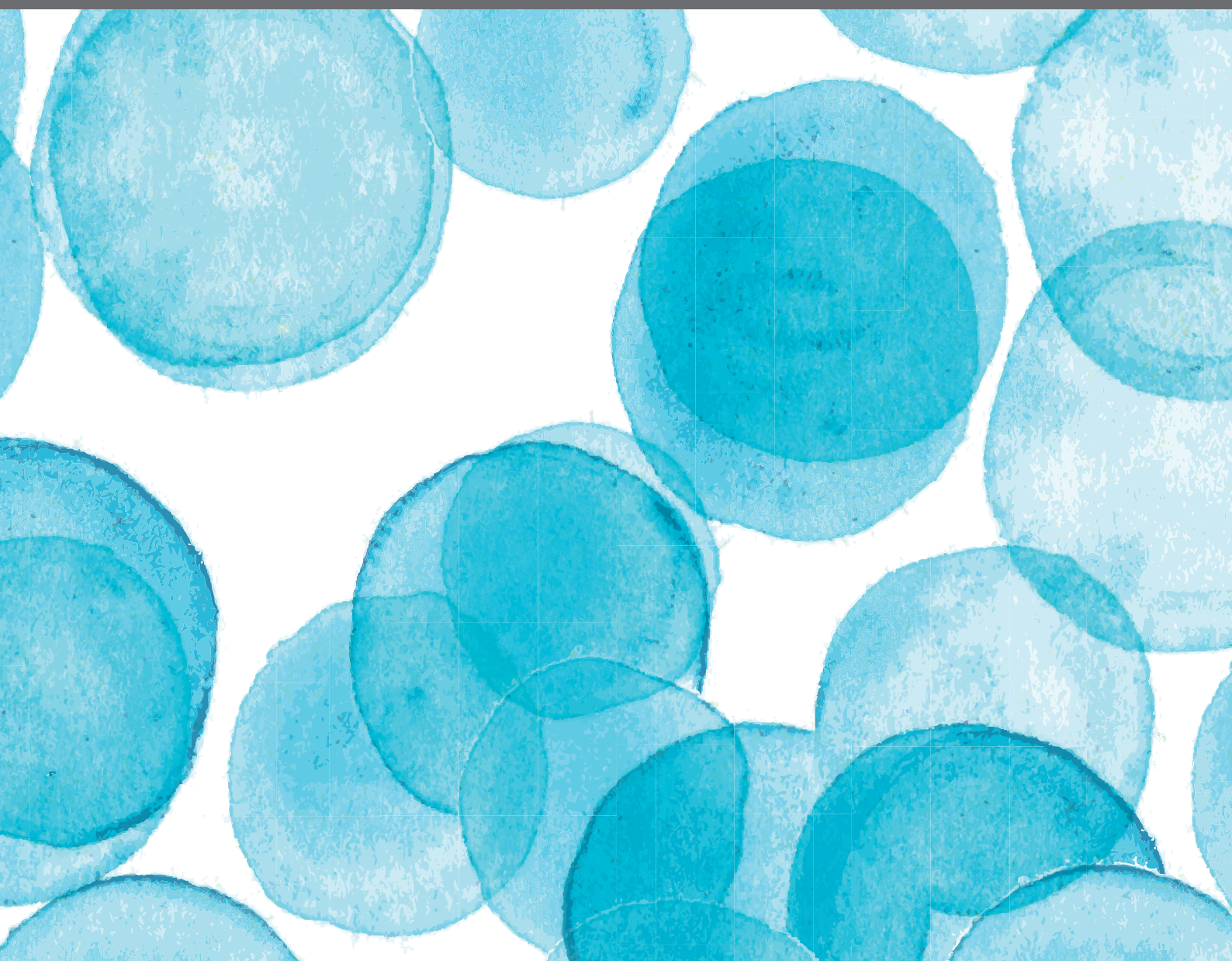
BACTERIAL TRANSCRIPTION FACTORS AND THE CELL CYCLE, 2nd Edition

EDITED BY: Morigen Morigen, Monika Glinkowska and Jianping Xie

PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-767-0
DOI 10.3389/978-2-88976-767-0

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 TRANSCRIPTION FACTORS AND THE CELL CYCLE, 2nd Edition

Topic Editors:
Morigen Morigen, Inner Mongolia University, China
Monika Glinkowska, University of Gdansk, Poland
Jianping Xie, Southwest University, China

Publisher's note: This is a 2nd edition due to an article retraction.

Citation: Morigen, M., Glinkowska, M., Xie, J., eds. (2022). Bacterial Transcription Factors and the Cell Cycle, 2nd Edition. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-767-0

Table of Contents

frontiers in Microbiology

EDITORIAL
published: 24 December 2021
doi: 10.3389/fmicb.2021.821394

Editorial: Bacterial Transcription Factors and the Cell Cycle

Morigen[1]*, *Monika Glinkowska*[2]* and *Jianping Xie*[3]*

[1] State Key Laboratory of Reproductive Regulation and Breeding of Grassland Livestock, School of Life Sciences, Inner Mongolia University, Hohhot, China, [2] Department of Bacterial Molecular Genetics, University of Gdansk, Gdańsk, Poland, [3] Institute of Modern Biopharmaceuticals, School of Life Sciences, Southwest University, Chongqing, China

Keywords: bacterial transcription factors, chromosome replication, cell division, cell growth, cell cycle

Editorial on the Research Topic

Bacterial Transcription Factors and the Cell Cycle

INTRODUCTION

The fluctuation in gene expression regulated by transcription factors are tightly associated with progression of the bacterial cell cycle. In fact, the replication initiator protein DnaA regulates DNA replication (Fuller and Kornberg, 1983; Bramhill and Kornberg, 1988) and also controls gene expression as a transcription factor (Atlung et al., 1984; Braun et al., 1985; Messer and Weigel, 1997; Wurihan et al., 2018). It is likely that DnaA may coordinate DNA replication and repair process by participating in control of the SOS response genes (Wurihan et al., 2018). As reviewed by Menikpurage et al., DnaA has been shown to regulate expression of genes involved in different cellular processes including quorum sensing, cell motility, heat shock response, DNA repair, and cell cycle regulation. In turn, the quorum sensing regulators QseB and QseC in *Escherichia coli* have been reported to modulate the expression of *dnaA* (Wu et al., 2021). Thus, dual function DnaA might temporally coordinate diverse cellular events with the onset of chromosome replication. Especially, in *Caulobacter crescentus*, DnaA couples DNA replication with the expression of the two oscillating regulators GcrA and CtrA, and the DnaA/GcrA/CtrA regulatory cascade drives forward the progression of the cell cycle (Collier et al., 2006). Other transcription factors and proteins involved in transcription control have also been shown to coordinate the cell cycle events. Namely, RpoS in *Coxiella* regulates the gene expression involved in the developmental cycle (Moormeier et al., 2019). Similarly, BolA in most Gram-negative bacteria turns off motility and turns on biofilm development as a transcription factor (Dressaire et al., 2015), whereas CtrA regulates cell division and outer membrane composition of the pathogen *Brucella abortus* (Francis et al., 2017). The Research Topic is a collection of articles focusing on the *Bacterial Transcription Factors and the Cell Cycle*, particularly: initiation of chromosome replication, cell division, and cell growth.

THE BACTERIAL TRANSCRIPTION FACTORS

Mfd is a transcription-coupled DNA repair factor, which couples a stalled RNA polymerase with the nucleotide excision repair pathway to preferentially repair lesions in the template strand of actively transcribing genes (Hanawalt and Spivak, 2008). Interestingly, Martin et al. find that Mfd modulates global transcription in stationary-phase and during disulfide stress in *Bacillus subtilis* cells. Loss of Mfd leads to changes in expression of genes involved in translation, endospore formation, motility, and pyrimidine metabolism. The profound effects of Mfd on the modulation of the transcriptome and on bacterial physiology, particularly in cells experiencing nutritional and oxidative stress suggest that Mfd is multifunctional protein coordinating transcription and DNA repair.

OPEN ACCESS

Edited and reviewed by:
Marc Strous,
University of Calgary, Canada

***Correspondence:**
Morigen
morigenm@hotmail.com
Monika Glinkowska
monika.glinkowska@ug.edu.pl
Jianping Xie
georgex@swu.edu.cn

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *24 November 2021*
Accepted: *06 December 2021*
Published: *24 December 2021*

Citation:
Morigen, Glinkowska M and Xie J (2021) Editorial: Bacterial Transcription Factors and the Cell Cycle. Front. Microbiol. 12:821394. doi: 10.3389/fmicb.2021.821394

Regulation of gene expression by various transcription factors enables bacteria to adapt to various environmental conditions and utilize different nutrients. In this context, transcription factors and their functions are investigated in various bacterial species in the collection. Firstly, Tsevelkhoroloo et al. elucidates DagR-mediated regulation of expression of three agarase genes (*dagA*, *dagB*, and *dagC* in *Streptomyces coelicolor*). They show that DagR operates by binding to a palindromic sequence present in the upstream regions of the three genes. This is the first report on the regulation of expression of genes involved in agar metabolism in *S. coelicolor* A3(2). In *Mycobacterium smegmatis*, Jeong-Il Oh lab demonstrates the distinct roles of additional RsbW homologs: RsbW1, RsbW2, and RsbW3 in the SigF Partner Switching System (PSS). They show that RsbW1 and RsbW2 serve as the anti-sigma factor of SigF and the protein kinase phosphorylating RsfB, respectively, while RsbW3 functions as an anti-SigF antagonist through its protein interaction with RsbW1. Moreover, they also demonstrate that Crp1 plays a predominant role in induction of the cydAB operon in *M. smegmatis* exposed to the respiratory electron transport chain (ETC)-inhibitory conditions. In a similar context, Wan et al. identify that OxyR as a master transcription regulator mediates cellular responses to hydrogen peroxide in *Yersinia pseudotuberculosis.* Using genomic and transcriptomic analyses, they show that OxyR activates transcription of diverse genes encoding for catalases, peroxidases, and thiol reductases. Those findings provide not only novel insights into the structural and functional diversity of bacterial hydrogen peroxide-scavenging systems but also a basic understanding of how *Y. pseudotuberculosis* copes with oxidative stress. Finally, in *Bacillus anthracis*, the NprR is a protein that exhibits moonlighting functions as either a phosphatase or a neutral protease regulator that belongs to the RNPP family. Wang et al. show that the extracellular protease activity of an *nprR* deletion mutant is significantly decreased. Further, the data suggests that NprR is a transcription regulator that controls genes involved in extracellular proteolysis and also the oxidative stress response.

THE BACTERIAL CELL CYCLE

The fast-growing *E. coli* cell cycle consists of a chromosome replication and a cell division cycle (Boye and Nordström, 2003). As reviewed by Grimwade and Leonard, chromosome replication is a critical event in the cell cycle since all daughter cells must inherit a complete genome. In bacteria, chromosome replication starts when nucleoprotein complexes (orisomes) unwind replication origin DNA and build new replication forks. Functional orisomes comprising DnaA bound to a set of high and low affinity recognition sites in *oriC* must be assembled once, and only once, per cell cycle. During the orisome assembly in *E. coli*, three known transcriptional factors, SeqA, Fis (factor for inversion stimulation), and IHF (integration host factor), interact directly with the *oriC* site to regulate orisome assembly and replication initiation timing. Interestingly, Kasho et al. determine the *E. coli* genome-wide, cell cycle-dependent binding of IHF with base-pair resolution using GeF-seq (genome footprinting with high-throughput sequencing). They show that *oriC* is a main site for stable IHF binding before replication initiation whereas all other loci have a reduced IHF binding. Thus, they believe that IHF rapidly dissociates from *oriC* after replication initiation, and IHF binding to other sites is sustained or stimulated. The findings provide mechanistic insight into cell cycle-dependent IHF binding and dissociation in the genome. In another work concerning nucleoid associated proteins, Pal et al. characterize the YbaB/EbfC family DNA binding protein *of C. crescentus*, which is conserved in bacteria. Unlike *Borrelia burgdorferi* EbfC protein, YbaBCc protein is a non-specific DNA-binding protein that protects DNA from enzymatic degradation. The findings reveal that YbaBCc is a small histone-like protein and may play a role in bacterial chromosome structuring and gene regulation.

Cell division is tightly coordinated with chromosome replication, and cell growth. In *E. coli*, many cell division proteins interact with FtsZ and modulate Z-ring assembly, cell wall insertion and peptidoglycan remodeling. DiBiasio et al. show that ZapE directly interacts with FtsZ and phospholipids *in vitro*. ZapE is an ATPase that accumulates during late constriction, and induces bundling of GTP-induced FtsZ polymers in the presence of ATP. ZapE is proposed to ensure cell proliferation during and after stress, especially when cells are also deleted for *minC*. This suggests that ZapE is a stress-regulated cell division protein, promoting growth and division after exposure to environmental stress. In another study, Patil et al. show that excess DnaA(L366K) or deletion of *fis* allows for growth of cells that otherwise would undergo cell cycle arrest as a result of accumulation of lipoprotein Lpp(C21G). The DnaA(L366K)-mediated restoration of growth occurs by reduced expression of Lpp(C21G) via a σE-dependent small-regulatory RNA (sRNA), MicL-S. On the other hand, deletion of *fis* rescues Lpp(C21G)-evoked growth arrest in cells lacking physiological levels of PG and cardiolipin (CL), independently of MicL-S. The results suggest a close link between the physiological state of the bacterial cell membrane and DnaA- and Fis-dependent growth.

Quantitative models of the cell-cycle and cell-size control in *E. coli* and *B. subtilis* have been proposed over the last decade to explain single-cell experimental data generated with high-throughput methods. Le Treut et al. examine predictive power of these models against experimental data. They find that Independent Double Adder (IDA) model is more consistent with the data than the other models and conclude that the IDA model is so far the most consistent one with both the data and the biology of bacterial cell-cycle and cell-size control.

PERSPECTIVES

The location of genes relative to *ori* and *ter* of the bacterial circular chromosome has profound effects on transcription regulatory networks and physiological processes (Sobetzko et al., 2012). The CtrA phosphorelay is conserved in most alphaproteobacteria as a gene regulatory system. Tomasch et al. show that the CtrA phosphorelay genes are conserved closer to *ori* in the Caulobacterales, while the genes are highly conserved

closer to *ter* in the Rhodobacterales. However, the location of the phosphorelay genes is the least conserved in the Rhodospirillales and Sphingomonadales. The findings suggest that selection pressure results in differential positioning of CtrA phosphorelay and associated genes in alphaproteobacteria. Thus, future deeper investigations on correlation of chromosomal locations and function of a broad spectrum of bacterial transcription factors are worth pursuing both theoretically and experimentally.

AUTHOR CONTRIBUTIONS

M, MG, and JX participated in the first manuscript draft. M edited the final version and all authors revised it. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was supported by a grants from the National Natural Science Foundation of China NSFC (Grant no. 32060016 to M).

ACKNOWLEDGMENTS

We deeply thank all the authors and reviewers who have participated in this Research Topic.

REFERENCES

Atlung, T., Clausen, E., and Hansen, F. G. (1984). Autorepression of the *dnaA* gene of *Escherichia coli*. *Adv. Exp. Med. Biol.* 179, 199–207. doi: 10.1007/978-1-4684-8730-5_20

Boye, E., and Nordström, K. (2003). Coupling the cell cycle to cell growth. *EMBO Rep.* 4, 757–760. doi: 10.1038/sj.embor.embor895

Bramhill, D., and Kornberg, A. (1988). A model for initiation at origins of DNA replication. *Cell* 54, 915–918. doi: 10.1016/0092-8674(88)90102-X

Braun, R. E., O'day, K., and Wright, A. (1985). Autoregulation of the DNA replication gene *dnaA* in *E. coli* K-12. *Cell* 40, 159–169. doi: 10.1016/0092-8674(85)90319-8

Collier, J., Murray, S. R., and Shapiro, L. (2006). DnaA couples DNA replication and the expression of two cell cycle master regulators. *EMBO J.* 25, 346–356. doi: 10.1038/sj.emboj.7600927

Dressaire, C., Moreira, R. N., Barahona, S., Alves De Matos, A. P., and Arraiano, C. M. (2015). BolA is a transcriptional switch that turns off motility and turns on biofilm development. *mBio* 6, e02352–e02314. doi: 10.1128/mBio.02352-14

Francis, N., Poncin, K., Fioravanti, A., Vassen, V., Willemart, K., Ong, T. A., et al. (2017). CtrA controls cell division and outer membrane composition of the pathogen *Brucella abortus*. *Mol. Microbiol.* 103, 780–797. doi: 10.1111/mmi.13589

Fuller, R. S., and Kornberg, A. (1983). Purified DnaA protein in initiation of replication at the *Escherichia coli* chromosomal origin of replication. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5817–5821. doi: 10.1073/pnas.80.19.5817

Hanawalt, P. C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* 9, 958–970. doi: 10.1038/nrm2549

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

Moormeier, D. E., Sandoz, K. M., Beare, P. A., Sturdevant, D. E., Nair, V., Cockrell, D. C., et al. (2019). *Coxiella burnetii* RpoS regulates genes involved in morphological differentiation and intracellular growth. *J. Bacteriol.* 201:e00009-19. doi: 10.1128/JB.00009-19

Sobetzko, P., Travers, A., and Muskhelishvili, G. (2012). Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc. Natl. Acad. Sci. U.S.A.* 109, E42–50. doi: 10.1073/pnas.1108229109

Wu, D., Baigalmaa, L., Yao, Y., Li, G., Su, M., Fan, L., and Morigen (2021). The *Escherichia coli* QseB/QseC signaling is required for correct timing of replication initiation and cell motility. *Gene* 773:145374. doi: 10.1016/j.gene.2020.145374

Wurihan, Gezi, Brambilla, E., Wang, S., and Morigen (2018). DnaA and LexA proteins regulate transcription of the *uvrB* gene in *Escherichia coli*: the role of DnaA in the control of the SOS regulon. *Front. Microbiol.* 09:1212. doi: 10.3389/fmicb.2018.01212

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 11 November 2020
doi: 10.3389/fmicb.2020.588487

The Partner Switching System of the SigF Sigma Factor in *Mycobacterium smegmatis* and Induction of the SigF Regulon Under Respiration-Inhibitory Conditions

*Yuna Oh†, Su-Yeon Song†, Hye-Jun Kim, Gil Han, Jihwan Hwang, Ho-Young Kang and Jeong-Il Oh**

Department of Integrated Biological Science, Pusan National University, Busan, South Korea

OPEN ACCESS

Edited by:
Morigen Morigen, Inner Mongolia University, China

Reviewed by:
Vinay Kumar Nandicoori, National Institute of Immunology (NII), India
Kaixia Mi, Chinese Academy of Sciences, China

***Correspondence:**
Jeong-Il Oh
joh@pusan.ac.kr

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *29 July 2020*
Accepted: *22 October 2020*
Published: *11 November 2020*

Citation:
Oh Y, Song S-Y, Kim H-J, Han G, Hwang J, Kang H-Y and Oh J-I (2020) The Partner Switching System of the SigF Sigma Factor in Mycobacterium smegmatis and Induction of the SigF Regulon Under Respiration-Inhibitory Conditions. Front. Microbiol. 11:588487. doi: 10.3389/fmicb.2020.588487

The partner switching system (PSS) of the SigF regulatory pathway in *Mycobacterium smegmatis* has been previously demonstrated to include the anti-sigma factor RsbW (MSMEG_1803) and two anti-sigma factor antagonists RsfA and RsfB. In this study, we further characterized two additional RsbW homologs and revealed the distinct roles of three RsbW homologs [RsbW1 (MSMEG_1803), RsbW2 (MSMEG_6129), and RsbW3 (MSMEG_1787)] in the SigF PSS. RsbW1 and RsbW2 serve as the anti-sigma factor of SigF and the protein kinase phosphorylating RsfB, respectively, while RsbW3 functions as an anti-SigF antagonist through its protein interaction with RsbW1. Using relevant mutant strains, RsfB was demonstrated to be the major anti-SigF antagonist in *M. smegmatis*. The phosphorylation state of Ser-63 was shown to determine the functionality of RsfB as an anti-SigF antagonist. RsbW2 was demonstrated to be the only protein kinase that phosphorylates RsfB in *M. smegmatis*. Phosphorylation of Ser-63 inactivates RsfB to render it unable to interact with RsbW1. Our comparative RNA sequencing analysis of the wild-type strain of *M. smegmatis* and its isogenic Δaa_3 mutant strain lacking the aa_3 cytochrome *c* oxidase of the respiratory electron transport chain revealed that expression of the SigF regulon is strongly induced under respiration-inhibitory conditions in an RsfB-dependent way.

Keywords: aa_3 cytochrome *c* oxidase, anti-sigma factor, anti-anti-sigma factor, electron transport chain, *Mycobacterium*, partner switching system, protein kinase, SigF

INTRODUCTION

Sigma factors reversibly associate with the core RNA polymerase and function as specific factors that direct transcription of a specific subset of genes. 28 sigma factor genes were found to occur in *Mycobacterium smegmatis* in contrast to 13 genes in *Mycobacterium tuberculosis* (Cole et al., 1998; Manganelli et al., 2004; Waagmeester et al., 2005; Rodrigue et al., 2006). SigF belongs to group III sigma factors and is dispensable for growth of *M. tuberculosis* and *M. smegmatis* (Chen et al., 2000; Williams et al., 2007; Provvedi et al., 2008; Singh et al., 2015). SigF is phylogenetically and

functionally in close relation to the well-studied stress sigma factor SigB in *Bacillus* species (DeMaio et al., 1996, 1997; Hecker and Volker, 2001; Hecker et al., 2007; Bouillet et al., 2018).

The *sigF* gene is widely conserved in mycobacterial species (Singh and Singh, 2008; Sachdeva et al., 2010). In *M. tuberculosis*, SigF (Rv3286c) was shown to be involved in virulence, biofilm formation, and diverse stress responses (Chen et al., 2000; Geiman et al., 2004; Karls et al., 2006; Williams et al., 2007; Hartkoorn et al., 2012; Manganelli, 2014). SigF (MSMEG_1804) in *M. smegmatis* was suggested to play roles in adaptation to stationary phase and conditions of heat and oxidative stress (Gebhard et al., 2008; Humpel et al., 2010; Singh et al., 2015). Overexpression or deletion of *sigF* was reported to alter cell wall architectures in *M. tuberculosis*, *Mycobacterium bovis*, and *M. smegmatis* (Forrellad et al., 2013; Singh et al., 2015; Dutta et al., 2019). The inactivation of *sigF* in *M. smegmatis* was shown to result in the loss of carotenoid (isorenieratene) pigmentation, accompanying increased susceptibility to hydrogen peroxide (Provvedi et al., 2008; Humpel et al., 2010; Singh et al., 2015). The consensus sequence (GGWWT-N_{16-17}-GGGTAY) was suggested for the mycobacterial SigF-recognizing promoters (Gebhard et al., 2008; Provvedi et al., 2008; Humpel et al., 2010). The *sigF* genes of *M. tuberculosis* and *M. smegmatis* were demonstrated to be cotranscribed with their cognate anti-sigma factor genes *usfX* (*Rv3287c*) and *rsbW* (*MSMEG_1803*), respectively (DeMaio et al., 1997; Gebhard et al., 2008). Transcription of the *usfX-sigF* operon in *M. tuberculosis* is driven from a SigF-dependent promoter, whereas the *rsbW-sigF* operon in *M. smegmatis* is transcribed from two promoters, a SigF-independent promoter located immediately upstream of *rsbW* and a SigF-dependent promoter upstream of the *chaB* (*MSMEG_1802*) gene that is located 103 bp upstream of *rsbW* (Gebhard et al., 2008).

M. tuberculosis sigF was found to be strongly induced within cultured human macrophages, during stationary phase of growth, and upon exposure to cold shock, nutrient depletion, oxidative stress, and several antibiotics (rifampicin, ethambutol, streptomycin, and cycloserine), as well as in persister cells (DeMaio et al., 1996; Graham and Clark-Curtiss, 1999; Michele et al., 1999; Mariani et al., 2000; Betts et al., 2002; Keren et al., 2011; Forrellad et al., 2013), while *M. smegmatis sigF* was shown to be expressed at similar levels throughout its growth phase and only marginally increased under SigF-activating conditions (Gebhard et al., 2008; Singh and Singh, 2008).

The functionality of SigF is regulated by the so-called partner switching system (PSS) including its cognate anti-sigma factor (anti-SigF) and anti-anti-sigma factors (anti-SigF antagonists) (**Figure 1**; DeMaio et al., 1997; Singh et al., 2015; Bouillet et al., 2018). Under non-stress (SigF-non-activating) conditions, SigF is held in an inactive state in complex with the anti-SigF (RsbW or UsfX). Under stress (SigF-activating) conditions, the release of SigF from its anti-SigF is accomplished by two anti-SigF antagonists, RsfA (Rv1365c in *M. tuberculosis* and MSMEG_1786 in *M. smegmatis*) and RsfB (Rv3687c in *M. tuberculosis* and MSMEG_6127 in *M. smegmatis*), which sequester the anti-SigF (Beaucher et al., 2002; Parida et al., 2005; Singh et al., 2015). RsfA is inactivated when a disulfide bond is formed between its two redox-responsive cysteine residues, while RsfB was suggested to be inactivated by phosphorylation (Beaucher et al., 2002; Manganelli et al., 2004). MSMEG_6129 was identified to be a protein kinase that phosphorylates RsfB in *M. smegmatis* (Bowman and Ghosh, 2014), while the kinase that is responsible for RsfB phosphorylation remains unknown in *M. tuberculosis*. Since an *MSMEG_6129* mutant strain of *M. smegmatis* paradoxically displayed decreased expression of the SigF regulon relative to the wild-type (WT) strain (Bowman and Ghosh, 2014), it was uncertain whether MSMEG_6129 is the kinase that inactivates RsfB by phosphorylation. Dephosphorylation of RsfB homologs in *Bacillus* species was demonstrated to be catalyzed by the PP2C family of phosphatases (Voelker et al., 1996; Vijay et al., 2000; Chen et al., 2003). However, no study has been published regarding which gene product is responsible for dephosphorylation of phosphorylated RsfB in mycobacteria. The *Rv1364c* gene in *M. tuberculosis* was found to encode a multi-domain protein consisting of the sensor, PP2C phosphatase, GHKL (gyrase, Hsp90, histidine kinase, MutL) kinase, and anti-sigma antagonist domains (Sachdeva et al., 2008; Greenstein et al., 2009; Misra et al., 2019). Rv1364c was shown to interact with SigF *in vitro*, suggesting the possibility that it might serve as an anti-SigF along with the major anti-SigF UsfX (Misra et al., 2019). Although Rv1364c was demonstrated to possess both kinase and phosphatase activities that autophosphorylate and autodephosphorylate its anti-sigma antagonist domain (Greenstein et al., 2009), the question remains unanswered regarding whether it can phosphorylate and dephosphorylate RsfB to modulate the anti-SigF antagonist activity of RsfB.

Using relevant mutant strains and protein interaction analyses, we here reveal the distinct roles of three RsbW homologs and two anti-SigF antagonists (RsfA and RsfB) in the SigF PSS of *M. smegmatis*. This study provides several lines of evidence showing that MSMEG_6129 is the only kinase in *M. smegmatis* that phosphorylates RsfB on Ser-63 to inactivate the functionality of RsfB as an anti-SigF antagonist. This study also presents the novel finding that expression of the SigF regulon in *M. smegmatis* is strongly induced under respiration-inhibitory conditions in an RsfB-dependent way.

MATERIALS AND METHODS

Bacterial and Yeast Strains, Plasmids, and Culture Conditions

The bacterial and yeast strains and plasmids used in this study are listed in **Table S1** in the supplementary material. *Escherichia coli* strains were cultivated in Luria-Bertani (LB) medium on a gyratory shaker (200 rpm) at 37°C. *M. smegmatis* strains were grown aerobically in Middlebrook 7H9 medium (Difco, Sparks, MD, United States) supplemented with 0.2% (w/v) glucose (7H9-glucose) and 0.02% (v/v) Tween 80 as an anti-clumping agent on a gyratory shaker at 37°C. Ampicillin (100 or 200 μg/ml for *E. coli*), kanamycin (50 μg/ml for *E. coli* and 15 or 30 μg/ml for *M. smegmatis*), chloramphenicol (34 μg/ml for *E. coli*) and hygromycin (200 μg/ml for *E. coli* and 50 μg/ml for *M. smegmatis*) were added to the growth medium when required.

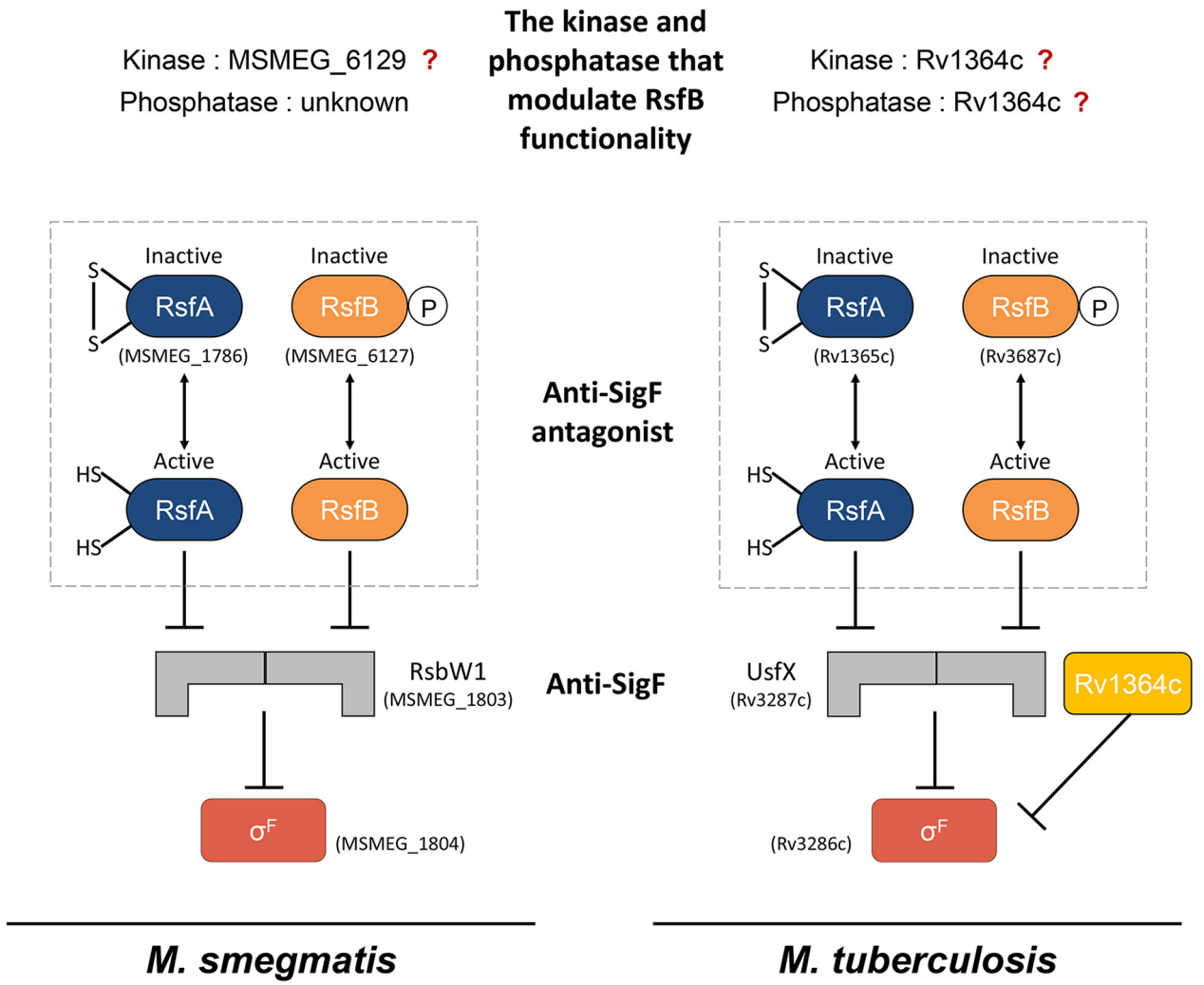


FIGURE 1 | The schematic illustration outlining the current knowledge of the SigF PSS in *M. smegmatis* and *M. tuberculosis*. The "?" mark indicates the ambiguity of the roles of the corresponding proteins in the SigF PSS.

For treatment of *M. smegmatis* cultures with KCN, the cultures were grown to an optical density at 600 nm (OD_{600}) of 0.45–0.5 and further incubated for 15 min following the addition of KCN to the cultures to a final concentration of 0.5 mM.

Saccharomyces cerevisiae strains were cultivated in YPD medium (Difco) or synthetic defined dropout (SD) medium (Clontech, Palo Alto, CA, United States) on a gyratory shaker at 30°C.

DNA Manipulation and Transformation

Standard protocols and manufacturers' instructions were followed for recombinant DNA manipulations (Green and Sambrook, 2012). Transformation of *M. smegmatis* and *S. cerevisiae* with plasmids was conducted by electroporation and the lithium acetate (LiAc)-mediated method, respectively, as previously described (Snapper et al., 1990; Guthrie and Fink, 1991).

Site-Directed Mutagenesis

To introduce point mutations into the *rsfB* genes, PCR-based mutagenesis was performed using the Quick Change site-directed mutagenesis procedure (Stratagene, La Jolla, CA, United States). Synthetic oligonucleotides containing a mutated codon in the middle of their sequences were used to mutagenize the original codons. The primers used for mutagenesis are listed in Table S2. Mutations were verified by DNA sequencing.

β-Galactosidase Assay and Determination of the Protein Concentration

β-Galactosidase activity in *M. smegmatis* was assayed spectrophotometrically as described elsewhere (Oh and Kaplan, 1999). Protein concentration was determined by using a Bio-Rad protein assay kit (Bio-Rad, Hercules, CA, United States) with bovine serum albumin as a standard protein.

Reverse Transcription-PCR and Quantitative Real-Time PCR

RNA isolation from *M. smegmatis* strains, preparation of cDNA, reverse transcription PCR (RT-PCR), and quantitative real-time PCR (qRT-PCR) were conducted as described previously (Kim et al., 2010). The primers used for cDNA synthesis, RT-PCR, and qRT-PCR are listed in Table S2.

Protein Purification

The C-terminally His_6-tagged WT and mutant forms of RsfB were expressed in *E. coli* BL21 (DE3) strains harboring the pT7-7 derivative plasmids (pT7-7rsfB, pT7-7rsfBS63A, and pT7-7rsfBS63E). The strains harboring the pT7-7 derivatives were cultivated aerobically at 37°C in LB medium containing 100 μg/ml ampicillin to an OD_{600} of 0.4–0.6. Expression of the *rsfB* gene was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, and then cells were further grown for 4 h at 30°C. For purification of RsfB from *M. smegmatis*, the *M. smegmatis* strains containing pMHRsfB or pMHRsfBS63E were grown aerobically to an OD_{600} of 0.5–0.8 at 37°C in 7H9-glucose medium supplemented with 15 μg/ml kanamycin. Expression of the *rsfB* gene was induced by the addition of acetamide to a final concentration of 0.2% (w/v), and then cells were further grown for 7 h at 37°C. After 400 ml of *E. coli* or *M. smegmatis* cultures were harvested, cells were resuspended in 10 ml of buffer A [(20 mM Tris–HCl pH 8.0), 100 mM NaCl] containing 10 U/ml DNase I and 10 mM $MgCl_2$. The resuspended cells were disrupted twice for *E. coli* or five times for *M. smegmatis* using a French pressure cell, and cell-free crude extracts were obtained by centrifugation twice at 14,000 × *g* for 15 min. 500 μl of 50% (v/v) slurry (bed volume: 250 μl) of Ni-Sepharose high performance resin (GE Healthcare, Piscataway, NJ, United States) was packed into a column. After equilibration of the resin with 10 bed volumes of buffer A, cell-free crude extracts were loaded into the column. The resin was washed with 40 bed volumes of buffer A containing 10 mM imidazole, 20 bed volumes of buffer A containing 30 mM imidazole, and then His_6-tagged RsfB was eluted with 10 bed volumes of buffer A containing 100 mM imidazole. The eluted His_6-tagged RsfB was diluted with buffer A to 10 mM imidazole and subjected to affinity chromatography again. The resin was washed with 20 bed volumes of buffer A containing 30 mM imidazole, and then His_6-tagged RsfB was finally eluted with 10 bed volumes of buffer A containing 100 mM imidazole. Imidazole and NaCl were removed from purified RsfB by means of a PD-10 desalting column (GE Healthcare) equilibrated with 20 mM Tris–HCl (pH 8.0).

Purification of RsbW1 was conducted using *E. coli* BL21 (DE3) strains carrying pT7-7rsbW1. Cell-free crude extracts were loaded into the column packed with Ni-Sepharose resin. The resin was washed with 40 bed volumes of buffer A containing 10 mM imidazole, 20 bed volumes of buffer A containing 30 mM imidazole, and then His_6-tagged RsfB was eluted with 10 bed volumes of buffer A containing 100 mM imidazole. RsbW2 and RsbW3 were purified from the *E. coli* BL21(DE3) strain carrying pT7-7rsbW2 and the *E. coli* Rosetta-gami 2 (DE3) strain with pT7-7rsbW3, respectively, in the same way as that of RsbW1 except for the modified wash and elution steps (40 bed volumes of buffer A containing 5 mM imidazole and 30 bed volumes of buffer A containing 50 mM imidazole for the wash step; 10 bed volumes of buffer A containing 250 mM imidazole for the elution step). SigF was purified from the *E. coli* BL21(DE3) strain carrying pT7-7sigF in the same way as RsbW2 except for the modified wash step (40 bed volumes of buffer A containing 10 mM imidazole and then 40 bed volumes of buffer A containing 25 mM imidazole).

Western Blotting Analysis

To detect expressed SigF, RsbW1, and RsfB in cells, Western blotting analyses using rabbit polyclonal antibodies against the corresponding proteins were performed as described previously (Mouncey and Kaplan, 1998). For detection of His_6-tagged proteins a mouse monoclonal antibody against His-3 (Santa Cruz Biotechnology, Santa Cruz, CA, United States; sc8036) was employed. The rabbit polyclonal antibodies and His-3 monoclonal antibody were used at a 1:20,000 and 1:2,000 dilution, respectively. To detect GroEL, a mouse monoclonal antibody against HSP65 (Santa Cruz Biotechnology; sc58170) was used at a 1:2,000 dilution. Alkaline phosphatase-conjugated anti-rabbit IgG produced in goat (Sigma, St. Louis, CA, United States; A0545) or alkaline phosphatase-conjugated anti-mouse IgG produced in rabbit (Sigma; A4312) was used at a 1:10,000 dilution for the detection of the primary antibodies.

Analysis of *in vitro* Protein–Protein Interactions Using Non-denaturing PAGE

The mixture of two purified proteins in 20 mM Tris–HCl (pH 8.0) solution containing 20 mM β-mercaptoethanol was mixed with the same volume of 2× Binding buffer [40 mM Tris–HCl (pH 8.0), 0.01 mM EDTA (pH 8.0), 10 mM $MgCl_2$, 20% (v/v) glycerol] and incubated for 25 min at room temperature. After the addition of 10× sample buffer [50 mM Tris–HCl (pH 6.8), 40% (w/v) sucrose, 0.05% (w/v) bromophenol blue], the mixtures were subjected to non-denaturing PAGE [7.5% (w/v) acrylamide] using electrophoresis buffer [2.5 mM Tris–HCl (pH 8.3), 19.2 mM glycine], which was initially run at 80 V for 1 h and subsequently at 100 V for 4 h. Non-denaturing PAGE was conducted at 4°C.

Analysis of *in vivo* Protein–Protein Interactions Using Copurification Assay

To examine protein interactions of RsbW1 with RsbW2 and RsbW3 in *M. smegmatis*, copurification assay using Ni-Sepharose resin was performed. The C-terminally His_6-tagged RsbW2 and RsbW3 were expressed in the WT strains of *M. smegmatis* harboring pMHRsbW2 and pMHRsbW3, respectively. The strains were grown aerobically to an OD_{600} of 0.45–0.5 at 37°C in 7H9-glucose medium supplemented with 15 μg/ml kanamycin and 0.1% (w/v) acetamide. After 200 ml of *M. smegmatis* cultures were harvested, cells were resuspended in 8 ml of buffer A containing 10 U/ml DNase I and 10 mM $MgCl_2$. The resuspended cells were disrupted five times using a French pressure cell, and cell-free crude extracts were obtained by centrifugation twice at

14,000 × *g* for 15 min. 500 μl of the 50% (v/v) slurry (bed volume 250 μl) of Ni-Sepharose resin was packed into a column. After equilibration of the resin with 10 bed volumes of buffer A, cell-free crude extracts were loaded into the column. The resin was washed with 125 bed volumes of buffer A containing 10 mM imidazole and then His_6-tagged RsbW2 and RsbW3 were eluted with 10 bed volumes of buffer A containing 250 mM imidazole. RsbW2 and RsbW3 in the eluents were detected by Western blotting analysis with a His-3 monoclonal antibody. The presence of RsbW1 in the eluents was determined using Western blotting analysis with RsbW1 polyclonal antibodies.

Analysis of *in vivo* Protein–Protein Interactions Using Yeast Two-Hybrid Assay

S. cerevisiae AH109 strains cotransformed with both pGADT7linker and pGBKT7 derivatives were grown in SD medium (Clontech, Palo Alto, CA, United States) lacking leucine and tryptophan (SD/-Leu/-Trp). The overnight cultures were diluted with distilled water to an OD_{600} of 0.6 and spotted onto both solid SD/-Leu/-Trp plates and histidine-deficient SD/-Leu/-Trp/-His plates containing various concentrations of 3-amino-1,2,4-triazole (3-AT) for spotting assays. The plates were incubated at 30°C for 3–5 days.

RNA Sequencing and Gene Expression Profiling

Three biological replicate cultures of the WT and Δaa_3 strains were grown aerobically to an OD_{600} of 0.45–0.5. Total RNA of each culture was isolated as described previously (Kim et al., 2010). rRNA was removed from the isolated total RNA using a Ribo-Zero rRNA Removal Kit (Bacteria) (Illumina, San Diego, CA, United States). The RNA sequencing libraries were created using a TruSeq RNA Sample Prep Kit v2 (Illumina) with the standard low-throughput protocol. Sequencing of the six libraries was conducted on an Illumina HiSeq 4000 platform at Macrogen Inc. (Seoul, South Korea) using the HiSeq 3000–4000 sequencing protocol and TruSeq 3000–4000 SBS Kit v3 reagent (Illumina). Paired-end reads (101 bp) were then mapped to the reference genome sequence of *M. smegmatis* mc^2155 (GCF_000015005.1_ASM1500v1) with the program Bowtie 1.1.2 using default settings. Summarized statistics of RNA sequencing alignment are listed in Table S3. The differentially expressed genes (DEGs) were subsequently identified pair-wise by the edgeR package in R language (Robinson et al., 2010). In this analysis, the genes with *P*-value < 0.05 and $|\log_2$ fold change of gene expression (FC)$| > 2$ were regarded as DEGs. The data described in this study have been deposited in NCBI's Gene Expression Omnibus and are accessible through the GEO Series accession number GSE155251.

In vitro Kinase Assay

Purified RsfB was mixed with purified RsbW1, RsbW2, or RsbW3 in 30 μl of reaction buffer [20 mM Tris–Cl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, and 10 mM $MnCl_2$]. The reactions were started by adding 100 μM ATP and incubated for 30 min at 30°C. The reactions were terminated by adding 7.5 μl of gel-loading buffer [250 mM Tris–Cl (pH 6.8), 50% (w/v) glycerol, 500 mM dithiothreitol (DTT), 10% (w/v) SDS, 5% (v/v) β-mercaptoethanol, and 0.5% (w/v) bromophenol blue]. Proteins were resolved by Phos-tag SDS-PAGE prepared as described elsewhere (Barbieri and Stock, 2008). The duplicated reaction mixtures were subjected to normal SDS-PAGE. The gels were stained with Coomassie brilliant blue (CBB).

RESULTS

Induction of the SigF Regulon Under Respiration-Inhibitory Conditions and the Genetic Organization of the Genes Involved in the SigF PSS

Comparative RNA sequencing analysis of the WT strain of *M. smegmatis* and its isogenic Δaa_3 mutant strain with a deletion in *ctaC* encoding subunit III of the aa_3 cytochrome *c* oxidase led us to identify 103 DEGs whose expression is upregulated in the Δaa_3 mutant by more than four-fold with a *P*-value less than 0.05 relative to the WT strain. As shown in **Figure 2A** and Table S4, 61 genes of the 103 DEGs were found to overlap with the genes belonging to the known SigF regulon (Singh et al., 2015), suggesting that the genes of the SigF regulon are strongly upregulated in *M. smegmatis*, when the major terminal oxidase of the electron transport chain (ETC) is inactivated. Among the 61 identified genes, we selected two genes (*MSMEG_1777* and *MSMEG_1782*) with the large FC and RPKM (reads per kilo base pair per million mapped reads) values in the Δaa_3 mutant, and examined the expression levels of the genes in the WT (control) and Δaa_3 mutant strains grown aerobically, as well as in the WT strain treated with KCN, an inhibitor of the aa_3 cytochrome *c* oxidase (**Figure 2B**). Consistent with the RNA sequencing result, expression of *MSMEG_1777* and *MSMEG_1782* was significantly increased in the Δaa_3 mutant and the WT strain treated with KCN as compared to that in the control WT strain. We also included the Δ*f1f2f3* mutant strain of *M. smegmatis* with deletions in three *furA* paralogous genes in this experiment, since it had been reported that the genes of the SigF regulon are strongly downregulated in the Δ*f1f2f3* mutant relative to the WT strain (Lee et al., 2018). As expected, expression of *MSMEG_1777* and *MSMEG_1782* was significantly decreased in the mutant, confirming the *MSMEG_1777* and *MSMEG_1782* genes belong to the SigF regulon. Based on this result, we hereafter used the *MSMEG_1777* gene as a marker gene of the SigF regulon to determine the functionality of SigF.

As a first step to understand the mechanism underlying the strong induction of the SigF regulon under respiration-inhibitory conditions, we decided to investigate the SigF PSS in detail. The *rsbW-sigF* (*MSMEG_1803-MSMEG_1804*) operon has been previously identified, and the role of RsbW (MSMEG_1803) as an anti-SigF in *M. smegmatis* has been suggested on the basis of its overexpression phenotype and its protein interaction with SigF (Singh et al., 2015). The genes encoding the proposed

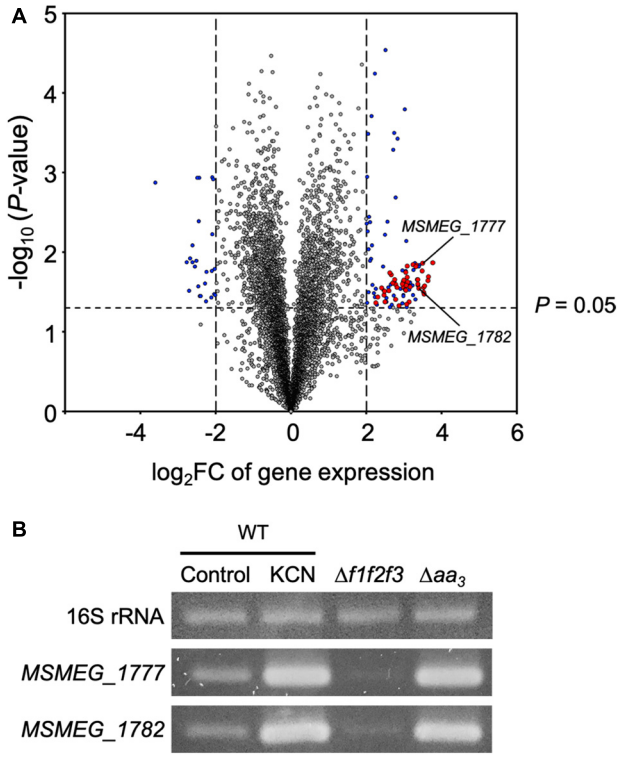


FIGURE 2 | Overlap between the SigF regulon and the genes induced in the Δaa_3 mutant strain of *M. smegmatis*. **(A)** Volcano plot showing the DEGs in the Δaa_3 mutant strain relative to the WT strain. RNA sequencing was carried out using RNA extracted from three biological replicate cultures of the WT and Δaa_3 strains aerobically grown to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. The *x*-axis represents $\log_2$ fold change of gene expression ($\log_2$FC) in the Δaa_3 mutant strain relative to the WT strain, and the y-axis represents $-\log_{10}$ (*P*-value). The vertical dashed lines on the graph mark the border lines indicating the $\log_2$FC values of –2 and 2. The genes, which are differently regulated by more than $|\log_2 FC| > 2$ with *P*-value < 0.05, are depicted by blue-filled circles. Among the identified DEGs, the genes belonging to the SigF regulon are denoted by red-filled circles. The reported downregulated genes ($\log_2 FC < -2$) in a *sigF* mutant strain relative to the WT strain were regarded as the genes belonging to the SigF regulon (Singh et al., 2015). Two representative genes (*MSMEG_1777* and *MSMEG_1782*), whose expression was confirmed by RT-PCR, are indicated by the arrows. **(B)** RT-PCR analysis showing the expression levels of the *MSMEG_1777* and *MSMEG_1782* in the WT, *Δf1f2f3*, and Δaa_3 mutant strains. The *M. smegmatis* strains were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. For treatment of *M. smegmatis* cultures with KCN, the WT strain was grown to an OD_{600} of 0.45–0.5, followed by the addition of KCN to the final concentration of 0.5 mM. The WT cultures were further grown under KCN-untreated (control) or treated (0.5 mM KCN) conditions for 15 min. RT-PCR for 16S rRNA was performed to assure that equal amounts of total RNA were employed for RT-PCR.

anti-SigF antagonists RsfA and RsfB have been also previously identified (Singh et al., 2015). To identify the additional genes whose products are likely to be involved in the SigF PSS, a BLAST search using the RsbW (MSMEG_1803) sequence as a query sequence was performed against the *M. smegmatis* mc^2155 genome, which revealed two additional genes (*MSMEG_1787* and *MSMEG_6129*) that encode the RsbW homologs. As shown in **Figure 3**, The *MSMEG_1787* and *MSMEG_6129* genes were found to be located in the vicinity of the *rsfA* and *rsfB* genes, respectively. The RsbW homologs, whose genes are adjacent to *sigF*, *rsfB*, and *rsfA*, were named as RsbW1, RsbW2, and RsbW3, respectively. Among the three RsbW homologs, RsbW1 composed of 138 amino acids shows the highest homology (66.2% identity) to UsfX of *M. tuberculosis*. The *rsbW2* gene, whose product consists of 148 amino acids, appears to form an operon with the upstream gene *rsfB*. Downstream of *rsbW2* occurs a putative operon that contains the gene (*MSMEG_6130*) encoding a histidine kinase and two adjacent genes (*MSMEG_6128* and *MSMEG_6131*) encoding N-terminally receiver domain-containing proteins. Sequence analysis revealed that MSMEG_6128 and MSMEG_6131 contain a DNA-binding domain and a PP2C phosphatase domain at their C-termini, respectively. RsbW2 was previously reported to encode the protein kinase that phosphorylates the anti-SigF antagonist RsfB (Bowman and Ghosh, 2014). However, its inactivation by mutation was reported not to produce the anticipated phenotype like an increase in expression of the SigF regulon (Bowman and Ghosh, 2014), casting doubt as to whether RsbW2 is the kinase that inactivates RsfB by phosphorylation. The *rsbW3* gene was found to code for the largest RsbW homolog composed of 194 amino acids. When the amino acid sequence of RsbW3 was aligned with that of RsbW1, RsbW3 was found to have an N-terminal extension of 30 amino acids which is not present in RsbW1 (**Supplementary Figure S1**).

The Roles of Three RsbW Homologs in the Regulation of SigF Functionality

To examine whether the identified RsbW homologs function as anti-SigF, we individually inactivated the *rsbW1*, *rsbW2*, and *rsbW3* genes by deleting the corresponding genes, yielding the Δ*rsbW1*, Δ*rsbW2*, and Δ*rsbW3* mutant strains of *M. smegmatis* (**Supplementary Figure S2**). As shown in **Supplementary Figure S3**, both Δ*rsbW1* and Δ*rsbW2* mutants formed yellow-colored colonies on solid agar plates unlike the WT, Δ*rsbW3*, and Δ*sigF* strains of *M. smegmatis*, implying that biosynthesis of the carotenoid isorenieratene is increased in the Δ*rsbW1* and Δ*rsbW2* mutant strains. Given the previous report that the genes involved in biosynthesis of isorenieratene belong to the SigF regulon in *M. smegmatis* (Provvedi et al., 2008; Humpel et al., 2010), it is likely that expression of the SigF regulon is increased in the Δ*rsbW1* and Δ*rsbW2* mutants. Unfortunately, the Δ*rsbW1* and Δ*rsbW2* mutants were found to be instable in terms of yellow pigmentation. They lost the yellow color during cultivation in liquid growth medium, and the altered strains did not restore the yellow pigmentation in solid growth plates. For this reason, we did not examine expression of the SigF-dependent genes in the mutants. Instead, overexpression effects of *rsbW1*, *rsbW2*, and *rsbW3* on *MSMEG_1777* expression were examined to assess the anti-SigF activity of the three RsbW homologs. The genes of the RsbW homologs were overexpressed from an acetamide-inducible promoter on pMHRsbW1, pMHRsbW2, and pMHRsbW3 that are derivatives of the pMH201 integration vector. The expression

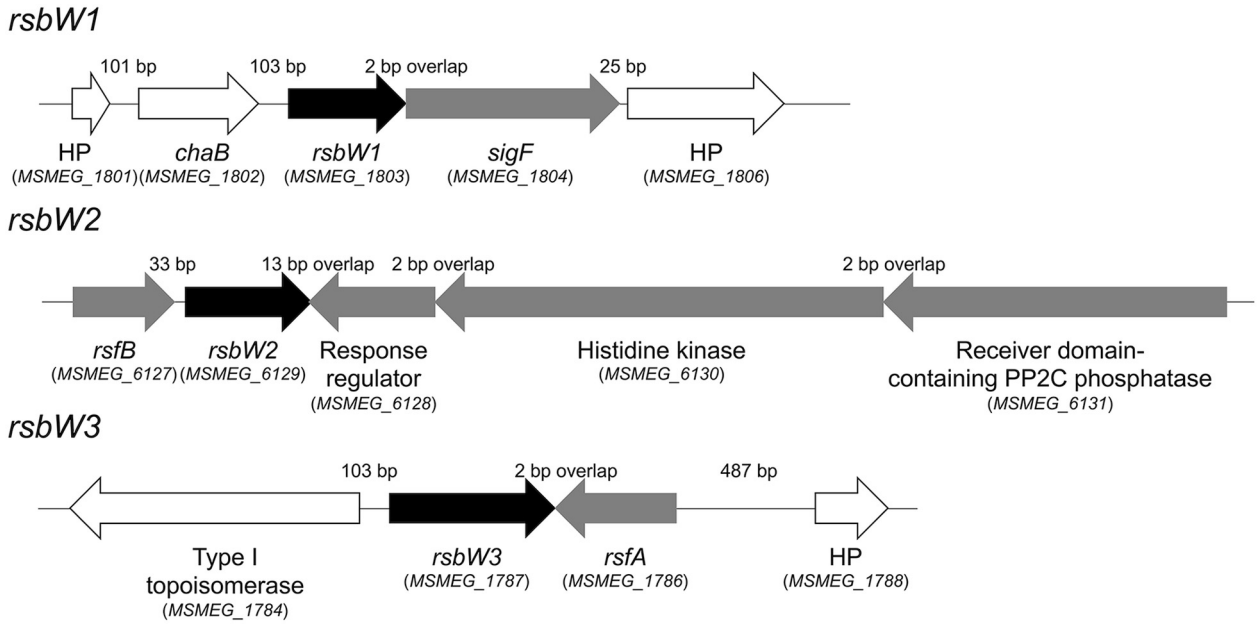


FIGURE 3 | Genetic organization of the *rsbW1*, *rsbW2*, and *rsbW3* loci of *M. smegmatis* mc^2 155. The arrows indicate the open reading frames and their transcriptional direction. The locus tag numbers of the genes are presented in parentheses below the gene names. The lengths of the intergenic and overlapping regions are given as the nucleotide numbers above the schematic diagrams. The genes encoding hypothetical proteins are denoted by HP.

level of *MSMEG_1777* in *M. smegmatis* was determined using an *MSMEG_1777:lacZ* translational fusion, pNCII1777. As shown in **Figure 4**, overexpression of *rsbW1* or *rsbW2* led to a significant decrease in *MSMEG_1777* expression in *M. smegmatis* compared to the control strain with pNCII1777 and pMH201, while overexpression of *rsbW3* rather strongly increased the expression of *MSMEG_1777*. Overexpression of *rsbW1*, *rsbW2*, and *rsbW3* in the *M. smegmatis* strains was verified by Western blotting analysis using a His-tag antibody. Altogether, the results indicate that overexpression of *rsbW1* and *rsbW2* inhibits the transcriptional activity of SigF, while that of *rsbW3* increases the functionality of SigF.

The prerequisite for a protein to act as an anti-sigma factor is protein–protein interactions between the protein and its cognate sigma factor. To determine protein–protein interactions between the RsbW homologs and SigF, we performed yeast two-hybrid assay (Y2H). For the Y2H assay, the *rsbW1*, *rsbW2*, and *rsbW3* genes were cloned into the prey vector pGADT7linker, whereas the *sigF* gene was cloned into the bait vector pGBKT7. As shown in **Figure 5A**, the yeast strain coexpressing RsbW1 and SigF grew well on solid growth medium without histidine (−His) in the presence of up to 5 mM 3-AT. In contrast, coexpression of either RsbW2 or RsbW3 with SigF did not lead to growth of yeast on −His medium in the presence of 3-AT. As expected, the yeast strains coexpressing either RsbW1 or SigF alone did not grow on −His medium in the presence of 3-AT. We next examined *in vitro* protein–protein interactions between SigF and RsbW homologs by means of non-denaturing PAGE analysis using purified SigF and RsbW homologs. Since the RsbW3 protein was not purified to homogeneity, the partially purified protein was used in the experiment. As shown in **Figure 5B**, RsbW1 was shown to interact with SigF as judged by the formation of a new band representing the SigF-RsbW1 complex and disappearance of the SigF band in the non-denaturing PAGE gel. In contrast, the presence of RsbW2 and RsbW3 in the binding mixtures did not result in a decrease in the SigF band intensity in the non-denaturing PAGE gel, indicating that RsbW2 and RsbW3 do not interact with SigF. Taken together, the Y2H and non-denaturing PAGE results suggest that SigF interacts only with RsbW1 among the three RsbW homologs.

The observed overexpression effect of *rsbW3* on *MSMEG_1777* led us to assume that RsbW3 might serve as an anti-SigF antagonist. To examine this assumption, protein–protein interactions between RsbW3 and three RsbW homologs were assessed using Y2H analysis (**Figure 6A**). For the Y2H assay, the *rsbW1*, *rsbW2*, and *rsbW3* genes were cloned into pGADT7linker, and the *rsbW3* gene was cloned into pGBKT7. Only the yeast strain coexpressing RsbW1 and RsbW3 grew on −His medium containing 0.5 mM 3-AT, indicating a possible protein interaction between RsbW1 and RsbW3. Protein–protein interactions between RsbW1 and RsbW3 were also assessed by copurification analysis using affinity chromatography (**Figure 6B**). RsbW1 was copurified with His_6-tagged RsbW3 from crude extracts of the WT strain of *M. smegmatis* expressing His_6-tagged RsbW3, whereas RsbW1 was not copurified with His_6-tagged RsbW2 from the *M. smegmatis* strain expressing His_6-tagged RsbW2, confirming protein–protein interactions between RsbW1 and RsbW3.

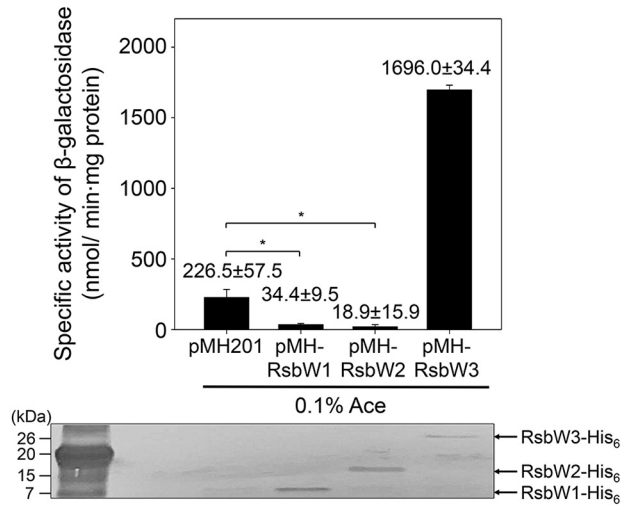


FIGURE 4 | Overexpression effects of *rsbW1*, *rsbW2*, and *rsbW3* on *MSMEG_1777* expression. The *rsbW1*, *rsbW2*, and *rsbW3* genes were overexpressed from the pMH201 derivatives pMHRsbW1, pMHRsbW2, and pMHRsbW3, respectively. The overexpression effects of *rsbW1*, *rsbW2*, and *rsbW3* on the expression level of *MSMEG_1777* were determined in the *M. smegmatis* strains harboring the *MSMEG_1777:lacZ* translational fusion plasmid (pNCII1777) and the pMH201 derivatives. As a control, the *M. smegmatis* strain harboring both pNCII1777 and the empty pMH201 vector was employed in the experiment. The promotor activity of *MSMEG_1777* was measured by determining β-galactosidase activity in the strains grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium in the presence of 0.1% (w/v) acetamide (Ace). All values are the means of the results from three biological replicates. The error bars indicate the standard deviations. Western blotting analysis was conducted for the detection of expressed His_6-tagged RsbW1, RsbW2, and RsbW3. Cell-free crude extracts (30 μg) were separated on SDS-PAGE, followed by Western blotting analysis with a His-tag antibody. *$P < 0.05$.

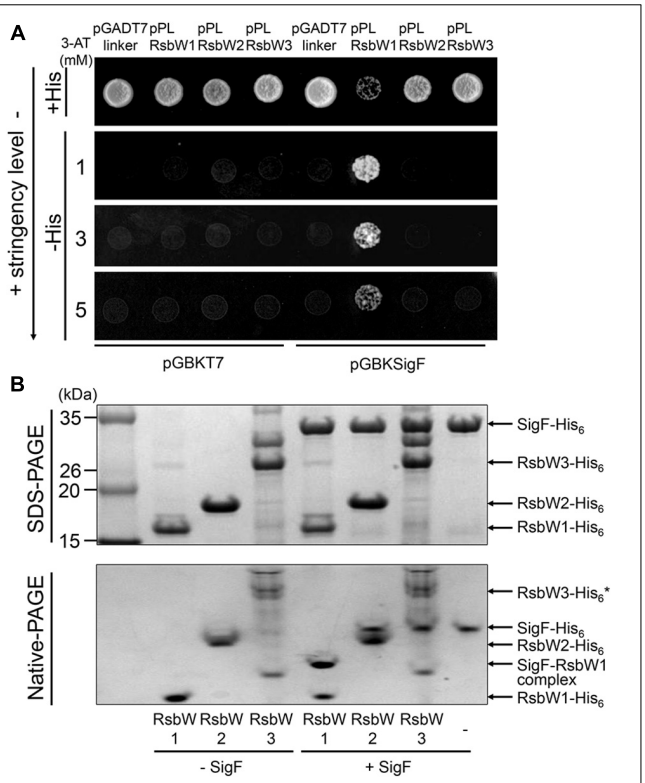


FIGURE 5 | Determination of protein–protein interactions between SigF and RsbW homologs using Y2H assay and non-denaturing PAGE. **(A)** Y2H assay. The *sigF* gene is cloned into pGBKT7 (encoding the GAL4 DNA-binding domain), yielding pGBKSigF. The *rsbW1*, *rsbW2*, *rsbW3* genes were cloned into the pGADT7linker (encoding the GAL4 activation domain), resulting in pPLRsbW1, pPLRsbW2, and pPLRsbW3, respectively. The yeast strains cotransformed with both pGBKSigF and pGADT7linker derivative plasmids were employed for Y2H assay. To discriminate false positive interactions, the yeast strains with empty pGBKT7 and either pGADT7linker or the pGADT7linker derivatives were included in Y2H assay as negative controls. All yeast strains were spotted onto SD/-Leu/-Trp plates (+His) and histidine-deficient SD/-Leu/-Trp/-His plates (–His) containing different concentrations of 3-AT. The numbers to the left indicate the concentration of 3-AT. **(B)** Non-denaturing PAGE analysis. 100 pmol of purified SigF was mixed with purified RsbW1 (200 pmol), RsbW2 (200 pmol), or RsbW3 in binding buffer and incubated for 30 min at 25°C (+SigF). The mixtures were subjected to both SDS-PAGE (upper panel) and native PAGE (lower panel). As controls, the binding mixtures without the addition of SigF, which contain the RsbW homologs alone, were included in the experiment (–SigF). The gels were stained with CBB. The bands representing the RsbW homologs, SigF, and the RsbW1-SigF complex are indicated by the arrows. * Since RsbW3 was not purified to homogeneity, the RsbW3 band in the native PAGE gel is uncertain.

Both RsfA and RsfB Are Functional as Anti-SigF Antagonists, and RsfB Is the Major Anti-SigF Antagonist in *M. smegmatis*

After having established the distinct roles of three RsbW homologs in the SigF PSS, we next examined the roles of the suggested anti-SigF antagonists RsfA and RsfB *in vivo*. We constructed deletion mutants of *rsfA* and *rsfB* in the background of both WT and Δaa_3 strains, and the expression level of *MSMEG_1777* in the WT and mutant strains of *M. smegmatis* was comparatively determined (**Figures 7A,B**). The inactivation of the aa_3 cytochrome *c* oxidase was used for an induction condition of the SigF regulon. The expression level of *MSMEG_1777* was shown to be increased by 4.9-fold in the Δaa_3 strain of *M. smegmatis* relative to that in the WT strain grown under the same conditions. Expression of *MSMEG_1777* was abolished in the Δ*sigF* and Δaa_3Δ*sigF* mutant strains. Both the results confirmed that transcription of *MSMEG_1777* depends on SigF, and that expression of the SigF regulon is induced under respiration-inhibitory conditions. Expression of *MSMEG_1777* was decreased by 26% in the Δ*rsfA* mutant compared to the WT strain. The Δaa_3Δ*rsfA* mutant also showed a 31% decrease in *MSMEG_1777* expression relative to the Δaa_3 mutant strain. It is noteworthy that inactivation of *rsfB* almost abolished expression of *MSMEG_1777* in both WT and Δaa_3 mutant strains. These results suggest that both RsfA and RsfB serve as anti-SigF antagonists, and that RsfB is the major anti-SigF antagonist in *M. smegmatis* under both SigF-activating and SigF-non-activating conditions. To confirm the roles of RsfA and RsfB as anti-SigF antagonists, we examined whether *MSMEG_1777* expression correlates with

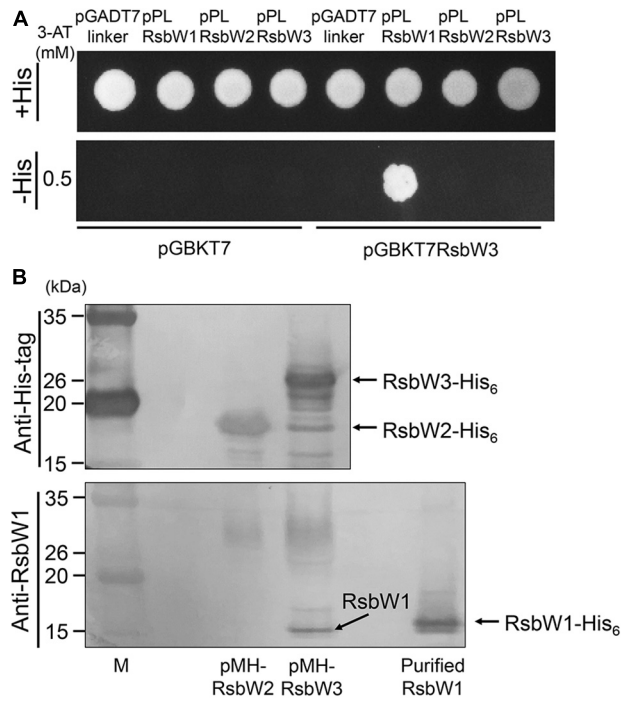


FIGURE 6 | Determination of protein–protein interactions between RsbW3 and RsbW homologs using Y2H and copurification assays. **(A)** Y2H assay. The yeast strains carrying both pGBKT7-derived pGBKRsbW3 and pGADT7linker-derived plasmids (pPLRsbW1, pPLRsbW2, and pPLRsbW3) were employed for Y2H assay. The yeast strains harboring the empty pGBKT7 or pGADT7linker vector were used as negative controls. All yeast strains were spotted onto SD/-Leu/-Trp plates (+His) and histidine-deficient SD/-Leu/-Trp/-His plates (–His) containing 0.5 mM 3-AT. **(B)** Copurification assay. His_6-tagged RsbW2 and RsbW3 were expressed in the WT strains of *M. smegmatis* