*, PolyP granule biogenesis is temporally and functionally tied to cell cycle exit indicated by the inhibition of reinitiation of DNA replication, completion of open rounds of DNA replication, segregation of daughter chromosomes, and septation (Racki et al., 2017). PolyP interacts with *Escherichia coli* Lon via ATPase domain (Nomura et al., 2004), as in the case of DNA (Karlowicz et al., 2017), which implies that both phosphate-containing molecules can compete for Lon binding. Indeed, the equimolar concentration of PolyP was shown to disrupt the Lon–DNA complex (Nomura et al., 2004) and Lon colocalization with nucleoid (Zhao et al., 2008). Lon loses DNA-binding ability when cells are exposed to heat-shock conditions, which is directly linked to an increase in the amount of damaged proteins (Sonezaki et al., 1995). During starvation, Lon is associated with PolyP granules (Kuroda, 2006). PolyP stimulates Lon to proteolyze ribosomal proteins such

as L1, L3, and L24 but inhibits proteolysis of SulA protein (an inhibitor of cell division accumulated in response to DNA damage) (Nomura et al., 2004). When Lon is pre-incubated with PolyP, the proteolysis of L24 ribosomal protein is the most efficient (Nomura et al., 2004). Not all PolyP-interacting proteins are degraded by Lon, but all proteins degraded by Lon in a PolyP-dependent manner do form a complex with PolyP (Kuroda, 2006).

Although a complex of protease with PolyP and its general role was uncovered almost two decades ago, we still lack the full mechanistic and physiological insight into this complex formation. To date, no data are available on how/if PolyP affects other proteases in bacterial cells.

THE PROTEOLYSIS OF REPLICATION PROTEINS AND PROTEINS ASSOCIATED WITH DNA REPLICATION

Not only proteases but also their substrate can interact with DNA or PolyP. Depending on the substrate, the process of proteolysis is specifically controlled and fine-tuned (**Table 1**). Here, we discuss the proteolysis of selected replication factors and how it affects cell survival.

Replication Initiators

The replication initiation proteins are the prerequisite factors responsible for initiating DNA replication in various replicons; thus, their degradation allows for rapid arrest of DNA replication. The DnaA, a highly conserved replication initiation protein in bacteria, is an obvious target for cellular proteases. In *Caulobacter crescentus*, DnaA protein is degraded mainly by Lon, under optimal and stress conditions (Gorbatyuk and Marczynski, 2005; Jonas et al., 2013; Leslie et al., 2015; Liu et al., 2016). It was demonstrated that the DnaA intracellular levels depend on a reduction in DnaA synthesis and fast degradation by the Lon protease. Constitutively, ATP-bound DnaA mutant was shown to be degraded more slowly than wild-type (wt) protein, indicating that degradation of DnaA is linked to DnaA activity or DnaA nucleotide bound state (Liu et al., 2016). Under proteotoxic stress, DnaA is degraded as a result of allosteric activation of Lon by accumulated unfolded substrates and increase in Lon intracellular concentration (Jonas et al., 2013; **Figure 1A**). Under normal growth conditions, *C. crescentus* DnaA is proteolyzed at the end of S-phase to ensure that only newly synthesized DnaA is available at the start of each replication round (Jenal, 2009). The overexpression of ClpA in *lon*-depleted strain restores DnaA degradation, indicating that fail-safe systems are present (Liu et al., 2016).

The regulatory mechanism that controls DNA replication in *Escherichia coli* by directed proteolysis of replication initiator was termed PolyP-induced DnaA proteolysis (PDAP) (Gross and Konieczny, 2020; **Figure 1B**). In *E. coli* cells during amino acid starvation, PolyP induces Lon activity to specifically degrade, DnaA when bound to ADP, but not ATP. When PolyP-synthesizing enzyme (PPK) or Lon protease is depleted in *E. coli* during stress, DnaA level remains high. Also, the level of DnaA protein variant permanently bound to ATP does not change in stress conditions (Gross and Konieczny, 2020). Both *in vivo* and *in vitro* data indicate that when DnaA is converted to ADP-bound form, it is degraded by Lon (Gross and Konieczny, 2020). PolyP interacts with DnaA-ADP, but not DnaA-ATP, which provides an explanation on how Lon targets only DnaA-ADP for proteolysis. In starvation, as a result of an increase in Lon level and Lon activation by PolyP, the overall DnaA concentration decreases, which leads to the inhibition of DNA replication initiation (Gross and Konieczny, 2020). Since in *E. coli* (Gross and Konieczny, 2020) and in *C. crescentus* (Liu et al., 2016) DnaA protein degradation depends on its nucleotide state, it may be crucial for the control of DNA replication. Such possibility is discussed in a recent review on the regulation of *Caulobacter* DnaA (Felletti et al., 2019). It was also shown that in stress in *E. coli*, ppGpp affects RNA polymerase activity and thereby superhelicity of replication origin, which leads to DNA replication initiation inhibition (Kraemer et al., 2019). Because ppGpp is not required for PolyP synthesis in *E. coli* (Crooke et al., 1994), it is very likely that the regulations by ppGpp (Kraemer et al., 2019) and PDAP (Gross and Konieczny, 2020) are independent mechanisms responsible for controlling DNA replication initiation during stress in *E. coli*.

DnaA participates in the replication initiation of many plasmids, which implies that the replication of plasmid and chromosome in one cell may be coordinately regulated by the inducible degradation of DnaA during stress conditions. This possibility requires to be investigated. It was shown that stability of plasmid DNA is decreased in *E. coli* protease-deficient mutants (Bury et al., 2017; Dubiel et al., 2018). Plasmid-encoded replication initiators (Rep), e.g., RK2 plasmid TrfA protein, are degraded by Lon and other cytosolic proteases (Wojtkowiak et al., 1993; Wickner et al., 1994; Levchenko et al., 1995; Pierechod et al., 2009; Kubik et al., 2012). The selective proteases activity may affect Rep monomer/dimer ratio and therefore the ability of replication initiator to initiate plasmid DNA replication. DNA stimulates TrfA degradation by Lon (**Figure 1C**) and ClpAP but inhibits proteolysis by ClpXP and HslUV (Kubik et al., 2012). Similarly, binding of λO protein, i.e., replication initiator of bacteriophage Lambda, to *ori*λ DNA protects it from degradation by ClpXP (Zylicz et al., 1998). Despite replication initiation control by the Rep concentration and monomer/dimer ratio, the RK2 plasmid replication is also controlled by joining two DNA plasmid particles via TrfA to form handcuff complex, thereby preventing replication reinitiation. *E. coli* Lon disrupts the handcuff complex by proteolyzing TrfA (Bury et al., 2017).

CtrA

The response regulator CtrA in *C. crescentus* is another DNA-binding protein whose level is controlled by proteases. CtrA not only controls transcription of more than a hundred genes (Wojtkowiak et al., 1993) but also inhibits DNA replication initiation (Quon et al., 1996, 1998; Laub et al., 2002). For replication to occur, CtrA must be eliminated at the G1–S transition, and this is carried out by dephosphorylation (Jacobs et al., 2003) and ClpXP-mediated proteolysis (Jenal and Fuchs, 1998). Under nutritional stress, CtrA proteolysis is inhibited by ppGpp and PolyP

TABLE 1 | Comparison of degradation conditions of replication proteins and proteins associated with DNA replication.

Substrate	Function of a substrate	Organism	Protease	Factors affecting the proteolysis	References
DnaA	Required for bacterial DNA replication initiation	*Caulobacter crescentus*	Lon, ClpAP	Unfolded substrates (+)	Jonas et al., 2013; Liu et al., 2016
		Escherichia coli	Lon	PolyP (+)[a]	Gross and Konieczny, 2020
TrfA–wt (dimer)	Participates in the formation of "handcuff" of RK2 plasmid particles	*E. coli*	ClpAP, Lon	DNA (+)[b]	Kubik et al., 2012
			ClpXP, HslUV	DNA (–)	Kubik et al., 2012
TrfA G254D/S256L (monomer)	Participates in replication initiation of RK2 plasmid	*E. coli*	Lon, ClpAP	DNA (+)	Kubik et al., 2012
RepE	Participates in replication initiation of F plasmid	*E. coli*	Lon	DNA (+)	Karlowicz et al., 2017
λO	Participates in replication initiation of phage lambda	*E. coli*	Lon	DNA (+)	Karlowicz et al., 2017
HimA	As a heterodimer with HimD bends DNA in the structure of oriC, thus facilitating the replication initiation	*E. coli*	Lon	PolyP (+)	Nomura et al., 2004
Dps	Protects DNA during starvation and oxidative stress	*E. coli*	ClpAP, ClpXP	?[c]	Stephani et al., 2003
CspD	Inhibits DNA replication; plays a regulatory role in chromosomal replication in nutrient-depleted cells	*E. coli*	Lon	?	Langklotz and Narberhaus, 2011
CtrA	Controls transcription and inhibits DNA replication initiation.	*C. crescentus*	ClpXP	?	Jenal and Fuchs, 1998
CcrM	Inhibits DNA replication initiation	*C. crescentus*	Lon	DNA (+)	Gonzalez et al., 2014
DnaX	Participates in the loading of sliding clamp	*C. crescentus*	ClpXP	?	Vass and Chien, 2013
SocB	Binds to sliding clamp and inhibits elongation of DNA replication	*C. crescentus*	ClpXP	?	Aakre et al., 2013

[a] *(+) indicates that the proteolysis is stimulated by DNA, PolyP, or unfolded substrates.*
[b] *(–) indicates that the proteolysis is inhibited by DNA, PolyP, or unfolded substrates.*
[c] *? indicates that there are no data about impact of DNA, PolyP, or unfolded substrates.*

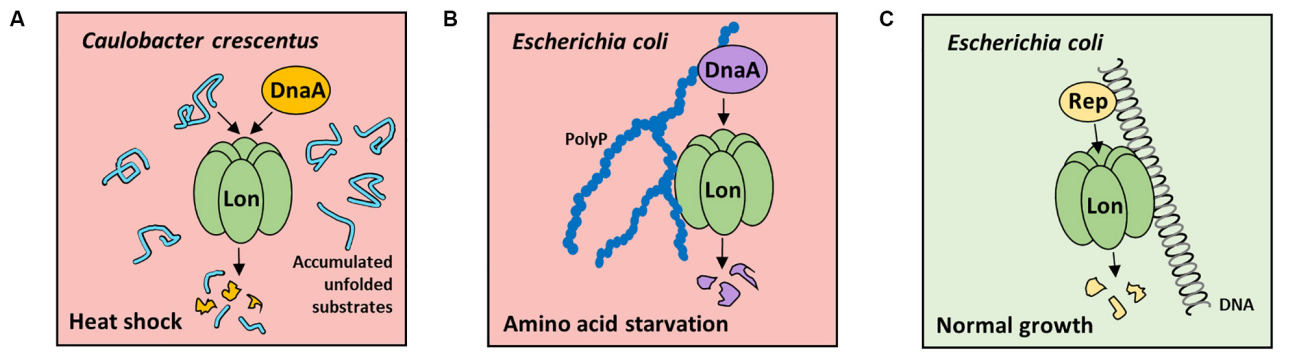


FIGURE 1 | Factors stimulating Lon-dependent proteolysis of replication initiators in *Caulobacter crescentus* and *Escherichia coli* during stress or normal growth conditions. **(A)** In *C. crescentus* in heat-shock conditions, unfolded substrate accumulation stimulate Lon protease for DnaA degradation, which results in inhibition of DNA replication initiation. **(B)** In *E. coli* cells during amino acid starvation, PolyP-induced DnaA proteolysis (PDAP) is launched. PolyP activates Lon protease to degrade DnaA, thereby resulting in the decreasing DnaA level and, consequently, DNA replication initiation arrest. **(C)** Under normal growth conditions, plasmid replication initiation protein (Rep) degradation by Lon is induced by nucleoprotein complex formation. The protease and substrate interaction with DNA is crucial for efficient degradation. No data are available about Rep proteins stability in stress conditions.

accumulation (Boutte et al., 2012). The proteolysis of CtrA is carried out by ClpXP only when both proteins are localized in the cell pole (Iniesta et al., 2006). This process occurs in the presence of accessory proteins, i.e., CpdR, RcdA, PopA, and cyclic diguanylate (cdG), which accelerate CtrA degradation *in vitro*. Those accessory proteins are also essential for proteolysis of CtrA bound to DNA (Smith et al., 2014).

CcrM

In order to complete cell division, the chromosome needs to be fully methylated by the CcrM DNA methyltransferase. This

methyltransferase CcrM is proteolyzed by Lon to restrict CcrM to most of the cell cycle that prolongs the hemimethylation state of chromosomal DNA during DNA synthesis in *C. crescentus* (Wright et al., 1996). The *ccrM* gene transcription is regulated by a positive global regulator CtrA, and the CcrM protein is constitutively degraded by Lon (Wright et al., 1996). Not only DNA was shown to stimulate Lon-mediated proteolysis of CcrM but also CcrM has 10-fold higher affinity for Lon in the presence of DNA, when compared with CcrM to Lon alone (Zhou et al., 2019). The C-terminus of CcrM binds DNA and is recognized by Lon (Zhou et al., 2019). Lon interaction with DNA is not crucial for CcrM proteolysis because CcrM degradation is still observed in cells expressing Lon mutant defective in DNA binding (Zeinert et al., 2018). Therefore, the CcrM level and correct completion of cell cycle depend on the balance between the synthesis and proteolysis of CcrM. CcrM degradation by Lon can also affect the dNTP production in a cell. In Δ*lon* strains, an increase in the ribonucleotide reductase (RNR) expression level is observed, which is driven by stabilization of the transcription factor CcrM (Zeinert et al., 2020).

Integration Host Factor

The integration host factor (IHF) (Nomura et al., 2004) is a histone-like protein responsible for modulation of the DNA condensation (Pettijohn, 1988). IHF is a HimA/HimD heterodimer, which interacts with DNA through specific binding sequence (IBS, IHF binding sequence) and bends DNA in the structure of *oriC*, thus facilitating the process of replication initiation in *E. coli* (Ozaki and Katayama, 2012). IHF also participates in regulating the nucleotide state of DnaA. IHF dimers bound to *datA* sequence promote DnaA-ATP hydrolysis in the DDAH system, thus increasing the pool of DnaA-ADP to prevent overinitiation (Kasho and Katayama, 2013). Moreover, IHF, together with Fis, binds to DARS2 sequence and participates in DnaA-ATP regeneration, which is coupled to cell cycle and growth phase (Kasho et al., 2014). IHF interacts with PolyP (Kornberg, 1995), and its level is regulated by Lon in a PolyP-dependent manner (Nomura et al., 2004). The IHF oligomeric state has an impact on this process. HimA degradation is dependent on PolyP and Lon, as opposed to HimD. When both monomers formed heterodimers, neither HimA nor HimD is degraded (Nomura et al., 2004). This suggests that either Lon recognition for HimA is buried at the interface of monomers within heterodimer or a significant structural rearrangement occurs upon dimerization.

CspD

Upon entry into the stationary phase in *E. coli*, CspD is expressed and acts as an inhibitor of replication (Yamanaka and Inouye, 1997). Expression of CspD was shown to be activated by (p)ppGpp (Yamanaka and Inouye, 1997). This allows for the adaptation to nutritional changes. CspD was found to be related to persister cell formation (Kim and Wood, 2010). Cellular level of CspD is regulated in response to growth phase and growth rate by proteolysis. Using electron microscopy (EM), it was shown that CspD condenses ssDNA; however, those nucleoprotein complexes are distinct from the complex of single-stranded binding protein (SSB) with DNA (Yamanaka et al., 2001). When growth is resumed in nutrient-rich environment, CspD is degraded by Lon (Langklotz and Narberhaus, 2011). The proteolysis of CspD by Lon was reconstituted *in vitro* and did not require any additives, besides ATP, which indicates that during growth, unknown factors must regulate either Lon activity or CspD availability for degradation.

Dps

Known as the most abundant protein in a stationary phase in *E. coli*, Dps was shown to protect DNA during starvation and oxidative stress, by self-aggregation and DNA condensation (Almirón et al., 1992; Azam et al., 1999; Ceci et al., 2004; Frenkiel-Krispin et al., 2004; Melekhov et al., 2015). During DNA damage, Dps also interacts with DnaA in order to delay replication initiation and allow for DNA repair (Chodavarapu et al., 2008). ClpAP and ClpXP degrade Dps during the exponential phase, which leads to a significant reduction in Dps level (Ninnis et al., 2009). Considering the involvement of Dps in various important functions, its level must be tightly controlled.

SocB

Sliding clamp (a protein responsible for the replisome processivity in DNA replication) is inhibited by SocB, a component of SocB toxin–SocA antitoxin system in *C. crescentus* (Aakre et al., 2013). The SocB is unstable and constitutively proteolyzed by ClpXP in the presence of SocA. SocB interacts with sliding clamp and inhibits elongation of DNA replication, presumably by outcompeting other proteins from binding sliding clamp. The excessive sliding clamp occupation by SocB leads to premature collapse of replication fork and incomplete cell cycle (Aakre et al., 2013).

DnaX

For the sliding clamp to be loaded onto DNA, a clamp loader complex is required. In *E. coli*, this complex contains tau and gamma subunits, which are produced from the same gene, but gamma is shorter due to ribosomal frameshifting (Lee and Walker, 1987). In *C. crescentus*, which lacks a frameshifting site, ClpXP generates the shorter version, i.e., gamma subunit, which is necessary under normal growth conditions as well as for DNA damage tolerance (Vass and Chien, 2013).

CONCLUSION AND PERSPECTIVES

In this review, we highlight that directed proteolysis can be stimulated by protease interaction with phosphate-containing molecules such as DNA and PolyP. To date, no structural data are available on such complexes. This specific interaction affects protease activity and selectivity against substrates especially those important in regulation of DNA replication. The current knowledge indicates that among all cytosolic proteases, Lon plays the most important role in the regulation of DNA replication in bacterial cells. We propose that during normal growth,

it is the nucleoid DNA that provides matrix for Lon and its substrate proteins. During stress, Lon binds to PolyP granules, thereby stimulating degradation of substrates, which also interact with PolyP in stress. The exact molecular mechanism for this activation remains to be elucidated and needs further validation. Application of the cutting-edge structural research, single-molecule experiments, and trapping approach (Aubin-Tam et al., 2011; Arends et al., 2018; Hu et al., 2018; Fei et al., 2020) should provide insight into the structure–function relationship of Lon, its substrates, adaptor proteins, and complexes with phosphate-containing molecules. Growing evidence indicates that proteolysis is crucial for virulence in many pathogens (Butler et al., 2006; Ingmer and Brøndsted, 2009; Willett et al., 2015). Understanding how directed proteolysis is regulated by phosphate-containing molecules will give insight into microbial stress responses and the regulation of DNA replication.

AUTHOR CONTRIBUTIONS

MR, MG, and IK compiled the concept of the mini review. MR and MG wrote the initial version of text. MR and IK prepared the figure and table and wrote the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Foundation for Polish Science (TEAM, POIR.04.04.00-00-5C75/17-00). Funding for open access charge: Fundacja na rzecz Nauki Polskiej (TEAM, POIR.04.04.00-00-5C75/17-00).

REFERENCES

Aakre, C. D., Phung, T. N., Huang, D., and Laub, M. T. (2013). A bacterial toxin inhibits DNA replication elongation through a direct interaction with the β sliding clamp. *Mol. Cell.* 52, 617–628. doi: 10.1016/j.molcel.2013.10.014

Almirón, M., Link, A. J., Furlong, D., and Kolter, R. (1992). A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev.* 6, 2646–2654. doi: 10.1101/gad.6.12b.2646

Arends, J., Griego, M., Thomanek, N., Lindemann, C., Kutscher, B., Meyer, H. E., et al. (2018). An integrated proteomic approach uncovers novel substrates and functions of the Lon protease in *Escherichia coli*. *Proteomics* 18:1800080. doi: 10.1002/pmic.201800080

Aubin-Tam, M. E., Olivares, A. O., Sauer, R. T., Baker, T. A., and Lang, M. J. (2011). Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* 145, 257–267. doi: 10.1016/j.cell.2011.03.036

Azam, T. A., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* 181, 6361–6370. doi: 10.1128/jb.181.20.6361-6370.1999

Boutte, C. C., Henry, J. T., and Crosson, S. (2012). ppGpp and polyphosphate modulate cell cycle progression in *Caulobacter crescentus*. *J. Bacteriol.* 194, 28–35. doi: 10.1128/jb.05932-11

Burton, R. E., Baker, T. A., and Sauer, R. T. (2005). Nucleotide-dependent substrate recognition by the AAA+ HslUV protease. *Nat. Struct. Mol. Biol.* 12, 245–251. doi: 10.1038/nsmb898

Bury, K., Wegrzyn, K., and Konieczny, I. (2017). Handcuffing reversal is facilitated by proteases and replication initiator monomers. *Nucleic Acids Res.* 45, 3953–3966. doi: 10.1093/nar/gkx166

Butler, S. M., Festa, R. A., Pearce, M. J., and Darwin, K. H. (2006). Self-compartmentalized bacterial proteases and pathogenesis. *Mol. Microbiol.* 60, 553–562. doi: 10.1111/j.1365-2958.2006.05128.x

Ceci, P., Cellai, S., Falvo, E., Rivetti, C., Rossi, G. L., and Chiancone, E. (2004). DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res.* 32, 5935–5944. doi: 10.1093/nar/gkh915

Charette, M. F., Henderson, G. W., Doane, L. L., and Markovitz, A. (1984). DNA-stimulated ATPase activity on the lon (CapR) protein. *J. Bacteriol.* 158, 195–201. doi: 10.1128/jb.158.1.195-201.1984

Charette, M. F., Henderson, G. W., and Markovitz, A. (1981). ATP hydrolysis-dependent protease activity of the lon (capR) protein of *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. U.S.A.* 78, 4728–4732. doi: 10.1073/pnas.78.8.4728

Chodavarapu, S., Gomez, R., Vicente, M., and Kaguni, J. M. (2008). *Escherichia coli* Dps interacts with DnaA protein to impede initiation: a model of adaptive mutation. *Mol. Microbiol.* 67, 1331–1346. doi: 10.1111/j.1365-2958.2008.06127.x

Crooke, E., Akiyama, M., Rao, N. N., and Kornberg, A. (1994). Genetically altered levels of inorganic polyphosphate in *Escherichia coli*. *J. Biol. Chem.* 269, 6290–6295.

Dubiel, A., Wegrzyn, K., Kupinski, A. P., and Konieczny, I. (2018). ClpAP protease is a universal factor that activates the parDE toxin-antitoxin system from a broad host range RK2 plasmid. *Sci. Rep.* 8, 1–12.

Fei, X., Bell, T. A., Jenni, S., Stinson, B. M., Baker, T. A., Harrison, S. C., et al. (2020). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *eLife* 9:e52774.

Felletti, M., Omnus, D. J., and Jonas, K. (2019). Regulation of the replication initiator DnaA in *Caulobacter crescentus*. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 697–705. doi: 10.1016/j.bbagrm.2018.01.004

Frees, D., Qazi, S. N., Hill, P. J., and Ingmer, H. (2003). Alternative roles of ClpX and ClpP in *Staphylococcus aureus* stress tolerance and virulence. *Mol. Microbiol.* 48, 1565–1578. doi: 10.1046/j.1365-2958.2003.03524.x

Frenkiel-Krispin, D., Ben-Avraham, I., Englander, J., Shimoni, E., Wolf, S. G., and Minsky, A. (2004). Nucleoid restructuring in stationary-state bacteria. *Mol. Microbiol.* 51, 395–405. doi: 10.1046/j.1365-2958.2003.03855.x

Fu, G. K., and Markovitz, D. M. (1998). The human LON protease binds to mitochondrial promoters in a single-stranded, site-specific, strand-specific manner. *Biochemistry* 37, 1905–1909. doi: 10.1021/bi970928c

Goff, S. A., and Goldberg, A. L. (1987). An increased content of protease La, the lon gene product, increases protein degradation and blocks growth in *Escherichia coli*. *J. Biol. Chem.* 262, 4508–4515.

Goldberg, A. L., Strnad, N. P., and Swamy, K. S. (1980). Studies of the ATP dependence of protein degradation in cells and cell extracts. *Ciba Found. Symp.* 75, 227–251. doi: 10.1002/9780470720585.ch15

Gonzalez, D., Kozdon, J. B., McAdams, H. H., Shapiro, L., and Collier, J. (2014). The functions of DNA methylation by CcrM in *Caulobacter crescentus*: a global approach. *Nucleic Acids Res.* 42, 3720–3735. doi: 10.1093/nar/gkt1352

Gora, K. G., Cantin, A., Wohlever, M., Joshi, K. K., Perchuk, B. S., Chien, P., et al. (2013). Regulated proteolysis of a transcription factor complex is critical to cell cycle progression in *Caulobacter crescentus*. *Mol. Microbiol.* 87, 1277–1289. doi: 10.1111/mmi.12166

Gorbatyuk, B., and Marczynski, G. T. (2005). Regulated degradation of chromosome replication proteins DnaA and CtrA in *Caulobacter crescentus*. *Mol. Microbiol.* 55, 1233–1245. doi: 10.1111/j.1365-2958.2004.04459.x

Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell Dev. Biol.* 19, 565–587. doi: 10.1146/annurev.cellbio.19.110701.153228

Gray, M. J. (2019). Inorganic polyphosphate accumulation in *Escherichia coli* is regulated by DksA but not by (p) ppGpp. *J. Bacteriol.* 201:e0664-18.

Gross, M. H., and Konieczny, I. (2020). Polyphosphate induces the proteolysis of ADP-bound fraction of initiator to inhibit DNA replication initiation upon stress in *Escherichia coli*. *Nucleic Acids Res.* 48, 5457–5466. doi: 10.1093/nar/gkaa217

Hu, C., Su, C., Yun, Z., Wang, S., He, C., Gao, X., et al. (2018). Real-time identification of the singleness of a trapped bead in optical tweezers. *Appl. Opt.* 57, 1241–1246. doi: 10.1364/ao.57.001241

Ingmer, H., and Brøndsted, L. (2009). Proteases in bacterial pathogenesis. *Res. Microbiol.* 160, 704–710. doi: 10.1016/j.resmic.2009.08.017

Iniesta, A. A., McGrath, P. T., Reisenauer, A., McAdams, H. H., and Shapiro, L. (2006). A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10935–10940. doi: 10.1073/pnas.0604554103

Jacobs, C., Ausmees, N., Cordwell, S. J., Shapiro, L., and Laub, M. T. (2003). Functions of the CckA histidine kinase in *Caulobacter* cell cycle control. *Mol. Microbiol.* 47, 1279–1290. doi: 10.1046/j.1365-2958.2003.03379.x

Jenal, U. (2009). The role of proteolysis in the *Caulobacter crescentus* cell cycle and development. *Res. Microbiol.* 160, 687–695. doi: 10.1016/j.resmic.2009.09.006

Jenal, U., and Fuchs, T. (1998). An essential protease involved in bacterial cell-cycle control. *EMBO J.* 17, 5658–5669. doi: 10.1093/emboj/17.19.5658

Jenal, U., and Hengge-Aronis, R. (2003). Regulation by proteolysis in bacterial cells. *Curr. Opin. Microbiol.* 6, 163–172. doi: 10.1016/s1369-5274(03)00029-8

Jonas, K., Liu, J., Chien, P., and Laub, M. T. (2013). Proteotoxic stress induces a cell-cycle arrest by stimulating Lon to degrade the replication initiator DnaA. *Cell* 154, 623–636. doi: 10.1016/j.cell.2013.06.034

Kannan, G., Wilks, J. C., Fitzgerald, D. M., Jones, B. D., BonDurant, S. S., and Slonczewski, J. L. (2008). Rapid acid treatment of *Escherichia coli*: transcriptomic response and recovery. *BMC Microbiol.* 8:37. doi: 10.1186/1471-2180-8-37

Karlowicz, A., Wegrzyn, K., Gross, M., Kaczynska, D., Ropelewska, M., Siemiątkowska, M., et al. (2017). Defining the crucial domain and amino acid residues in bacterial Lon protease for DNA binding and processing of DNA-interacting substrates. *J. Biol. Chem.* 292, 7507–7518. doi: 10.1074/jbc.m116.766709

Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T., and Katayama, T. (2014). Timely binding of IHF and Fis to DARS2 regulates ATP-DnaA production and replication initiation. *Nucleic Acids Res.* 42, 13134–13149. doi: 10.1093/nar/gku1051

Kasho, K., and 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. doi: 10.1073/pnas.1212070110

Khattar, M. M. (1997). Overexpression of the hslVU operon suppresses SOS-mediated inhibition of cell division in *Escherichia coli*. *FEBS Lett.* 414, 402–404. doi: 10.1016/s0014-5793(97)01024-7

Kim, Y., and Wood, T. K. (2010). Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 391, 209–213. doi: 10.1016/j.bbrc.2009.11.033

Kornberg, A. (1995). Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *J. Bacteriol.* 177, 491–496. doi: 10.1128/jb.177.3.491-496.1995

Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The stringent response inhibits DNA replication initiation in *E. coli* by modulating supercoiling of oriC. *mBio* 10:e001330-19.

Kubik, S., Wegrzyn, K., Pierechod, M., and Konieczny, I. (2012). Opposing effects of DNA on proteolysis of a replication initiator. *Nucleic Acids Res.* 40, 1148–1159. doi: 10.1093/nar/gkr813

Kuroda, A. (2006). A polyphosphate-lon protease complex in the adaptation of *Escherichia coli* to amino acid starvation. *Biosci. Biotechnol. Biochem.* 70, 325–331. doi: 10.1271/bbb.70.325

Kuroda, A., Murphy, H., Cashel, M., and Kornberg, A. (1997). Guanosine tetra- and pentaphosphate promote accumulation of inorganic polyphosphate in *Escherichia coli*. *J. Biol. Chem.* 272, 21240–21243. doi: 10.1074/jbc.272.34.21240

Kuroda, A., Nomura, K., Ohtomo, R., Kato, J., Ikeda, T., Takiguchi, N., et al. (2001). Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in *E. coli*. *Science* 293, 705–708. doi: 10.1126/science.1061315

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*. *EMBO J.* 18, 6642–6652. doi: 10.1093/emboj/18.23.6642

Langklotz, S., and Narberhaus, F. (2011). The *Escherichia coli* replication inhibitor CspD is subject to growth-regulated degradation by the Lon protease. *Mol. Microbiol.* 80, 1313–1325. doi: 10.1111/j.1365-2958.2011.07646.x

Laub, M. T., Chen, S. L., Shapiro, L., and McAdams, H. H. (2002). Genes directly controlled by CtrA, a master regulator of the *Caulobacter* cell cycle. *Proc. Natl. Acad. Sci. U.S.A.* 99, 4632–4637. doi: 10.1073/pnas.062065699

Lee, A. L., Chen, Y. D., Chang, Y. Y., Lin, Y. C., Chang, C. F., Huang, S. J., et al. (2014). Structural basis for DNA-mediated allosteric regulation facilitated by the AAA+ module of Lon protease. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70, 218–230. doi: 10.1107/s139900471302631x

Lee, A. Y. L., Hsu, C. H., and Wu, S. H. (2004). Functional domains of *Brevibacillus thermoruber* Lon protease for oligomerization and DNA binding role of N-terminal and sensor and substrate discrimination domains. *J. Biol. Chem.* 279, 34903–34912. doi: 10.1074/jbc.m403562200

Lee, S. H., and Walker, J. R. (1987). *Escherichia coli* DnaX product, the tau subunit of DNA polymerase III, is a multifunctional protein with single-stranded DNA-dependent ATPase activity. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2713–2717. doi: 10.1073/pnas.84.9.2713

Leslie, D. J., Heinen, C., Schramm, F. D., Thüring, M., Aakre, C. D., Murray, S. M., et al. (2015). Nutritional control of DNA replication initiation through the proteolysis and regulated translation of DnaA. *PLoS Genet.* 11:e10053432. doi: 10.1371/journal.pgen.1005342

Levchenko, I., Luo, L., and Baker, T. A. (1995). Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes Dev.* 9, 2399–2408. doi: 10.1101/gad.9.19.2399

Lin, Y. C., Lee, H. C., Wang, I., Hsu, C. H., Liao, J. H., Lee, A. Y. L., et al. (2009). DNA-binding specificity of the Lon protease α-domain from *Brevibacillus thermoruber* WR-249. *Biochem. Biophys. Res. Commun.* 388, 62–66. doi: 10.1016/j.bbrc.2009.07.118

Liu, J., Francis, L. I., Jonas, K., Laub, M. T., and Chien, P. (2016). ClpAP is an auxiliary protease for DnaA degradation in *Caulobacter crescentus*. *Mol. Microbiol.* 102, 1075–1085. doi: 10.1111/mmi.13537

Liu, T., Lu, B., Lee, I., Ondrovièová, G., Kutejová, E., and Suzuki, C. K. (2004). DNA and RNA binding by the mitochondrial lon protease is regulated by nucleotide and protein substrate. *J. Biol. Chem.* 279, 13902–13910. doi: 10.1074/jbc.m309642200

Lu, B., Liu, T., Crosby, J. A., Thomas-Wohlever, J., Lee, I., and Suzuki, C. K. (2003). The ATP-dependent Lon protease of Mus musculus is a DNA-binding protein that is functionally conserved between yeast and mammals. *Gene* 306, 45–55. doi: 10.1016/s0378-1119(03)00403-7

Magnusson, L. U., Farewell, A., and Nystrom, T. (2005). ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol.* 13, 236–242. doi: 10.1016/j.tim.2005.03.008

Martin, A., Baker, T. A., and Sauer, R. T. (2008). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol. Cell.* 29, 441–450. doi: 10.1016/j.molcel.2008.02.002

Melekhov, V. V., Shvyreva, U. S., Timchenko, A. A., Tutukina, M. N., Preobrazhenskaya, E. V., Burkova, D. V., et al. (2015). Modes of *Escherichia coli* Dps interaction with DNA as revealed by atomic force microscopy. *PLoS One* 10:e0126504. doi: 10.1371/journal.pbio.000126504

Minami, N., Yasuda, T., Ishii, Y., Fujimori, K., and Amano, F. (2011). Regulatory role of cardiolipin in the activity of an ATP-dependent protease, Lon, from *Escherichia coli*. *J. Biochem.* 149, 519–527. doi: 10.1093/jb/mvr036

Nicoloff, H., Perreten, V., and Levy, S. B. (2007). Increased genome instability in *Escherichia coli* lon mutants: relation to emergence of multiple-antibiotic-resistant (Mar) mutants caused by insertion sequence elements and large tandem genomic amplifications. *Antimicrob. Agents Chemother.* 51, 1293–1303. doi: 10.1128/aac.01128-06

Ninnis, R. L., Spall, S. K., Talbo, G. H., Truscott, K. N., and Dougan, D. A. (2009). Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli*. *EMBO J.* 28, 1732–1744. doi: 10.1038/emboj.2009.134

Nomura, K., Kato, J., Takiguchi, N., Ohtake, H., and Kuroda, A. (2004). Effects of inorganic polyphosphate on the proteolytic and DNA-binding activities of Lon in *Escherichia coli*. *J. Biol. Chem.* 279, 34406–34410. doi: 10.1074/jbc.m404725200

Osbourne, D. O., Soo, V. W., Konieczny, I., and Wood, T. K. (2014). Polyphosphate, cyclic AMP, guanosine tetraphosphate, and c-di-GMP reduce in vitro Lon activity. *Bioengineered* 5, 264–268. doi: 10.4161/bioe.29261

Ozaki, S., and Katayama, T. (2012). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.* 40, 1648–1665. doi: 10.1093/nar/gkr832

Pettijohn, D. E. (1988). Histone-like proteins and bacterial chromosome structure. *J. Biol. Chem.* 263, 12793–12796.

Pierechod, M., Nowak, A., Saari, A., Purta, E., Bujnicki, J. M., and Konieczny, I. (2009). Conformation of a plasmid replication initiator protein affects its proteolysis by ClpXP system. *Protein Sci.* 18, 637–649.

Puri, N. (2016). *HspQ Functions as a Novel Specificity-Enhancing Factor for the AAA+ Lon Protease*. Doctoral thesis, The Graduate School, Stony Brook University, Stony Brook, NY.

Quon, K. C., Marczynski, G. T., and Shapiro, L. (1996). Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* 84, 83–93. doi: 10.1016/s0092-8674(00)80995-2

Quon, K. C., Yang, B., Domian, I. J., Shapiro, L., and Marczynski, G. T. (1998). Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. *Proc. Natl. Acad. Sci. U.S.A.* 95, 120–125. doi: 10.1073/pnas.95.1.120

Racki, L. R., Tocheva, E. I., Dieterle, M. G., Sullivan, M. C., Jensen, G. J., and Newman, D. K. (2017). Polyphosphate granule biogenesis is temporally and functionally tied to cell cycle exit during starvation in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* 114, E2440–E2449.

Rao, N. N., Gomez-Garcia, M. R., and Kornberg, A. (2009). Inorganic polyphosphate: essential for growth and survival. *Annu. Rev. Biochem.* 78, 605–647. doi: 10.1146/annurev.biochem.77.083007.093039

Rao, N. N., and Kornberg, A. (1996). Inorganic polyphosphate supports resistance and survival of stationary-phase *Escherichia coli*. *J. Bacteriol.* 178, 1394–1400. doi: 10.1128/jb.178.5.1394-1400.1996

Schweder, T., Lee, K. H., Lomovskaya, O., and Matin, A. (1996). Regulation of *Escherichia coli* starvation sigma factor (sigma s) by ClpXP protease. *J. Bacteriol.* 178, 470–476. doi: 10.1128/jb.178.2.470-476.1996

Simmons, L. A., Grossman, A. D., and Walker, G. C. (2008). Clp and Lon proteases occupy distinct subcellular positions in *Bacillus subtilis*. *J. Bacteriol.* 190, 6758–6768. doi: 10.1128/jb.00590-08

Smith, S. C., Joshi, K. K., Zik, J. J., Trinh, K., Kamajaya, A., Chien, P., et al. (2014). Cell cycle-dependent adaptor complex for ClpXP-mediated proteolysis directly integrates phosphorylation and second messenger signals. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14229–14234. doi: 10.1073/pnas.1407862111

Sonezaki, S., Okita, K., Oba, T., Ishii, Y., Kondo, A., and Kato, Y. (1995). Protein substrates and heat shock reduce the DNA-binding ability of *Escherichia coli* Lon protease. *Appl. Microbiol. Biotechnol.* 44, 484–488. doi: 10.1007/s002530050586

Stephani, K., Weichart, D., and Hengge, R. (2003). Dynamic control of Dps protein levels by ClpXP and ClpAP proteases in *Escherichia coli*. *Mol. Microbiol.* 49, 1605–1614. doi: 10.1046/j.1365-2958.2003.03644.x

Sugiyama, N., Minami, N., Ishii, Y., and Amano, F. (2013). Inhibition of Lon protease by bacterial lipopolysaccharide (LPS) though inhibition of ATPase. *Adv. Biosci. Biotechnol.* 4, 590–598. doi: 10.4236/abb.2013.44077

Traxler, M. F., Summers, S. M., Nguyen, H. T., Zacharia, V. M., Hightower, G. A., Smith, J. T., et al. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol. Microbiol.* 68, 1128–1148. doi: 10.1111/j.1365-2958.2008.06229.x

Vass, R. H., and Chien, P. (2013). Critical clamp loader processing by an essential AAA+ protease in *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18138–18143. doi: 10.1073/pnas.1311302110

Wah, D. A., Levchenko, I., Baker, T. A., and Sauer, R. T. (2002). Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer. *Chem. Biol.* 9, 1237–1245. doi: 10.1016/s1074-5521(02)00268-5

Waxman, L., and Goldberg, A. L. (1985). Protease La, the lon gene product, cleaves specific fluorogenic peptides in an ATP-dependent reaction. *J. Biol. Chem.* 260, 12022–12028.

Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K., and Maurizi, M. R. (1994). A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc. Natl. Acad. Sci. U.S.A.* 91, 12218–12222. doi: 10.1073/pnas.91.25.12218

Willett, J. W., Herrou, J., Briegel, A., Rotskoff, G., and Crosson, S. (2015). Structural asymmetry in a conserved signaling system that regulates division, replication, and virulence of an intracellular pathogen. *Proc. Natl. Acad. Sci. U.S.A.* 112, E3709–E3718.

Wojtkowiak, D., Georgopoulos, C., and Zylicz, M. (1993). Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. *J. Biol. Chem.* 268, 22609–22617.

Wright, R., Stephens, C., Zweiger, G., Shapiro, L., and Alley, M. R. (1996). Caulobacter Lon protease has a critical role in cell-cycle control of DNA methylation. *Genes Dev.* 10, 1532–1542. doi: 10.1101/gad.10.12.1532

Xie, F., Li, G., Zhang, Y., Zhou, L., Liu, S., Liu, S., et al. (2016). The Lon protease homologue LonA, not LonC, contributes to the stress tolerance and biofilm formation of *Actinobacillus pleuropneumoniae*. *Microb. Pathog.* 93, 38–43. doi: 10.1016/j.micpath.2016.01.009

Yamanaka, K., and Inouye, M. (1997). Growth-phase-dependent expression of cspD, encoding a member of the CspA family in *Escherichia coli*. *J. Bacteriol.* 179, 5126–5130. doi: 10.1128/jb.179.16.5126-5130.1997

Yamanaka, K., Zheng, W., Crooke, E., Wang, Y. H., and Inouye, M. (2001). CspD, a novel DNA replication inhibitor induced during the stationary phase in *Escherichia coli*. *Mol. Microbiol.* 39, 1572–1584. doi: 10.1046/j.1365-2958.2001.02345.x

Zakrzewska-Czerwińska, J., Jakimowicz, D., Zawilak-Pawlik, A., and Messer, W. (2007). Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiol. Rev.* 31, 378–387. doi: 10.1111/j.1574-6976.2007.00070.x

Zehnbauer, B. A., Foley, E. C., Henderson, G. W., and Markovitz, A. (1981). Identification and purification of the Lon+ (capR+) gene product, a DNA-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 78, 2043–2047. doi: 10.1073/pnas.78.4.2043

Zeinert, R. D., Baniasadi, H., Tu, B. P., and Chien, P. (2020). The Lon protease links nucleotide metabolism with proteotoxic stress. *Mol. Cell* 79, 758–767.e6.

Zeinert, R. D., Liu, J., Yang, Q., Du, Y., Haynes, C., and Chien, P. (2018). A legacy role for DNA binding of Lon protects against genotoxic stress. *bioRxiv* [Preprint], doi: 10.1101/317677

Zhao, J., Niu, W., Yao, J., Mohr, S., Marcotte, E. M., and Lambowitz, A. M. (2008). Group II intron protein localization and insertion sites are affected by polyphosphate. *PLoS Biol.* 6:e150. doi: 10.1371/journal.pbio.0060150

Zhou, X., Wang, J., Herrmann, J., Moerner, W. E., and Shapiro, L. (2019). Asymmetric division yields progeny cells with distinct modes of regulating cell cycle-dependent chromosome methylation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 15661–15670. doi: 10.1073/pnas.1906119116

Zylicz, M., Liberek, K., Wawrzynow, A., and Georgopoulos, C. (1998). Formation of the preprimosome protects λ O from RNA transcription-dependent proteolysis by ClpP/ClpX. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15259–15263. doi: 10.1073/pnas.95.26.15259

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 02 October 2020
doi: 10.3389/fmicb.2020.581401

Identification of the Unwinding Region in the *Clostridioides difficile* Chromosomal Origin of Replication

Ana M. Oliveira Paiva[1,2†], Erika van Eijk[1†], Annemieke H. Friggen[1†], Christoph Weigel[3] and Wiep Klaas Smits[1,2†]*

[1] Department of Medical Microbiology, Section Experimental Bacteriology, Leiden University Medical Center, Leiden, Netherlands, [2] Centre for Microbial Cell Biology, Leiden, Netherlands, [3] Institute of Biotechnology, Technische Universität Berlin, Berlin, Germany

OPEN ACCESS

Edited by:
Monika Glinkowska,
University of Gdañsk, Poland

Reviewed by:
Julia Grimwade,
Florida Institute of Technology, United States
Shogo Ozaki,
Kyushu University, Japan

***Correspondence:**
Wiep Klaas Smits
w.k.smits@lumc.nl

†ORCID:
Ana M. Oliveira Paiva
orcid.org/0000-0002-6122-832X
Erika van Eijk
orcid.org/0000-0003-3003-9822
Annemieke H. Friggen
orcid.org/0000-0001-5780-2053
Wiep Klaas Smits
orcid.org/0000-0002-7409-2847

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: 08 July 2020
Accepted: 31 August 2020
Published: 02 October 2020

Citation:
Oliveira Paiva AM, van Eijk E, Friggen AH, Weigel C and Smits WK (2020) Identification of the Unwinding Region in the Clostridioides difficile Chromosomal Origin of Replication. Front. Microbiol. 11:581401. doi: 10.3389/fmicb.2020.581401

Faithful DNA replication is crucial for viability of cells across all kingdoms. Targeting DNA replication is a viable strategy for inhibition of bacterial pathogens. *Clostridioides difficile* is an important enteropathogen that causes potentially fatal intestinal inflammation. Knowledge about DNA replication in this organism is limited and no data is available on the very first steps of DNA replication. Here, we use a combination of *in silico* predictions and *in vitro* experiments to demonstrate that *C. difficile* employs a bipartite origin of replication that shows DnaA-dependent melting at *oriC2*, located in the *dnaA-dnaN* intergenic region. Analysis of putative origins of replication in different clostridia suggests that the main features of the origin architecture are conserved. This study is the first to characterize aspects of the origin region of *C. difficile* and contributes to our understanding of the initiation of DNA replication in clostridia.

Keywords: *oriC*, *Clostridioides difficile*, P1 nuclease, unwinding, DnaA

INTRODUCTION

Clostridioides difficile (formerly *Clostridium difficile*) (Lawson et al., 2016) is a Gram-positive anaerobic bacterium. *C. difficile* infections (CDI) can occur in individuals with a disturbed microbiota and is one of the main causes of hospital associated diarrhea, but can also be found in the environment (Smits et al., 2016). The incidence of CDI has increased worldwide since the beginning of the century (Smits et al., 2016; Warriner et al., 2017). Consequently, the interest in the physiology of the bacterium has increased as a way to understand its interaction with the host and the environment and to explore news pathways for intervention (van Eijk et al., 2017; Crobach et al., 2018).

One such pathway is the replication of the chromosome. Overall, DNA replication is a highly conserved process across different kingdoms (O'Donnell et al., 2013; Bleichert et al., 2017). In all bacteria, DNA replication is a tightly regulated process that occurs with high fidelity and efficiency, and is essential for cell survival. The process involves many different proteins that are required for the replication process itself, or to regulate and aid replisome assembly and activity (Katayama et al., 2010; Murray and Koh, 2014; Chodavarapu and Kaguni, 2016; Jameson and Wilkinson, 2017; Schenk et al., 2017). Replication initiation and its regulation arguably are candidates for the search for novel therapeutic targets (Fossum et al., 2008; Grimwade and Leonard, 2017; van Eijk et al., 2017).

In most bacteria, replication of the chromosome starts with the assembly of the replisome at the origin of replication (*oriC*) and proceeds bidirectionally (Chodavarapu and Kaguni, 2016).

In the majority of bacteria replication is initiated by the DnaA protein, an ATPase Associated with diverse cellular Activities (AAA+ protein) that binds specific sequences in the *oriC* region. The binding of DnaA induces DNA duplex unwinding, which subsequently drives the recruitment of other proteins, such as the replicative helicase, primase and DNA polymerase III proteins (Chodavarapu and Kaguni, 2016). Termination of replication eventually leads to disassembly of the replication complexes (Chodavarapu and Kaguni, 2016).

In *C. difficile*, knowledge on DNA replication is limited. Though many proteins appear to be conserved between well-characterized species and *C. difficile*, only certain replication proteins have been experimentally characterized for *C. difficile* (Torti et al., 2011; Briggs et al., 2012; van Eijk et al., 2016). DNA polymerase C (PolC, CD1305) of *C. difficile* has been studied in the context of drug-discovery and appears to have a conserved primary structure similar to other low-[G + C] gram-positive organisms (Torti et al., 2011). It is inhibited *in vitro* and *in vivo* by compounds that compete for binding with dGTP (van Eijk et al., 2019; Xu et al., 2019). Helicase (CD3657), essential for DNA duplex unwinding, was found to interact in an ATP-dependent manner with a helicase loader (CD3654) and loading was proposed to occur through a ring-maker mechanism (Davey and O'Donnell, 2003; van Eijk et al., 2016). However, in contrast to helicase of the Firmicute *Bacillus subtilis*, *C. difficile* helicase activity is dependent on activation by the primase protein (CD1454), as has also been described for *Helicobacter pylori* (Bazin et al., 2015; van Eijk et al., 2016). *C. difficile* helicase stimulates primase activity at the trinucleotide 5′d(CTA), but not at the preferred trinucleotide 5′-d(CCC) (van Eijk et al., 2016).

DnaA of *C. difficile* has not been studied to date. Although no full-length structure has been determined for DnaA, individual domains of the DnaA protein from different organisms have been characterized (Majka et al., 1997; Zawilak et al., 2003; Erzberger et al., 2006; Zawilak-Pawlik et al., 2017). DnaA proteins generally comprise four domains (Zawilak-Pawlik et al., 2017). Domain I is involved in protein-protein interactions and is responsible for DnaA oligomerization (Weigel et al., 1999; Abe et al., 2007; Natrajan et al., 2009; Jameson et al., 2014; Kim et al., 2017; Zawilak-Pawlik et al., 2017; Martin et al., 2018; Matthews and Simmons, 2019; Nowaczyk-Cieszewska et al., 2020). Little is known about a specific function of Domain II and this domain may even be absent (Erzberger et al., 2002). It is thought to be a flexible linker that promotes the proper conformation of the other DnaA domains (Abe et al., 2007; Nozaki and Ogawa, 2008). Domain III and Domain IV are responsible for the DNA binding. Domain III contains the AAA+ motif and is responsible for binding ATP, ADP and single-stranded DNA, as well as certain regulatory proteins (Kawakami et al., 2005; Cho et al., 2008; Ozaki et al., 2008; Ozaki and Katayama, 2012). Recent studies have also revealed the importance of this domain for binding phospholipids present in the bacterial membrane (Saxena et al., 2013). The C-terminal Domain IV contains a helix-turn-helix motif (HTH) and is responsible for the specific binding of DnaA to so called DnaA boxes (Blaesing et al., 2000; Erzberger et al., 2002; Fujikawa et al., 2003).

DnaA boxes are typically 9-mer non-palindromic DNA sequences, and the *Escherichia coli* DnaA box consensus sequence is TTWTNCACA (Schaper and Messer, 1995; Wolanski et al., 2014). The boxes can differ in their affinity for DnaA, and even demonstrate different dependencies on the ATP co-factor (Speck et al., 1999; Patel et al., 2017). Binding of Domain IV to the DnaA boxes promotes higher-order oligomerization of DnaA, forming a filament that wraps around DNA (Erzberger et al., 2006; Ozaki et al., 2012; Scholefield and Murray, 2013). It is thought that the interaction of the DnaA filament with the DNA helix introduces a bend in the DNA (Erzberger et al., 2006; Patel et al., 2017). The resulting superhelical torsion facilitates the melting of the adjacent [A + T]-rich DNA Unwinding Element (DUE) (Kowalski and Eddy, 1989; Erzberger et al., 2006; Zorman et al., 2012). Upon melting, the DUE provides the entry site for the replisomal proteins. Another conserved structural motif, a triplet repeat called DnaA-trio, is involved in the stabilization of the unwound region (Richardson et al., 2016, 2019).

The *oriC* region has been characterized for several bacterial species. These analyses show that *oriC* regions are quite diverse in sequence, length and even chromosomal location, all of which contribute to species-specific replication initiation requirements (Zawilak-Pawlik et al., 2005; Ekundayo and Bleichert, 2019). In Firmicutes, including *C. difficile*, the genomic context of the origin regions appears to be conserved and encompasses the *rnpA-rpmH-dnaA-dnaN* genes (Ogasawara and Yoshikawa, 1992; Briggs et al., 2012).

The *oriC* region can be continuous (i.e., located at a single chromosomal locus) or bipartite (Wolanski et al., 2014). Bipartite origins were initially identified in *B. subtilis* (Moriya et al., 1988) but more recently also in *H. pylori* (Donczew et al., 2012). The separated subregions of the bipartite origin, *oriC1* and *oriC2*, are usually separated by the *dnaA* gene. Both *oriC1* and *oriC2* contain clusters of DnaA boxes, and one of the regions contains the DUE region. The DnaA protein binds to both subregions and places them in close proximity to each other, consequently looping out the *dnaA* gene (Krause et al., 1997; Donczew et al., 2012). In *H. pylori*, DnaA Domain I and II are important for maintaining the interactions between both *oriC* regions (Nowaczyk-Cieszewska et al., 2020).

In this study, we identified the putative *oriC* of *C. difficile* through *in silico* analysis and demonstrate DnaA-dependent unwinding of the *oriC2* region *in vitro*. A clear conservation of the origin of replication organization is observed throughout the clostridia. The present study contributes to our understanding of clostridial DNA replication initiation in general, and replication initiation of *C. difficile* specifically.

MATERIALS AND METHODS

Sequence Alignments and Structure Modeling

Multiple sequence alignment of amino acid sequences was performed with Protein BLAST (blastP suite)[1] for individual

[1]https://blast.ncbi.nlm.nih.gov/Blast.cgi

alignment scores and the PRALINE program[2] (Bawono and Heringa, 2014) for multiple sequence alignment. Sequences were retrieved from the NCBI Reference Sequences. DnaA protein sequences from *C. difficile* 630Δ*erm* (CEJ96502.1), *C. acetobutylicum* DSM 1731 (AEI33799.1), *Bacillus subtilis* 168 (NP_387882.1), *Escherichia coli* K-12 (AMH32311.1), *Streptomyces coelicolor* A3(2) (TYP16779.1), *Mycobacterium tuberculosis* RGTB327 (AFE14996.1), *Helicobacter pylori J99* (Q9ZJ96.1) and *Aquifex aeolicus* (WP_010880157.1) were selected for alignment. Alignment was visualized in Jalview version 2.11, with coloring by percentage identity.

Secondary structure prediction and homology modeling were performed using Phyre2[3] (Kelley et al., 2015) using the intensive default settings. Phyre2 modeling of *C. difficile* 630Δ*erm* DnaA (CEJ96502.1) was performed with 3 templates from *A. aeolicus* (PDB 2HCB, chain C), *B. subtilis* (PDB 4TPS, chain D) and *E. coli* (PDB 2E0G, chain A) and 21 residues were modeled *ab initio*. 95% of the residues were modeled with >90% confidence. Graphical representation was performed with the PyMOL Molecular Graphics System, Version 1.76.6. (Schrödinger, LLC).

Prediction of the *C. difficile oriC*

To identify the *oriC* region of *C. difficile* the genome sequence of *C. difficile* 630Δ*erm* (GenBank accession no. LN614756.1) was analyzed through different software in a stepwise procedure (Mackiewicz et al., 2004).

The GenSkew Java Application[4] was used with default settings for the analysis of the normal and the cumulative skew of two selectable nucleotides of the genomic nucleotide sequence [(G−C)/(G + C)]. Calculations where performed with a window size of 4293 bp and a step size of 4293 bp. The inflection values of the cumulative GC skew plot are indicative of the chromosomal origin (*oriC*) and terminus of replication (*ter*).

Prediction of superhelicity-dependent helically unstable DNA stretches (SIDDs) was performed in the vicinity of the inflection point of the GC-skew plot, in 2.0 kb fragments comprising intergenic regions from nucleotide position 4291795 to 745 (*oriC1*) and 466 to 2465 (*oriC2*) of the *C. difficile* 630Δ*erm* chromosome. Prediction of the SIDDs in the different clostridia (**Table 1**) was performed in the vicinity of the inflection points of the GC-plot retrieved from DoriC 10.0 database[5] (Luo and Gao, 2019), in 2.0 kb fragments comprising intergenic regions summarized in **Table 1**. The SIST program[6] (Zhabinskaya et al., 2015) was used to predicted free energies $G_{(x)}$ by running the melting transition algorithm only (SIDD) with default values (copolymeric energetics; default: $\sigma = -0.06$; $T = 37°C$; $x = 0.01$ M) and with superhelical density $\sigma = -0.04$.

We performed the identification of the DnaA box clusters by search of the motif TTWTNCACA with one mismatch (**Supplementary Material**) in the leading strand on a 4432 bp sequence between the nucleotide position 4291488 to 2870 of the *C. difficile* 630Δ*erm* chromosome, using Pattern Locator[7] (Mrazek and Xie, 2006). Identification of the DnaA boxes in the different clostridia was performed with the same pattern motif in the leading strand of the intergenic regions summarized in **Table 1**.

TABLE 1 | Clostridia intergenic regions used for SIDD analysis.

Clostridia (GenBank accession no.)	*oriC1**1 *DoriC ID**2	*oriC2 DoriC ID**
C. difficile R20291 (NC_013316.1)	4189900 to 561 ORI93010593	780 to 2780 ORI93010592
C. botulinum A Hall (NC_009698.1)	3759361 to 800 ORI92010336	510 to 2510 ORI92010335
C. sordelli AM370 (NZ_CP014150	3549121 to 662 ORI97012279	561 to 2561 ORI97012278
C. acetobutylicum DSM 1731 (NC_015687.1)	3941422 to 961 ORI94010884	1040 to 3040 ORI94010883
C. perfringens str.13 (NC_003366.1)	3030241 to 810 ORI10010054	881 to 2881 ORI10010053
C. tetani E88 (NC_004557.1)	52001 to 54000 ORI10010089	50081 to 52081 ORI10010088

**1 2.0 kb fragments selected for SIDD analysis comprising the intergenic regions. *2 DoriC 10.0 intergenic regions from http://tubic.tju.edu.cn/doric/public/index.php.*

DnaA-trio sequences and ribosomal binding sites where manually predicted based on Richardson et al. (2016) and Vellanoweth and Rabinowitz (1992), respectively.

All output data was obtained as raw text files and further processed with Prism 8.3.1 (GraphPad, Inc., La Jolla, CA, United States) and CorelDRAW X7 (Corel).

Strains and Growth Conditions

Escherichia coli strains were grown aerobically at 37°C in lysogeny broth (LB, Affymetrix) supplemented with 15 μg/mL chloramphenicol or 50 μg/mL kanamycin when required. *E. coli* strains DH5α and MC1061 (**Table 2**) were used to maintain *dnaA*- and *oriC*-containing plasmids, respectively. *E. coli* strain MS3898, kindly provided by Alan Grossman (MIT, Cambridge, United States) (**Table 2**) was used for recombinant DnaA expression. *E. coli* transformation was performed using standard procedures (Sambrook et al., 1989). The growth was followed by monitoring the optical density at 600 nm (OD_{600}).

Construction of the Plasmids

For overexpression of DnaA, the *dnaA* nucleotide sequence (CEJ96502.1) from *C. difficile* 630Δ*erm* (GenBank accession no. LN614756.1) was amplified by PCR from *C. difficile* 630Δ*erm* genomic DNA using primers oEVE-7 and oEVE-21 (**Table 3**). The PCR product was subsequently digested with *Nco*I and *Bgl*II. The vector pAV13 (Smits et al., 2011; **Table 4**), containing *B. subtilis dnaA* cloned in pQE60 (Qiagen) was kindly provided by Alan Grossman (MIT, Cambridge, MA, United States) and was digested with the same enzymes and ligated to the digested fragment to yield vector pEVE40 (**Table 4**).

[2] http://www.ibi.vu.nl/programs/pralinewww/

[3] http://www.sbg.bio.ic.ac.uk/phyre2

[4] http://genskew.csb.univie.ac.at/

[5] http://tubic.tju.edu.cn/doric/public/index.php

[6] https://bitbucket.org/benhamlab/sist_codes/src/master/

[7] https://www.cmbl.uga.edu//downloads/programs/Pattern_Locator/patloc.c

To construct a plasmid carrying the complete predicted *oriC*, the predicted *oriC* region (nucleotide 4292150 to 1593 from *C. difficile* 630 GenBank accession no. LN614756.1) was amplified by PCR from *C. difficile* 630Δ*erm* genomic DNA using primers oAP40 and oAP41 (**Table 3**). The PCR product was subsequently digested with *Eco*RI and *Pst*I and ligated into pori1ori2 (**Table 4**), kindly provided by Anna Zawilak-Pawlik (Hirszfeld Institute of Immunology and Experimental Therapy, PAS, Wrocław, Poland), that was digested with the same enzymes, to yield vector pAP205 (**Table 4**).

For the cloning of the predicted *oriC1* region (nucleotide 4292150 to 24 of *C. difficile* 630Δ*erm* genomic DNA) the primer set oAP30/oAP31 (**Table 3**) was used. The amplified fragment was digested with *Eco*RI and *Pst*I and inserted onto pori1ori2 digested with same enzymes, yielding vector pAP83 (**Table 4**). For the cloning of the predicted *oriC2* region (nucleotide 1291 to the 1593 of *C. difficile* 630Δ*erm* genomic DNA) the primer set oAP32/oAP33 (**Table 3**) was used. The amplified fragment was digested with *Eco*RI and *Pst*I and inserted onto pori1ori2 (**Table 4**) digested with same enzymes, yielding vector pAP76 (**Table 4**).

All DNA sequences introduced into the cloning vectors were verified by Sanger sequencing. For *oriC* containing vectors primers oAP56 and oAP57 (**Table 3**) were used for sequencing.

TABLE 2 | *E. coli* strains used in this study.

Name	Relevant Genotype/Phenotype	Origin
DH5α	F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK–mK +), λ–	Laboratory collection
MC1061	str. K-12 F– λ– Δ(ara-leu)7697 [araD139]B/r Δ(codB-lacI)3 galK16 galE15 e14– mcrA0 relA1 rpsL150(StrR) spoT1 mcrB1 hsdR2(r–m +)	Laboratory Collection
CYB1002	ΔdnaA zia:pKN500(miniR1) asnB32 relA1 spoT1 thi-1 ilv192 mad1 recA1 λimm434 F-pBB42 (lacI; TetR)	Grossman lab

TABLE 3 | Oligonucleotides used in this study.

Name	Sequence (5′ > 3′)*
oEVE-7	CAGTCCATGGATATAGTTTCTTTATGGGACAAAACC
oEVE-21	CGGCAGATCTTCCCTTCAAATCTGATATAATTTTGTCTATTTTAG
oAP30	AATTGAATTCTTTGTCCCATAAAGAAACTATATCC
oAP31	TGGGCTGCAGTTCAACCCTTTAGTCCTATTAAAGTCC
oAP32	AATTGAATTCTTTGCTAGGATTTTTTGATTAC
oAP33	TGGGCTGCAGTTGACAAAATTATATCAGATTTG
oAP40	TGGGCTGCAGTTGCTAGGATTTTTTGATTAC
oAP41	AATTGAATTCTTTCAACCCTTTAGTCCTATTAAAGTCC
oAP56	CAGCGAGTCAGTGAGCGAGGAAG
oAP57	GATTGATTTAATTCTCATGTTTGAC

**Restriction enzyme cleavage sites used underlined.*

TABLE 4 | Plasmids used in this study.

Name	Relevant features*	Source/Reference
pAV13	$lacI^q$, P_{T5} expression vector; *km*	(Smits et al., 2011)
pEVE40	P_{T5} - DnaA-6xHis; *km*	This study
pori1ori2	*H. pylori oriC1oriC2; amp*	(Donczew et al., 2012)
pAP76	*C. difficile oriC2; amp*	This study
pAP83	*C. difficile oriC1*; amp	This study
pAP205	*C. difficile oriC1oriC2; amp*	This study

** amp, ampicillin resistance cassette; km, kanamycin resistance cassette.*

Overproduction and Purification of DnaA-6xHis

Overexpression of DnaA-6xHis was carried out in *E. coli* strain CYB1002 (**Table 2**), harboring the expression plasmid pEVE40 (**Table 4**). Cells were grown in 800 mL LB and induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at an OD_{600} of 0.6 for 3 h. The cells were collected by centrifugation at 4°C and stored at −80°C. Cells were resuspended in Binding buffer (1X Phosphate buffer pH7.4, 10 mM Imidazole, 10% glycerol) lysed by French Press and collected in phenylmethylsulfonyl fluoride (PMSF) at 0.1 mM (end concentration). Separation of the soluble fraction was performed by centrifugation at 13000 × *g* at 4°C for 20 min. Purification of the protein from the soluble fraction was done in Binding buffer on a 1 mL HisTrap Column (GE Healthcare) according to manufacturer's instructions. Elution was performed with Binding buffer in stepwise increasing concentrations of imidazole (20, 60, 100, 300, and 500 mM). DnaA-6xHis was mainly eluted at a concentration of imidazole equal to or greater than 300 mM.

Fractions containing the DnaA-6xHis protein were pooled together and applied to Amicon Ultra Centrifugal Filters with 30 kDa cutoff (Millipore). Buffer was exchanged to Buffer A (25 mM HEPES-KOH pH 7.5, 100 mM K-glutamate, 5 mM Mg-acetate, 10% glycerol). The concentrated DnaA-6xHis protein was subjected to size exclusion chromatography on an Äkta pure instrument (GE Healthcare). 200 μL of concentrated DnaA-6xHis was applied to a Superdex 200 Increase 10/30 column (GE Healthcare) in buffer A at a flow rate of 0.5 ml min^{-1}. UV detection was done at 280 nm. The column was calibrated with a mixture of proteins of known molecular weights (Mw): thyroglobulin (669 kDa), Apoferritin (443 kDa), β-amylase (200 kDa), Albumin (66 kDa), and Carbonic anhydrase (29 kDa). Eluted fractions containing DnaA-6xHis of the expected molecular weight (51 kDa) were quantified and visualized by Coomassie. Pure fractions were aliquoted and stored at −80°C for further experiments.

Immunoblotting and Detection

For immunoblotting, proteins were separated on a 12% SDS-PAGE gel and transferred onto nitrocellulose membranes (Amersham), according to the manufacturer's instructions. The membranes were probed in PBST (PBS pH 7.4, 0.05% (v/v) Tween-20) with a mouse anti-his antibody (1:3000, Invitrogen) and a secondary goat anti-mouse-HRP antibody (1:3000, DAKO) was used. The membranes were visualized using the chemiluminescence detection kit Clarity ECL Western

Blotting Substrates (Bio-Rad) in an Alliance Q9 Advanced machine (Uvitec).

P1 Nuclease Assay

For the P1 nuclease assay, 100 ng pAP205 plasmid was incubated with increasing concentrations of DnaA-6xHis (0.14, 0.54, 1, and 6.3 μM), when required, in P1 buffer (25 mM Hepes-KOH (pH 7.6), 12% (v/v) glycerol, 1 mM $CaCl_2$, 0.2 mM EDTA, 5 mM ATP, 0.1 mg/ml BSA), at 30°C for 12 min. 0.75 unit of P1 nuclease (Sigma), resuspended in 0.01 M sodium acetate (pH 7.6) was added to the reaction and incubated at 30°C for 5 min. 220 μl of buffer PB (Qiagen) was added and the fragments purified with the minElute PCR Purification Kit (Qiagen), according to manufacturer's instructions. Digestion with *Bgl*II, *Not*I or *Sca*I (NEB) of the purified fragments was performed according to manufacturer's instructions for 1 h at 37°C. Digested samples were resolved on 1% agarose gels in 0.5xTAE (40 mM Tris, 20 mM CH-COOH, 1 mM EDTA PH 8.0) and stained with 0.01 mg/mL ethidium bromide solution afterward. Visualization of the gels was performed on the Alliance Q9 Advanced machine (Uvitec). Images were processed in CorelDraw X7 software. For all experiments at least three independent replicates were performed with various concentrations of DnaA. To quantify the results, background-corrected band intensities were determined using ImageJ, values were normalized against the total signal in a lane in MS Excel, and plotted using GraphPad.

RESULTS

C. difficile DnaA Protein

Clostridioides difficile 630Δ*erm* encodes a homolog of the bacterial replication initiator protein DnaA (GenBank: CEJ96502.1; CD630DERM_00010). Alignment of the full-length *C. difficile* DnaA amino acid sequence with selected DnaA homologs from other organisms demonstrates a sequence identity of 35 to 67%, with an even higher similarity (57 to 83%, **Figure 1A**). *C. difficile* DnaA displays a greater sequence identity between the low-[G + C] Firmicutes (>60%). When compared with the extensively studied DnaA proteins from *E. coli* and *B. subtilis*, the full-length protein has 43 and 62% identity, and a similarity of 63 and 78%, respectively (**Figure 1A**).

To assess the structural properties of *C. difficile* DnaA, we predicted the secondary structure and generated a model of the protein using Phyre2 (Kelley et al., 2015; **Figure 1B**). The predicted DnaA model is based on three DnaA structures from different organisms: *A. aeolicus* (residues 101 to 318 and 334 to 437) (Erzberger et al., 2006) for Domain III and IV, and *B. subtilis* (residues 2 to 79) (Jameson et al., 2014) and *E. coli* (residues 5 to 97) (Abe et al., 2007) for Domain I and II.

Domain I of DnaA mediates interactions with a diverse set of regulators, and is involved in DnaA oligomerization (Zawilak-Pawlik et al., 2017; Nowaczyk-Cieszewska et al., 2020). We observe limited homology of *C. difficile* DnaA Domain I with the equivalent domain of the selected organisms (**Figure 1A**), although the overall fold is clearly conserved (**Figure 1B**). Nevertheless, some residues (P45, F48) appear to be conserved in most of the selected organisms (**Figure 1A**), though no functional role for these residues is known. Potentially, these residues might be involved in protein-protein interactions or DnaA oligomerization, as these functions have been mapped to Domain I of DnaA (Weigel et al., 1999; Abe et al., 2007; Natrajan et al., 2009; Jameson et al., 2014; Kim et al., 2017; Zawilak-Pawlik et al., 2017; Martin et al., 2018; Matthews and Simmons, 2019; Nowaczyk-Cieszewska et al., 2020).

Domain II is a flexible linker that is possibly involved in aiding the proper conformation of the DnaA domains, and thus requires a minimal length for DnaA function *in vivo* (Nozaki and Ogawa, 2008). No clear sequence similarity is observed on Domain II and modeling of the *C. difficile* DnaA protein suggests a putative disordered nature of this domain (**Figure 1**).

Domain III is responsible for binding to the co-factors ATP and ADP, and is in conjunction with Domain IV essential for DNA binding (Kawakami et al., 2005; Ozaki et al., 2008; Ozaki and Katayama, 2012). Within Domain III we readily identified the Walker A and Walker B motifs (WA and WB in **Figure 1A**) of the AAA+ fold (residues 135–317), crucial for binding and hydrolyzing ATP. This domain is highly conserved among all the selected organisms (**Figure 1A**) and comprises a structural center of β-sheets (**Figure 1B**, pink domain). Other features of the AAA+ ATPase fold are present and conserved between the organisms, such as the sensor I and sensor II motifs required for the nucleotide binding (I and II, **Figure 1A**). The arginine finger motif (the equivalent of R285 of *E. coli* DnaA in the VII box), important for the ATP dependent activation of DnaA (Kawakami et al., 2005), is conserved in *C. difficile* DnaA as well (R256 in motif box VII; **Figure 1A**).

The C-terminal Domain IV of the DnaA protein (residues 317–439, **Figure 1A**), contains the HTH motif required for the specific binding to DnaA-boxes (Erzberger et al., 2002; Zawilak et al., 2003). Previous studies identified several residues involved in specific interactions with the DnaA boxes, that bind through hydrogen bonds and van der Waals contacts with thymines present in the DNA sequence (Blaesing et al., 2000; Fujikawa et al., 2003; Tsodikov and Biswas, 2011). The residues are conserved among all Firmicutes and *E. coli,* including the residues R371 (position R399 in *E. coli*), P395 (P423), D405 (D433), H406 (H434), T407 (T435), and H411 (H439), (**Figure 1B** inset, red residues) (Fujikawa et al., 2003). Structural modeling of *C. difficile* DnaA predicts these residues to be exposed, providing an interface for DNA binding (**Figure 1B**). Residues involved in base-specific recognition of the DnaA box sequence are conserved between the Firmicutes and *E. coli* (**Figure 1A**), suggesting that *C. difficile* DnaA is likely to recognize the consensus DnaA box TTWTNCACA (Schaper and Messer, 1995). Notably, with the exception of a single arginine, these residues are not conserved between *C. difficile* and *Thermotoga maritima* DnaA (**Supplementary Figure S1**). As the latter recognizes an extended 12-bp motif (Ozaki et al., 2006; Richardson et al., 2019), this provides additional support for the notion that *C. difficile* DnaA recognizes a classical 9-bp DnaA box. In addition, residues found to be involved in non-specific interactions with the phosphate backbone of the DNA (some of which contribute

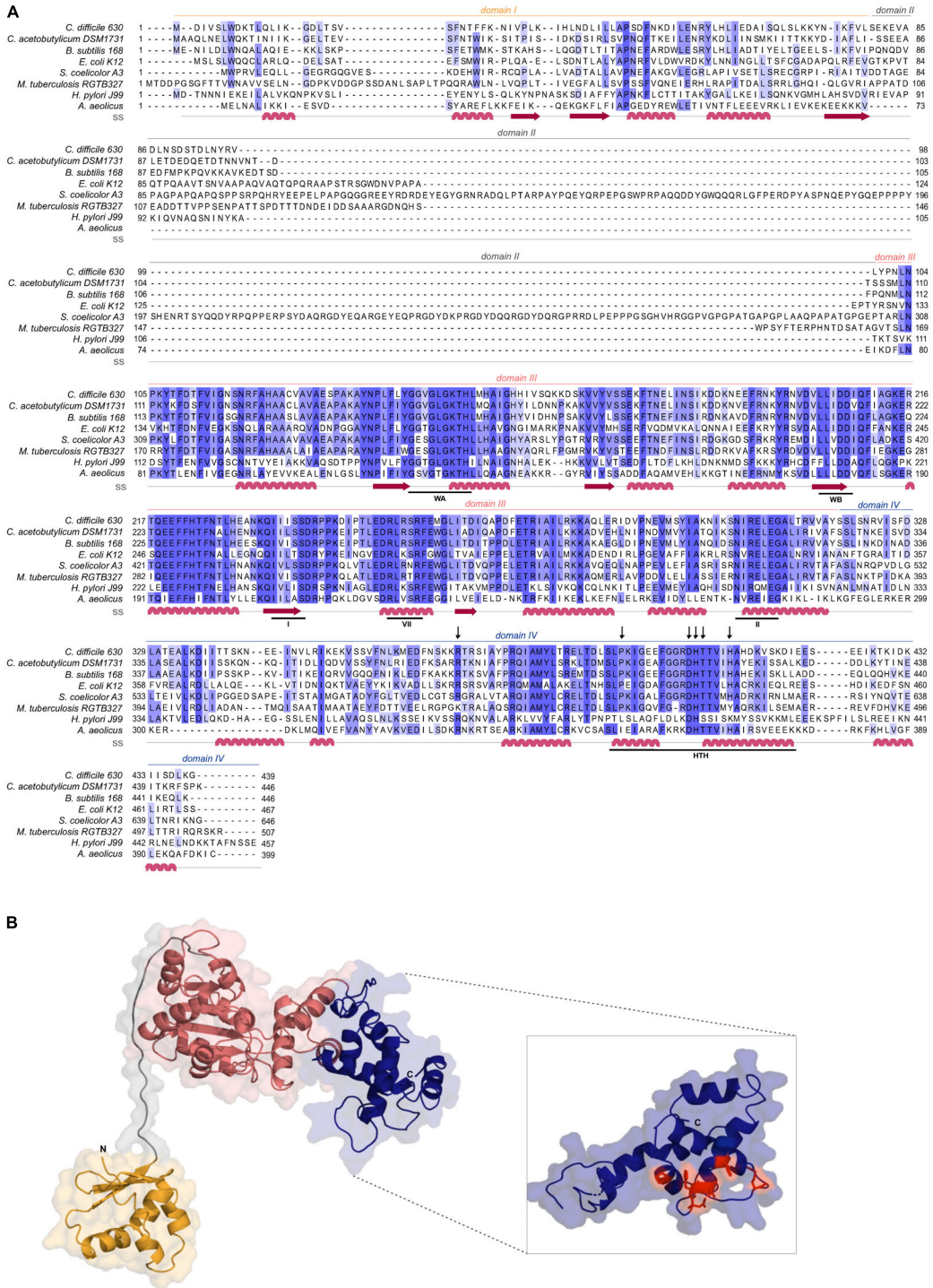


FIGURE 1 | *C. difficile* DnaA DNA-binding domain is conserved. **(A)** Multiple sequence alignment (PRALINE) of *C. difficile* DnaA with homologous proteins retrieved from GenBank. The amino acid sequences from *C. difficile* 630Δerm (CEJ96502.1), *C. acetobutylicum* DSM 1731 (AEI33799.1), *B. subtilis* 168 (NP_387882.1), *E. coli* K-12 (AMH32311.1), *S. coelicolor* A3(2) (TYP16779.1), *M. tuberculosis* RGTB327 (AFE14996.1), *H. pylori* J99 (Q9ZJ96.1) and *Aquifex aeolicus* (WP_010880157.1) were used. Residues are colored according to sequence identity conservation using blue shading (dark blue more conserved), as analyzed in Jalview. Secondary structure prediction (ss) is indicated, according to Phyre2 modeled structure. DnaA domains are represented, with the conserved AAA+ ATPase fold motifs Walker A, Walker B, VII box, sensor I and sensor II highlighted (WA, WB, I, VII, and II motifs), as well as the Domain IV helix-turn-helix (HTH). Residues involved in the base-specific recognition of the 9-mer DnaA box sequence are identified with an arrow. **(B)** Structural model of *C. difficile* DnaA determined by Phyre2. Domains are colored as in alignment. Both the N-terminus and the C-terminus are indicated in the figure. The DnaA Domain IV is enhanced (inset) with the DnaA-box binding specific residues represented in red sticks.

to sequence specificity) (Fujikawa et al., 2003; Tsodikov and Biswas, 2011) appear less conserved between the selected organisms (**Figure 1A**).

Expression and Purification of DnaA-6xHis

To allow for *in vitro* characterization of DnaA activity, we recombinantly expressed the *C. difficile* DnaA with a C-terminal 6xHis-tag in *E. coli* cells. To prevent the co-purification of *C. difficile* DnaA with host DnaA protein, *E. coli* strain CYB1002 was used (a kind gift of A. D. Grossman). This strain is a derivative of *E. coli* MS3898, that lacks the *dnaA* gene and replicates in a DnaA-independent fashion (Sutton and Kaguni, 1997). Induction of the DnaA-6xHis protein was confirmed by Coomassie staining and immunoblotting with anti-his antibody at the expected molecular weight of 51 kDa (**Supplementary Figure S2A**, red arrow). Upon overexpression of DnaA-6xHis, smaller fragments were observed, which accumulated with a prolonged time of expression (**Supplementary Figure S2A**), most likely corresponding to proteolytic fragments of the DnaA-6xHis protein.

Purification of the recombinant DnaA-6xHis showed a clear band at the expected size when eluted at 300 mM imidazole concentration, but several lower molecular size bands were observed (**Supplementary Figure S2B**). Therefore, the eluted fractions where further purified with size exclusion chromatography (SEC). This yielded a single product at the expected molecular weight of DnaA-6xHis, and its identity was confirmed by western-blot with anti-his antibody (**Supplementary Figure S2C**, red arrow). A minor band of lower molecular weight (approximately 38 kDa, < 1% of total protein) was observed (**Supplementary Figure S2C**, green asterisk), which may reflect some instability of the N-terminus of the DnaA-6xHis protein, as it appears to have retained the C-terminal 6xHis tag.

In silico Prediction of the *oriC* Region

To identify the *oriC* region and the elements that are part of it (DUE, DnaA-trio and DnaA boxes) we performed different prediction approaches in a stepwise procedure, as initially described (Mackiewicz et al., 2004).

We first analyzed the DNA asymmetry of the genome of *C. difficile* 630Δ*erm* (GenBank accession no. LN614756.1) (van Eijk et al., 2015), by plotting the normalized difference of the complementary nucleotides (GC-skew plot) (Necsulea and Lobry, 2007). *C. difficile* 630Δ*erm* has a circular genome of 4293049 bp and an average [G + C] content of 29.1%. We used the GenSkew Java Application[8] for determining the chromosomal asymmetry. Asymmetry changes in a GC-skew plot can be used to predict the origin of replication region and the terminus region of bacterial genomes. Based on this analysis, the origin is predicted at approximately position 1 of the chromosome. The terminus location is predicted at approximately 2.18 Mbp from the origin region (**Figure 2A**). These results were confirmed when artificially reassigning the starting position of the chromosomal assembly (data not shown). The gene organization in the putative origin region is *rnpA-rpmH-dnaA-dnaN* (position 4291488 to 2870, **Figure 2B**), identical to the origin of *B. subtilis* (Ogasawara et al., 1985; Briggs et al., 2012), and therefore encompasses the *dnaA* gene (CD630DERM_00010).

We next used the SIST program (Zhabinskaya et al., 2015) to localize putative DUEs in the intergenic regions in the chromosomal region predicted to contain the *oriC*. Hereafter we refer to these regions as *oriC1* (in the intergenic region of *rpmH-dnaA*) and *oriC2* (in the intergenic region *dnaA-dnaN*), in line with nomenclature in other organisms (Ogasawara et al., 1985; Donczew et al., 2012; **Figure 3B**). SIST identifies helically unstable AT-rich DNA stretches (Stress-Induced Duplex Destabilization regions; SIDDs) (Donczew et al., 2012; Zhabinskaya et al., 2015). In regions with a lower free energy ($G_{(x)}$ < y kcal/mol) the double-stranded helix has a high probability to become single-stranded DNA. With increasing negative superhelicity ($\sigma = -0.06$, **Figure 2C**, green line) regions of both *oriC1* and *oriC2* become single stranded DNA ($G_{(x)} < 2$ kcal/mol). At low negative superhelicity ($\sigma = -0.04$, **Figure 3C**, red line) short stretches of DNA of approximately 27 bp were identified with a significantly lower free energy. These regions with lower free energy at a negative superhelicity of −0.04 and −0.06 are potential DUE sites. The nucleotide sequence of the possible unwinding elements identified are represented in detail in **Figure 3** (gray boxes).

We then performed the identification of DnaA box clusters through a search of the consensus DnaA box TTWTNCACA containing up to one mismatch, using Pattern Locator (Mrazek and Xie, 2006). 22 putative DnaA boxes where identified in both the leading and lagging strand in the predicted *C. difficile oriC* regions (**Figure 3**, pink boxes), 14 in the *oriC1* region and 8 in the *oriC2* region. Both the consensus DnaA box TTWTNCACA and variant boxes are found. A cluster of DnaA boxes was proposed to contain at least three boxes with an average distance lower than 100 bp in between (Mackiewicz et al., 2004). At least one such cluster can be found in each origin region (**Figure 3**).

Though these are not crucial to origin function, we also manually identified the putative ribosomal binding sites for the annotated genes (**Figure 3**, dashed line) based on previously identified characteristics (Vellanoweth and Rabinowitz, 1992).

Finally, we manually predicted DnaA-trio sequences (3′-[G/A]A[T/A]$_{n>3}$-5′ preceded by a GC-cluster) in the predicted *oriC* regions, as this motif is required for successful replication in both *E. coli* and *B. subtilis* (Richardson et al., 2016) and can also be identified in *E. coli* (Katayama et al., 2017), though a role in binding of DnaA to ssDNA has yet to be experimentally demonstrated in this organism. We identified a clear DnaA-trio in the lagging strand upstream of a predicted DUE region in the *oriC2* region, with the nucleotide sequence 5′-CACCTACTACTATTACTACTATGA-3′ (**Figure 3**, light blue box), but no clear DnaA-trio was identified in the *oriC1* region.

From all the observations, we anticipate that a bipartite origin is located in the *dnaA* chromosomal region of *C. difficile* with unwinding occurring downstream of *dnaA*, at the *oriC2* region.

[8]http://genskew.csb.univie.ac.at/

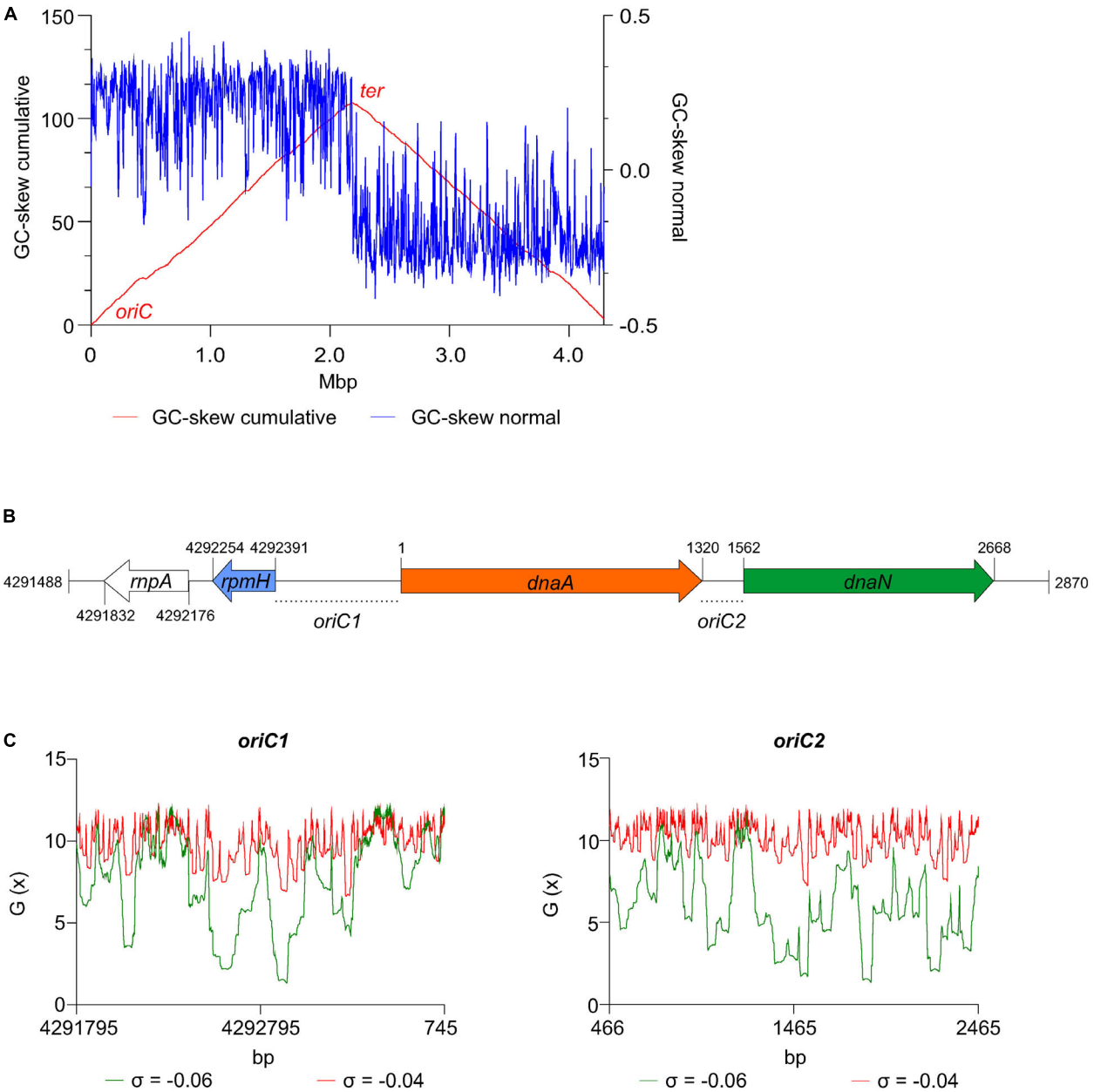


FIGURE 2 | Prediction of the *C. difficile* origin of replication. **(A)** GC skew analysis of the *C. difficile* 630Δ*erm* (LN614756.1) genome sequence. Normal GC skew analysis [(G – C)/(G + C)] performed on leading strand (blue line) and respective cumulative GC skew plot (red line). Calculations where performed with a window size of 4293 bp and a step size of 4293 bp. The *origin* (*oriC*) and *terminus (ter) regions* are indicated. **(B)** Representation of the predicted origin region and genomic context (from residues at position 4291488 to 2870 of the *C. difficile* 630Δ*erm* chromosome). The *rmpA, rpmH* (blue arrow), *dnaA* (orange arrow) and *dnaN* (green arrow) genes are indicated. Putative origins in intergenic regions are represented *oriC1* (*rpmH-dnaA*) and *oriC2* (*dnaA-dnaN*). **(C)** SIDD analysis of 2.0 kb fragments comprising *oriC1* (nucleotide 4291795 to 745) and *oriC2* (nucleotide 466 to 2465). Predicted free energies $G_{(x)}$ for duplex destabilization at a superhelical density of σ = –0.06 (green) or σ = –0.04 (red).

DnaA-Dependent Unwinding

To analyze DnaA-dependent unwinding of *oriC*, we used the purified *C. difficile* DnaA-6xHis protein and the predicted *oriC* sequence, to perform P1 nuclease assays as previously described (Sekimizu et al., 1988; Donczew et al., 2012). Localized melting resulting from DnaA activity exposes ssDNA to the action of the ssDNA-specific P1 nuclease. After incubation of a vector containing the *oriC* fragment with DnaA protein and cleavage by the P1 nuclease, the vector is purified and digested with different endonucleases to map the location of the unwound region.

We constructed vectors, based on pori1ori2 (Donczew et al., 2012), harboring *C. difficile oriC1* (pAP76) or *oriC2* (pAP83)

FIGURE 3 | Identification of the *C. difficile oriC* region. Nucleotide sequence of the *oriC1* region (nucleotide 4292328 to 48 of the *C. difficile* 630Δerm LN614756.1 genome sequence) and *oriC2* region (nucleotide 1274 to 1587). Identification of the possible unwinding AT-rich regions previously identified in the SIDD analysis (gray boxes). The putative DnaA boxes found are represented (pink boxes) and orientation in the leading **(right)** and lagging strand **(left)** are shown. Possible DnaA-trio sequence are denoted (light blue boxes). Coding sequence of the genes *rpmH* (blue arrow), *dnaA* (orange arrow) and *dnaN* (green arrow) and respective putative ribosome binding sites (dashed line) are indicated. Pattern identification is described in section "Materials and Methods."

individually (**Supplementary Figure S3A**), as well as the complete *oriC* region (pAP205) (**Figure 4A**). For a more accurate determination of the unwound region, the vectors were subjected to digestion by two different restriction enzymes (*Bgl*II and *Not*I), resulting in different restriction patterns. Limited spontaneous unwinding of the plasmid was observed in the *C. difficile oriC*-containing vectors (**Figure 4B** and **Supplementary Figure S3B**). No DnaA-dependent change in restriction pattern was observed when using the single *oriC* regions (**Supplementary Figure S3B**), suggesting *oriC1* and *oriC2* individually lack the requirements for DnaA-dependent unwinding.

We did observe a DnaA-dependent change in digestion patterns for the *oriC1oriC2*-containing vector pAP205 (**Figure 4**). Digestion of this vector with *Bgl*II in the absence of DnaA-6xHis and P1 nuclease resulted in a linear DNA fragment (4638 bp) due to the presence of a unique *Bgl*II restriction site (**Figure 4B**, upper panel, first lane). The addition of P1 nuclease leads to the appearance of a faint band between 1650 and 3000 bp (**Figure 4B**, upper panel, second lane), consistent with previous observations that the presence of a plasmid DUE can result in low-level spontaneous unwinding due to the inherent instability of these AT-rich regions (Jaworski et al., 2016). Upon the addition of the DnaA-6xHis protein the observed band becomes more intense, suggesting a strong increase in unwinding (**Figure 4B**, upper panel, red arrow).

Digestion of pAP205 with *Not*I in the absence of DnaA-6xHis and P1 nuclease results in fragments of 3804 and 842 bp, due to two *Not*I recognition sites in the vector (**Figure 4B**, lower panel, first lane). In the presence of just P1 nuclease, a similar low level of spontaneous unwinding is observed, resulting in the appearance of two additional faint bands, one between 1650 and 3000 bp and other between 1000 and 1650 bp (**Figure 4B**, lower panel, second lane). The addition of DnaA-6xHis results in an increase in intensity of both these bands in a dose dependent manner (**Figure 4B**, lower panel, red arrows).

We quantified the intensity of the bands from three independent P1 nuclease assays in order to determine the reproducibility of the assay (**Figures 4C,D** and **Supplementary Figure S4**). For the *Bgl*II-digested vector, we observed a DnaA-dependent increase of 20 to 60% of the total signal for the band between 1650 and 3000 bp (**Figure 4C**, band 2). For the *Not*I-digested vector, the signals of the second and third band increase from approximately 10% of the total signal to approximately 35% (1650–3000 bp, band 2) and 20% (1000–1650 bp, band 3) of total signal in the lane (**Figure 4D**). The observed increase was highly consistent, and appeared to saturate around 0.54–1 uM of DnaA (**Figures 4C,D**). The quantification also revealed a concomitant decrease in the signal for the upper bands in the gels of the *Bgl*II and *Not*I digests (**Supplementary Figure S4**, band 1).

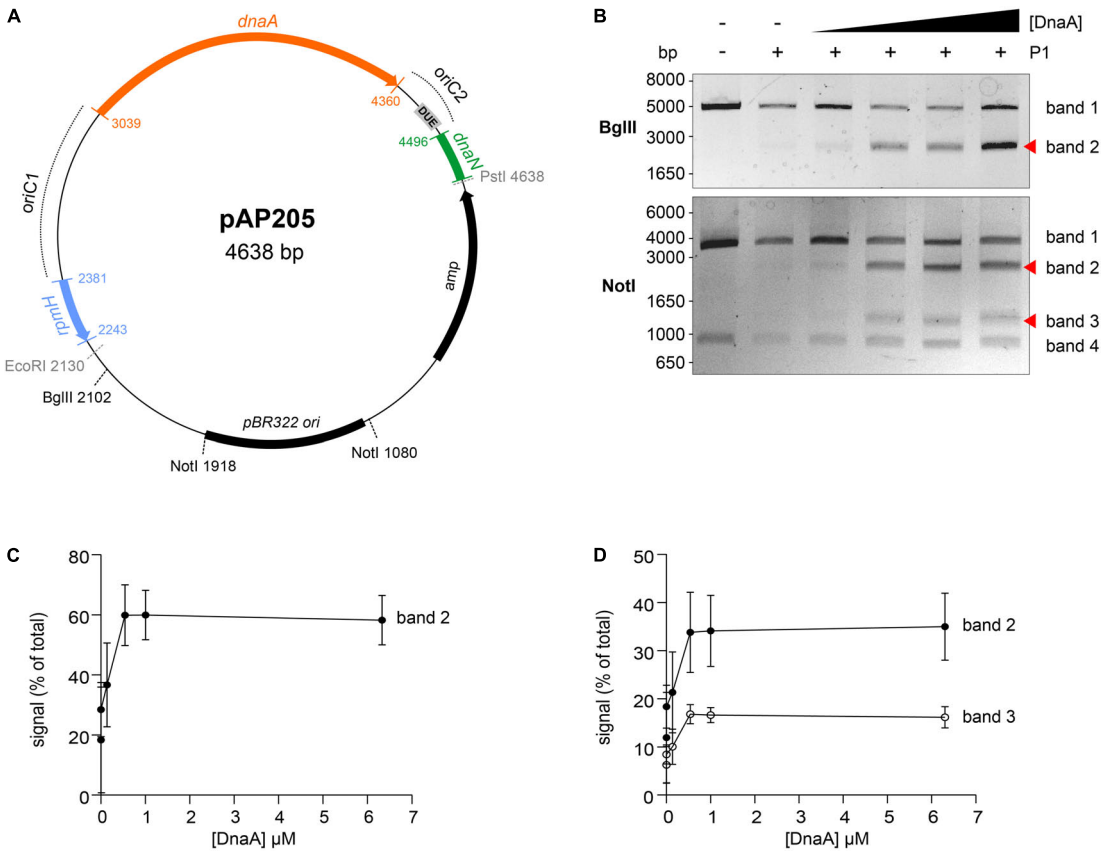


FIGURE 4 | Identification of the unwinding region in *C. difficile oriC*. **(A)** Representation of the *oriC1oriC2* containing vector pAP205 used in the P1 nuclease assay. The predicted *oriC1* and *oriC2* regions (dotted lines) and included genes are represented, *rpmH* (blue), *dnaA* (orange), and *dnaN* (green). The *bla* gene, the pBR322 plasmid origin of replication and the positions of used restriction sites are marked. The unwinding region (DUE) is denoted in a gray circle. **(B)** P1 nuclease assay of the *oriC1oriC2*-containing vector pAP205. Digestion of the vector (lane 1) with different restriction enzymes *Bgl*II (upper panel) or *Not*I (lower panel). Treatment of the fragments with P1 nuclease only (lane 2) and incubated with increasing amounts of *C. difficile* DnaA protein (lanes 3–6). The DNA fragments were separated in a 1% agarose gel and analyzed after ethidium bromide staining. Fragments resulting from DnaA-dependent unwinding are indicated with a red arrow (see Results for details). A typical result is shown. **(C)**. Quantification of band 2 (black circles) of the P1/*Bgl*II digested vector. **(D)**. Quantification of bands 2 (black circles) and 3 (open circles) of the P1/*Not*I digested vector. For panels **(C,D)**, error bars indicate the standard deviation of the mean of $n = 3$ independent experiments.

The DnaA-dependent appearance of the ∼2000 bp band in the *Bgl*II digest, and the ∼1200 and ∼2200 bp bands in the *Not*I digest localize the DnaA-dependent unwinding of the *C. difficile oriC* in the *oriC2* region (**Figure 4A**, gray rectangle, DUE). Moreover, these results suggest that *C. difficile* has a bipartite origin of replication, as successful DnaA-dependent unwinding of *C. difficile* in the *oriC2* region requires both *oriC* regions (*oriC1* and *oriC2*).

Conservation of the Origin Organization in Related Clostridia

Our results suggest that the origin organization of *C. difficile* resembles that of a more distantly related Firmicute, *B. subtilis*. To extend our observations, we evaluated the genomic organization of the *oriC* region in different organisms phylogenetically related to *C. difficile*. We followed a similar approach as described above for *C. difficile* 630Δ*erm*, taking advantage of the DoriC 10.0 database (Luo and Gao, 2019). Importantly, our results with respect to the *C. difficile* origin of replication described above were largely congruent with the DoriC 10.0 database despite being based on different methods (a notable exception is the prediction for *C. difficile* strain 630; data not shown). We retrieved the predicted *oriC* regions from the DoriC 10.0 database and performed an in-depth analysis of these regions for the closely related *C. difficile* strain R20291 (NC_013316.1), as well as the more distantly related *C. botulinum* A Hall (NC_009698.1), *C. sordelli* AM370 (NZ_CP014150), *C. acetobutylicum* DSM 1731 (NC_015687.1), *C. perfringens* str.13 (NC_003366.1), and *C. tetani* E88 (NC_004557.1) (**Table 1**).

Similar to *C. difficile* 630Δ*erm*, the genomic context of the origin contains the *rpmH-dnaA-dnaN* region for most of the clostridia selected and mirrors that of *B. subtilis* (**Figure 5**). The only exception is *C. tetani* E88 where the uncharacterized CLOTE0041 gene lies upstream of the *dnaA-dnaN* cluster (**Figure 5**).

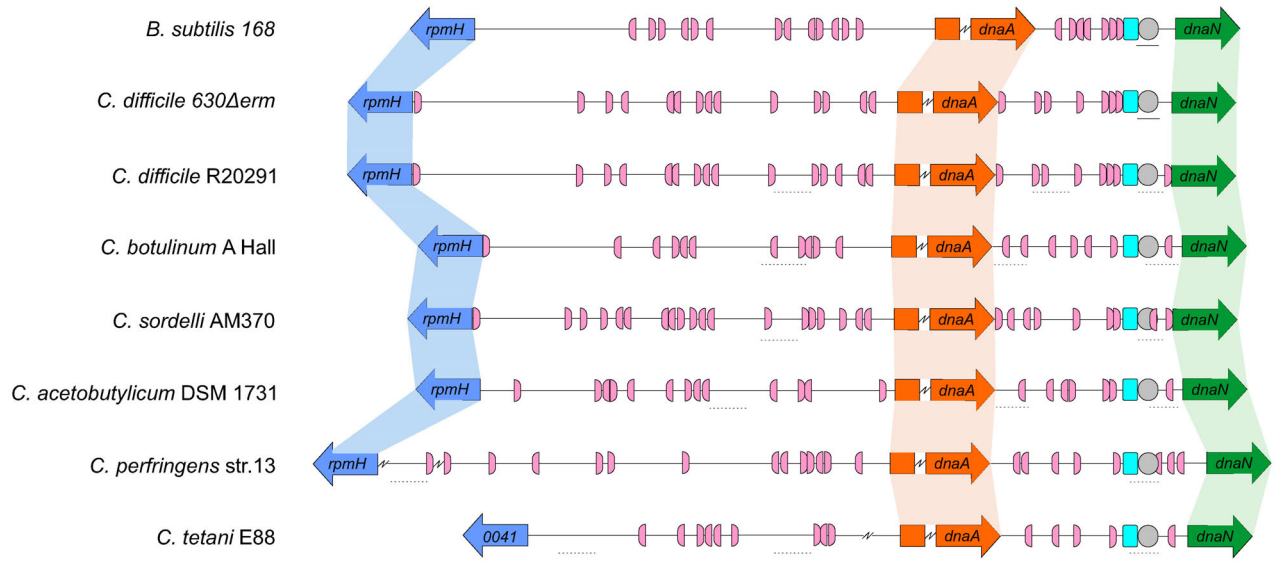


FIGURE 5 | Comparison of the clostridial *oriC* regions. Representation of the origin region and genomic context of *B. subtilis*, *C. difficile* 630Δ*erm* chromosome and the predicted regions for *C. difficile* R20291, *C. botulinum A Hall, C. sordelli* AM370, *C. acetobutylicum* DSM 1731, *C. perfringens* str.13, *C. tetani* E88 (see **Table 1**). The *rpmH* (blue arrow), *dnaA* (orange arrow) and *dnaN* (green arrow) genes are indicated. Predicted DnaA-boxes are indicated by pink boxes and orientation on the leading (right) and lagging strand (left) are shown. Identification of the experimentally identified unwinding AT-rich regions (lines) and the SIDD-predicted helical instability are shown (dashed lines). The putative DUE is denoted (gray circle). Possible DnaA-trio sequences are shown in light blue boxes. See section "Materials and Methods" for detailed information. Alignment of the represented chromosomal regions is based on the location of the DnaA-trio.

We also identified the possible DnaA boxes for the selected clostridia (**Figure 5**, pink semi-circle). Across the analyzed clostridia, *oriC1* region presented more variability in the number of putative DnaA boxes, from 9 to 19, whereas *oriC2* contained 5 to 9 DnaA boxes, with *C. tetani* E88 with the lowest number of possible DnaA boxes, both at the *oriC1* (9 boxes) and *oriC2* (5 boxes) regions (**Figure 5**, pink semi-circle). In all the organisms we observe at least 1 DnaA cluster in each origin region, as also observed for *C. difficile* 630Δ*erm*.

Prediction of DUEs using the SIST program (Zhabinskaya et al., 2015) identified several helically unstable regions that are candidate sites for unwinding (**Figure 5**, dashed lines, and **Supplementary Figure S5**). Notably, in all cases one such region in *oriC2* (**Figure 5**, gray circle) is preceded immediately by the manually identified DnaA-trio (**Figure 5**, light blue circle). Based on our experimental data for *C. difficile* 630Δ*erm*, we suggest that in all analyzed clostridia, DnaA-dependent unwinding occurs at a conserved DUE downstream of the DnaA-trio in the *oriC2* region (**Figure 5**).

DISCUSSION

Chromosomal replication is an essential process for the survival of the cell. In most bacteria DnaA protein is the initiator protein for replication and through a cascade of events leads to the successful loading of the replication complex onto the origin of replication (Duderstadt et al., 2011; Chodavarapu and Kaguni, 2016).

Initial characterization of bacterial replication has been assessed in the model organisms *E. coli* and *B. subtilis* (Jameson and Wilkinson, 2017). Despite the similarities (location in an intergenic region, presence of a DUE, several DnaA boxes in both orientations) the structure of the replication origins and the regulation mechanisms are variable among bacteria (Wolanski et al., 2014). In contrast to *E. coli*, the *B. subtilis* origin region is bipartite, with two intergenic regions upstream and downstream of the *dnaA* gene. In *C. difficile* the genomic organization in the predicted cluster *rnpA-rpmH-dnaA-dnaN,* and the presence of [A + T]-rich sequences in the intergenic regions is consistent with a bipartite origin, as in *B. subtilis* (**Figure 3**).

The origin region contains several DnaA-boxes with different properties that are recognized by the DnaA protein. The specific binding of DnaA to the DnaA-boxes is mediated mainly through Domain IV of the DnaA protein. From DNA bound structures of DnaA it was possible to identify several residues involved in the contact with the DnaA boxes, some of which confer specificity (Blaesing et al., 2000; Fujikawa et al., 2003; Tsodikov and Biswas, 2011). Analysis of the of *C. difficile* DnaA homology in Domain IV did not show any difference in the residues involved on the DnaA-box specificity (**Figure 1**, vertical arrows), suggesting the same consensus motif conservation as the DnaA-box TTWTNCACA for *E. coli* (Schaper and Messer, 1995). The conserved DnaA-box motif allowed us to identify several DnaA boxes along the intergenic regions of the *oriC*. Like in the bipartite origin of *B. subtilis,* we identified at least one cluster of DnaA-boxes in the *C. difficile oriC1* and *oriC2* regions (**Figures 3**, **5**). In the case of *B. subtilis*, it has been shown that different DnaA boxes fulfill different roles in replication

initiation: two out of three DnaA boxes immediately upstream of the DnaA-trio are part of the basal unwinding system (i.e., required for DnaA-dependent strand separation), whereas other DnaA affect coordination and regulation of DNA replication (Richardson et al., 2019). For *C. difficile*, we also find three DnaA boxes immediately upstream of the DnaA trio (**Figure 3** and **Supplementary Figure S6**), but the role of these boxes has not been experimentally verified to date.

The P1 nuclease assays place a region in which DnaA-dependent unwinding occurs in the *oriC2* region of *C. difficile*, supported by the presence of the several features on the *oriC2*, such as the identified DUE and DnaA-trio, both required for unwinding (Kowalski and Eddy, 1989; Richardson et al., 2016). The presence of both *oriC* regions (*oriC1* and *oriC2*) is required for melting *in vitro*, as observed for other bipartite origins (Wolanski et al., 2014). In contrast to the bipartite origin identified in *H. pylori* (Donczew et al., 2012), we did not observe unwinding of the *oriC2* region alone. Though this may be a specific aspect of *C. difficile oriC2*, we cannot exclude that differences in the experimental setup (e.g. DnaA protein purification) could affect these observations. Nevertheless, our data are consistent with DnaA binding the DnaA-box clusters in both *oriC* regions, leading to potential DnaA oligomerization, loop formation, and unwinding at the [A + T]-rich DUE site.

When analyzing the origin region between different clostridia, features similar to those of *C. difficile* are observed, such as conservation of DnaA-box clusters within both *oriC* regions in the vicinity of the *dnaA* gene. Similar to *C. difficile* and *B. subtilis*, a putative DUE element, preceded by the DnaA-trio, was also located within the *oriC2* region (**Figures 4**, **6**). Thus, the overall origin organization and mechanism of DNA replication initiation is likely to be conserved within the Firmicutes (Briggs et al., 2012). As spacing of the DnaA-boxes are determinants for the species-specific effective replication (Zawilak et al., 2003; Zawilak-Pawlik et al., 2005), these similarities do no exclude the possibilities that subtle differences in replication initiation exist, and further studies are required. For instance, our work does not address which DnaA boxes in either *oriC1* or *oriC2* are important for unwinding, and whether the requirement is due to DnaA-dependent changes in structure of origin DNA (as has been shown for *B. subtilis*) (Richardson et al., 2019), or as a *cis*-acting regulatory element like DARS/DatA (Katayama et al., 2010, 2017). Further experiments could provide insights into the DnaA-box conservation and affinities and establish which DnaA boxes are crucial for origin firing and/or transcriptional regulation.

Several proteins can interact with the *oriC* region or DnaA, including YabA, Rok, DnaD/DnaB, Soj and HU (Briggs et al., 2012; Jameson and Wilkinson, 2017). In doing so they shape the origin conformation and/or stabilize the DnaA filament or the unwound region, consequently affecting replication initiation.

YabA or Rok affect *B. subtilis* replication initiation (Goranov et al., 2009; Schenk et al., 2017; Seid et al., 2017), but no homologs of these proteins have been identified in *C. difficile* (van Eijk et al., 2017). Similarly, no homologs are identified of other well-characterized DnaA-interacting proteins from gram-negative bacteria (van Eijk et al., 2017), such as Hda, DiaA/HobA (Zawilak-Pawlik et al., 2017) or HdaB (Frandi and Collier, 2020); it is unknown how *C. difficile* regulates DnaA activity.

In *B. subtilis*, DnaD, DnaB, and DnaI helicase loader proteins associate sequentially with the origin region resulting in the recruitment of the DnaC helicase protein (Marsin et al., 2001; Velten et al., 2003; Smits et al., 2010; Jameson and Wilkinson, 2017). In *B. subtilis*, DnaD binds to DnaA and it is postulated that this affects the stability of the DnaA filament and consequently the unwinding of the *oriC* (Ishigo-Oka et al., 2001; Martin et al., 2018; Matthews and Simmons, 2019). *B. subtilis* DnaB protein also affects the DNA topology and has been shown to be important for recruiting *oriC* to the membrane (Rokop et al., 2004; Zhang et al., 2005). *C. difficile* lacks a homolog for the DnaB protein, although the closest homolog of the DnaD protein (CD3653) (van Eijk et al., 2017) may perform similar functions in the origin remodeling (van Eijk et al., 2016). Direct interaction of DnaA-DnaD through the DnaA Domain I was structurally determined and the residues present at the interface were solved (Martin et al., 2018). Despite high variability of this domain between organisms, half of the identified contacts for the DnaA-DnaD interaction are conserved within *C. difficile*, the S22 (S23 in *B. subtilis* DnaA), T25 (T26), F48 (F49), D51 (D52) and L68 (L69) (**Figure 1**; Martin et al., 2018; Matthews and Simmons, 2019). This might suggest a similar interaction surface for CD3653 on *C. difficile* DnaA. A characterization of the putative interaction between CD3653 and DnaA, and the resulting effect on DnaA oligomerization and origin melting awaits purification and functional characterization of CD3653.

The Soj protein, also involved in chromosome segregation, has been shown to interact with DnaA via Domain III, regulating DnaA-filament formation (Scholefield et al., 2012) and the *C. difficile* encodes at least one uncharacterized Soj homolog, but a role in DNA replication has not been experimentally demonstrated.

Bacterial histone-like proteins (such as HU and HBsu) can modulate DNA topology and might therefore influence *oriC* unwinding and replication initiation. However, the importance of HU for replication initiation has only been clearly demonstrated for *E. coli* (Krause et al., 1997; Chodavarapu et al., 2008). Several studies have shown HU independent origin unwinding even in gram-negative bacteria (Donczew et al., 2012; Makowski et al., 2016; Jaworski et al., 2018; Plachetka et al., 2019), suggesting that HU-dependence of origin unwinding may be limited to a narrow phylogenetic group. *C. difficile* encodes a homolog of HU, HupA (Oliveira Paiva et al., 2019) but whether this protein plays a role in DNA replication initiation remains to be established.

Finally, Spo0A, the master regulator of sporulation, binds to several Spo0A-boxes present in this the *oriC* region in *B. subtilis* (Boonstra et al., 2013). Some of the Spo0A-boxes partially overlap with DnaA-boxes and binding of Spo0A can prevent the DnaA-mediated unwinding, thus playing a significant role on the coordination of between cell replication and sporulation (Boonstra et al., 2013). In *C. difficile,* Spo0A-binding has previously been investigated (Rosenbusch et al., 2012), but a role in DNA replication has not been assessed.

For all the regulators with a *C. difficile* homolog discussed above (i.e. CD3653, Soj, HupA, and Spo0A), further studies can

be envisioned employing the P1 nuclease assays described here to assess the effects on DnaA-mediated unwinding of the origin. Our experiments show, however, they are not strictly required for origin unwinding (**Figure 4**).

In summary, through a combination of different *in silico* predictions and *in vitro* studies, we have shown the DnaA-dependent unwinding in the *dnaA-dnaN* intergenic region in the bipartite *C. difficile* origin of replication. We have analyzed the putative origin of replication in different clostridia and a conserved organization is observed throughout the Firmicutes, although different mechanisms and regulation could be behind the initiation of replication. The present study is the first to characterize the origin region of *C. difficile* and forms the start to further unravel the mechanism behind the DnaA-dependent regulation of *C. difficile* initiation of replication.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

AO and WS designed the experiments, analyzed the data and wrote the manuscript. AO and CW performed the *in silico* analyses. AO, EE, and AF performed the experiments. All authors read and approved the final version for submission.

FUNDING

This work in the group of WS was supported by a Vidi Fellowship (Grant No: 864.13.003) of the Netherlands Organization for Scientific Research (NWO) and a Gisela Thier Fellowship from the Leiden University Medical Center.

ACKNOWLEDGMENTS

We thank Alan Grossman for kindly providing the pAV13 vector and *E. coli* strain CYB1002. We also thank Anna Zawilak-Pawlik for kindly providing the pori1ori2 vector and expert help in setting up the P1 assays. We would like to thank Luís Sousa for their help with the SIDD and Pattern Locator coding files.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.581401/full#supplementary-material

REFERENCES

Abe, Y., Jo, T., Matsuda, Y., Matsunaga, C., Katayama, T., and 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. doi: 10.1074/jbc.M701841200

Bawono, P., and Heringa, J. (2014). PRALINE: a versatile multiple sequence alignment toolkit. *Methods Mol. Biol.* 1079, 245–262. doi: 10.1007/978-1-62703-646-7_16

Bazin, A., Cherrier, M. V., Gutsche, I., Timmins, J., and Terradot, L. (2015). Structure and primase-mediated activation of a bacterial dodecameric replicative helicase. *Nucleic Acids Res.* 43, 8564–8576. doi: 10.1093/nar/gkv792

Blaesing, F., Weigel, C., Welzeck, M., and Messer, W. (2000). Analysis of the DNA-binding domain of *Escherichia coli* DnaA protein. *Mol. Microbiol.* 36, 557–569. doi: 10.1046/j.1365-2958.2000.01881.x

Bleichert, F., Botchan, M. R., and Berger, J. M. (2017). Mechanisms for initiating cellular DNA replication. *Science* 355:eaah6317. doi: 10.1126/science.aah6317

Boonstra, M., de Jong, I. G., Scholefield, G., Murray, H., Kuipers, O. P., and Veening, J. W. (2013). Spo0A regulates chromosome copy number during sporulation by directly binding to the origin of replication in *Bacillus subtilis*. *Mol. Microbiol.* 87, 925–938. doi: 10.1111/mmi.12141

Briggs, G. S., Smits, W. K., and Soultanas, P. (2012). Chromosomal replication initiation machinery of low-G+C-content Firmicutes. *J. Bacteriol.* 194, 5162–5170.

Cho, E., Ogasawara, N., and Ishikawa, S. (2008). The functional analysis of YabA, which interacts with DnaA and regulates initiation of chromosome replication in Bacillus subtils. *Genes Genet. Syst.* 83, 111–125. doi: 10.1266/ggs.83.111

Chodavarapu, S., Felczak, M. M., Yaniv, J. R., and Kaguni, J. M. (2008). Escherichia coli DnaA interacts with HU in initiation at the *E. coli* replication origin. *Mol. Microbiol.* 67, 781–792. doi: 10.1111/j.1365-2958.2007.06094.x

Chodavarapu, S., and Kaguni, J. M. (2016). Replication Initiation in Bacteria. *Enzymes* 39, 1–30. doi: 10.1016/bs.enz.2016.03.001

Crobach, M. J. T., Vernon, J. J., Loo, V. G., Kong, L. Y., Pechine, S., Wilcox, M. H., et al. (2018). Understanding *Clostridium difficile* Colonization. *Clin. Microbiol. Rev.* 31:e00021-17. doi: 10.1128/CMR.00021-17

Davey, M. J., and O'Donnell, M. (2003). Replicative helicase loaders: ring breakers and ring makers. *Curr. Biol.* 13, R594–R596. doi: 10.1016/s0960-9822(03)00523-2

Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwinska, J., and Zawilak-Pawlik, A. (2012). *Helicobacter pylori* oriC–the first bipartite origin of chromosome replication in Gram-negative bacteria. *Nucleic Acids Res.* 40, 9647–9660. doi: 10.1093/nar/gks742

Duderstadt, K. E., Chuang, K., and Berger, J. M. (2011). DNA stretching by bacterial initiators promotes replication origin opening. *Nature* 478, 209–213. doi: 10.1038/nature10455

Ekundayo, B., and Bleichert, F. (2019). Origins of DNA replication. *PLoS Genet.* 15:e1008320. doi: 10.1371/journal.pgen.1008320

Erzberger, J. P., Mott, M. L., and Berger, J. M. (2006). Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* 13, 676–683. doi: 10.1038/nsmb1115

Erzberger, J. P., Pirruccello, M. M., and Berger, J. M. (2002). The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. *EMBO J.* 21, 4763–4773.

Fossum, S., De Pascale, G., Weigel, C., Messer, W., Donadio, S., and Skarstad, K. (2008). A robust screen for novel antibiotics: specific knockout of the initiator of bacterial DNA replication. *FEMS Microbiol. Lett.* 281, 210–214. doi: 10.1111/j.1574-6968.2008.01103.x

Frandi, A., and Collier, J. (2020). HdaB: a novel and conserved DnaA-related protein that targets the RIDA process to stimulate replication initiation. *Nucleic Acids Res.* 48, 2412–2423. doi: 10.1093/nar/gkz1193

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., et al. (2003). Structural basis of replication origin recognition by the DnaA protein. *Nucleic Acids Res.* 31, 2077–2086.

Goranov, A. I., Breier, A. M., Merrikh, H., and Grossman, A. D. (2009). YabA of *Bacillus subtilis* controls DnaA-mediated replication initiation but not the

transcriptional response to replication stress. *Mol. Microbiol.* 74, 454–466. doi: 10.1111/j.1365-2958.2009.06876.x

Grimwade, J. E., and Leonard, A. C. (2017). Targeting the bacterial orisome in the search for new antibiotics. *Front. Microbiol.* 8:2352. doi: 10.3389/fmicb.2017.02352

Ishigo-Oka, D., Ogasawara, N., and Moriya, S. (2001). DnaD protein of *Bacillus subtilis* interacts with DnaA, the initiator protein of replication. *J. Bacteriol.* 183, 2148–2150. doi: 10.1128/JB.183.6.2148-2150.2001

Jameson, K. H., Rostami, N., Fogg, M. J., Turkenburg, J. P., Grahl, A., Murray, H., et al. (2014). Structure and interactions of the *Bacillus subtilis* sporulation inhibitor of DNA replication, SirA, with domain I of DnaA. *Mol. Microbiol.* 93, 975–991. doi: 10.1111/mmi.12713

Jameson, K. H., and Wilkinson, A. J. (2017). Control of initiation of DNA Replication in *Bacillus subtilis* and *Escherichia coli*. *Genes* 8:22. doi: 10.3390/genes8010022

Jaworski, P., Donczew, R., Mielke, T., Thiel, M., Oldziej, S., Weigel, C., et al. (2016). Unique and universal features of epsilonproteobacterial origins of chromosome replication and DnaA-DnaA Box Interactions. *Front. Microbiol.* 7:1555. doi: 10.3389/fmicb.2016.01555

Jaworski, P., Donczew, R., Mielke, T., Weigel, C., Stingl, K., and Zawilak-Pawlik, A. (2018). Structure and function of the campylobacter jejuni chromosome replication origin. *Front. Microbiol.* 9:1533. doi: 10.3389/fmicb.2018.01533

Katayama, T., Kasho, K., and Kawakami, H. (2017). The DnaA Cycle in *Escherichia coli*: activation, function and inactivation of the initiator protein. *Front. Microbiol.* 8:2496. doi: 10.3389/fmicb.2017.02496

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8, 163–170. doi: 10.1038/nrmicro2314

Kawakami, H., Keyamura, K., and 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. doi: 10.1074/jbc.M502764200

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kim, J. S., Nanfara, M. T., Chodavarapu, S., Jin, K. S., Babu, V. M. P., Ghazy, M. A., et al. (2017). Dynamic assembly of Hda and the sliding clamp in the regulation of replication licensing. *Nucleic Acids Res.* 45, 3888–3905. doi: 10.1093/nar/gkx081

Kowalski, D., and Eddy, M. J. (1989). The DNA unwinding element: a novel, cis-acting component that facilitates opening of the *Escherichia coli* replication origin. *EMBO J.* 8, 4335–4344.

Krause, M., Ruckert, B., Lurz, R., and Messer, W. (1997). Complexes at the replication origin of *Bacillus subtilis* with homologous and heterologous DnaA protein. *J. Mol. Biol.* 274, 365–380. doi: 10.1006/jmbi.1997.1404

Lawson, P. A., Citron, D. M., Tyrrell, K. L., and Finegold, S. M. (2016). Reclassification of *Clostridium difficile* as Clostridioides difficile (Hall and O'Toole 1935) Prevot 1938. *Anaerobe* 40, 95–99. doi: 10.1016/j.anaerobe.2016.06.008

Luo, H., and Gao, F. (2019). DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids Res.* 47, D74–D77. doi: 10.1093/nar/gky1014

Mackiewicz, P., Zakrzewska-Czerwinska, J., Zawilak, A., Dudek, M. R., and Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Res.* 32, 3781–3791. doi: 10.1093/nar/gkh699

Majka, J., Messer, W., Schrempf, H., and Zakrzewska-Czerwinska, J. (1997). Purification and characterization of the *Streptomyces lividans* initiator protein DnaA. *J. Bacteriol.* 179, 2426–2432.

Makowski, L., Donczew, R., Weigel, C., Zawilak-Pawlik, A., and Zakrzewska-Czerwinska, J. (2016). Initiation of Chromosomal Replication in Predatory Bacterium *Bdellovibrio* bacteriovorus. *Front. Microbiol.* 7:1898. doi: 10.3389/fmicb.2016.01898

Marsin, S., McGovern, S., Ehrlich, S. D., Bruand, C., and Polard, P. (2001). Early steps of *Bacillus subtilis* primosome assembly. *J. Biol. Chem.* 276, 45818–45825. doi: 10.1074/jbc.M101996200

Martin, E., Williams, H. E. L., Pitoulias, M., Stevens, D., Winterhalter, C., Craggs, T. D., et al. (2018). DNA replication initiation in *Bacillus subtilis*: structural and functional characterization of the essential DnaA-DnaD interaction. *Nucleic Acids Res.* 47, 2101–2112. doi: 10.1093/nar/gky1220

Matthews, L. A., and Simmons, L. A. (2019). Cryptic protein interactions regulate DNA replication initiation. *Mol. Microbiol.* 111, 118–130. doi: 10.1111/mmi.14142

Moriya, S., Fukuoka, T., Ogasawara, N., and Yoshikawa, H. (1988). Regulation of initiation of the chromosomal replication by DnaA-boxes in the origin region of the *Bacillus subtilis* chromosome. *EMBO J.* 7, 2911–2917.

Mrazek, J., and Xie, S. (2006). Pattern locator: a new tool for finding local sequence patterns in genomic DNA sequences. *Bioinformatics* 22, 3099–3100. doi: 10.1093/bioinformatics/btl551

Murray, H., and Koh, A. (2014). Multiple regulatory systems coordinate DNA replication with cell growth in *Bacillus subtilis*. *PLoS Genet.* 10:e1004731. doi: 10.1371/journal.pgen.1004731

Natrajan, G., Noirot-Gros, M. F., Zawilak-Pawlik, A., Kapp, U., and Terradot, L. (2009). The structure of a DnaA/HobA complex from *Helicobacter* pylori provides insight into regulation of DNA replication in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21115–21120. doi: 10.1073/pnas.0908966106

Necsulea, A., and Lobry, J. R. (2007). A new method for assessing the effect of replication on DNA base composition asymmetry. *Mol. Biol. Evol.* 24, 2169–2179. doi: 10.1093/molbev/msm148

Nowaczyk-Cieszewska, M., Zyla-Uklejewicz, D., Noszka, M., Jaworski, P., Mielke, T., and Zawilak-Pawlik, A. M. (2020). The role of *Helicobacter* pylori DnaA domain I in orisome assembly on a bipartite origin of chromosome replication. *Mol. Microbiol.* 113, 338–355. doi: 10.1111/mmi.14423

Nozaki, S., and Ogawa, T. (2008). Determination of the minimum domain II size of *Escherichia coli* DnaA protein essential for cell viability. *Microbiology* 154(Pt 11), 3379–3384. doi: 10.1099/mic.0.2008/019745-0

O'Donnell, M., Langston, L., and Stillman, B. (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.* 5:a010108. doi: 10.1101/cshperspect.a010108

Ogasawara, N., Moriya, S., von Meyenburg, K., Hansen, F. G., and Yoshikawa, H. (1985). Conservation of genes and their organization in the chromosomal replication origin region of *Bacillus subtilis* and *Escherichia coli*. *EMBO J.* 4, 3345–3350.

Ogasawara, N., and Yoshikawa, H. (1992). Genes and their organization in the replication origin region of the bacterial chromosome. *Mol. Microbiol.* 6, 629–634. doi: 10.1111/j.1365-2958.1992.tb01510.x

Oliveira Paiva, A. M., Friggen, A. H., Qin, L., Douwes, R., Dame, R. T., and Smits, W. K. (2019). The Bacterial Chromatin Protein HupA Can Remodel DNA and Associates with the Nucleoid in *Clostridium difficile*. *J. Mol. Biol.* 431, 653–672. doi: 10.1016/j.jmb.2019.01.001

Ozaki, S., Fujimitsu, K., Kurumizaka, H., and Katayama, T. (2006). The DnaA homolog of the hyperthermophilic eubacterium Thermotoga maritima forms an open complex with a minimal 149-bp origin region in an ATP-dependent manner. *Genes Cells* 11, 425–438. doi: 10.1111/j.1365-2443.2006.00950.x

Ozaki, S., and Katayama, T. (2012). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.* 40, 1648–1665. doi: 10.1093/nar/gkr832

Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S. Y., et al. (2008). A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J. Biol. Chem.* 283, 8351–8362. doi: 10.1074/jbc.M708684200

Ozaki, S., Noguchi, Y., Hayashi, Y., Miyazaki, E., and Katayama, T. (2012). Differentiation of the DnaA-oriC subcomplex for DNA unwinding in a replication initiation complex. *J. Biol. Chem.* 287, 37458–37471. doi: 10.1074/jbc.M112.372052

Patel, M. J., Bhatia, L., Yilmaz, G., Biswas-Fiss, E. E., and Biswas, S. B. (2017). Multiple conformational states of DnaA protein regulate its interaction with DnaA boxes in the initiation of DNA replication. *Biochim. Biophys. Acta* 1861, 2165–2174. doi: 10.1016/j.bbagen.2017.06.013

Plachetka, M., Zyla-Uklejewicz, D., Weigel, C., Donczew, R., Donczew, M., Jakimowicz, D., et al. (2019). Streptomycete origin of chromosomal replication with two putative unwinding elements. *Microbiology* 165, 1365–1375. doi: 10.1099/mic.0.000859

Richardson, T. T., Harran, O., and Murray, H. (2016). The bacterial DnaA-trio replication origin element specifies single-stranded DNA initiator binding. *Nature* 534, 412–416. doi: 10.1038/nature17962

Richardson, T. T., Stevens, D., Pelliciari, S., Harran, O., Sperlea, T., and Murray, H. (2019). Identification of a basal system for unwinding a bacterial chromosome origin. *EMBO J.* 38:e101649. doi: 10.15252/embj.2019101649

Rokop, M. E., Auchtung, J. M., and Grossman, A. D. (2004). Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of *Bacillus subtilis*. *Mol. Microbiol.* 52, 1757–1767. doi: 10.1111/j.1365-2958.2004.04091.x

Rosenbusch, K. E., Bakker, D., Kuijper, E. J., and Smits, W. K. (2012). *C. difficile* 630Derm Spo0A Regulates Sporulation, but Does Not Contribute to Toxin Production, by Direct High-Affinity Binding to Target DNA. *PLoS One* 7:e48608. doi: 10.1371/journal.pone.0048608

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning : A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Saxena, R., Fingland, N., Patil, D., Sharma, A. K., and Crooke, E. (2013). 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. doi: 10.3390/ijms14048517

Schaper, S., and Messer, W. (1995). Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. *J. Biol. Chem.* 270, 17622–17626. doi: 10.1074/jbc.270.29.17622

Schenk, K., Hervas, A. B., Rosch, T. C., Eisemann, M., Schmitt, B. A., Dahlke, S., et al. (2017). Rapid turnover of DnaA at replication origin regions contributes to initiation control of DNA replication. *PLoS Genet.* 13:e1006561. doi: 10.1371/journal.pgen.1006561

Scholefield, G., Errington, J., and Murray, H. (2012). Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. *EMBO J.* 31, 1542–1555. doi: 10.1038/emboj.2012.6

Scholefield, G., and Murray, H. (2013). YabA and DnaD inhibit helix assembly of the DNA replication initiation protein DnaA. *Mol. Microbiol.* 90, 147–159. doi: 10.1111/mmi.12353

Seid, C. A., Smith, J. L., and Grossman, A. D. (2017). Genetic and biochemical interactions between the bacterial replication initiator DnaA and the nucleoid-associated protein Rok in *Bacillus subtilis*. *Mol. Microbiol.* 103, 798–817. doi: 10.1111/mmi.13590

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

Smits, W. K., Goranov, A. I., and Grossman, A. D. (2010). Ordered association of helicase loader proteins with the *Bacillus subtilis* origin of replication *in vivo*. *Mol. Microbiol.* 75, 452–461. doi: 10.1111/j.1365-2958.2009.06999.x

Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H., and Kuijper, E. J. (2016). *Clostridium difficile* infection. *Nat. Rev. Dis. Prim.* 2:16020. doi: 10.1038/nrdp.2016.20

Smits, W. K., Merrikh, H., Bonilla, C. Y., and Grossman, A. D. (2011). Primosomal proteins DnaD and DnaB are recruited to chromosomal regions bound by DnaA in *Bacillus subtilis*. *J. Bacteriol.* 193, 640–648. doi: 10.1128/JB.01253-1210

Speck, C., Weigel, C., and Messer, W. (1999). ATP- and ADP-dnaA protein, a molecular switch in gene regulation. *EMBO J.* 18, 6169–6176. doi: 10.1093/emboj/18.21.6169

Sutton, M. D., and Kaguni, J. M. (1997). Novel alleles of the *Escherichia coli* dnaA gene. *J. Mol. Biol.* 271, 693–703. doi: 10.1006/jmbi.1997.1209

Torti, A., Lossani, A., Savi, L., Focher, F., Wright, G. E., Brown, N. C., et al. (2011). *Clostridium difficile* DNA polymerase IIIC: basis for activity of antibacterial compounds. *Curr. Enzym. Inhib.* 7, 147–153.

Tsodikov, O. V., and Biswas, T. (2011). Structural and thermodynamic signatures of DNA recognition by Mycobacterium tuberculosis DnaA. *J. Mol. Biol.* 410, 461–476. doi: 10.1016/j.jmb.2011.05.007

van Eijk, E., Anvar, S. Y., Browne, H. P., Leung, W. Y., Frank, J., Schmitz, A. M., et al. (2015). Complete genome sequence of the *Clostridium difficile* laboratory strain 630Deltaerm reveals differences from strain 630, including translocation of the mobile element CTn5. *BMC Genomics* 16:31. doi: 10.1186/s12864-015-1252-7

van Eijk, E., Boekhoud, I. M., Kuijper, E. J., Bos-Sanders, I., Wright, G., and Smits, W. K. (2019). Genome location dictates the transcriptional response to PolC inhibition in *Clostridium difficile*. *Antimicrob. Agents Chemother.* 63:e01363-18. doi: 10.1128/AAC.01363-18

van Eijk, E., Paschalis, V., Green, M., Friggen, A. H., Larson, M. A., Spriggs, K., et al. (2016). Primase is required for helicase activity and helicase alters the specificity of primase in the enteropathogen *Clostridium difficile*. *Open Biol.* 6:160272. doi: 10.1098/rsob.160272

van Eijk, E., Wittekoek, B., Kuijper, E. J., and Smits, W. K. (2017). DNA replication proteins as potential targets for antimicrobials in drug-resistant bacterial pathogens. *J. Antimicrob. Chemother.* 72, 1275–1284. doi: 10.1093/jac/dkw548

Vellanoweth, R. L., and Rabinowitz, J. C. (1992). The influence of ribosome-binding-site elements on translational efficiency in *Bacillus subtilis* and *Escherichia coli in vivo*. *Mol. Microbiol.* 6, 1105–1114. doi: 10.1111/j.1365-2958.1992.tb01548.x

Velten, M., McGovern, S., Marsin, S., Ehrlich, S. D., Noirot, P., and Polard, P. (2003). A two-protein strategy for the functional loading of a cellular replicative DNA helicase. *Mol. Cell* 11, 1009–1020. doi: 10.1016/s1097-2765(03)00130-8

Warriner, K., Xu, C., Habash, M., Sultan, S., and Weese, S. J. (2017). Dissemination of *Clostridium difficile* in food and the environment: Significant sources of *C. difficile* community-acquired infection? *J. Appl. Microbiol.* 122, 542–553. doi: 10.1111/jam.13338

Weigel, C., Schmidt, A., Seitz, H., Tungler, D., Welzeck, M., and Messer, W. (1999). The N-terminus promotes oligomerization of the *Escherichia coli* initiator protein DnaA. *Mol. Microbiol.* 34, 53–66. doi: 10.1046/j.1365-2958.1999.01568.x

Wolanski, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwinska, J. (2014). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol* 5:735. doi: 10.3389/fmicb.2014.00735

Xu, W. C., Silverman, M. H., Yu, X. Y., Wright, G., and Brown, N. (2019). Discovery and development of DNA polymerase IIIC inhibitors to treat Gram-positive infections. *Bioorg. Med. Chem.* 27, 3209–3217. doi: 10.1016/j.bmc.2019.06.017

Zawilak, A., Durrant, M. C., Jakimowicz, P., Backert, S., and Zakrzewska-Czerwinska, J. (2003). DNA binding specificity of the replication initiator protein, DnaA from *Helicobacter pylori*. *J. Mol. Biol.* 334, 933–947.

Zawilak-Pawlik, A., Kois, A., Majka, J., Jakimowicz, D., Smulczyk-Krawczyszyn, A., Messer, W., et al. (2005). Architecture of bacterial replication initiation complexes: orisomes from four unrelated bacteria. *Biochem. J.* 389(Pt 2), 471–481. doi: 10.1042/BJ20050143

Zawilak-Pawlik, A., Nowaczyk, M., and Zakrzewska-Czerwinska, 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* 8:136. doi: 10.3390/genes8050136

Zhabinskaya, D., Madden, S., and Benham, C. J. (2015). SIST: stress-induced structural transitions in superhelical DNA. *Bioinformatics* 31, 421–422. doi: 10.1093/bioinformatics/btu657

Zhang, W., Carneiro, M. J., Turner, I. J., Allen, S., Roberts, C. J., and Soultanas, P. (2005). The *Bacillus subtilis* DnaD and DnaB proteins exhibit different DNA remodelling activities. *J. Mol. Biol.* 351, 66–75. doi: 10.1016/j.jmb.2005.05.065

Zorman, S., Seitz, H., Sclavi, B., and Strick, T. R. (2012). Topological characterization of the DnaA-oriC complex using single-molecule nanomanipuation. *Nucleic Acids Res.* 40, 7375–7383. doi: 10.1093/nar/gks371

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership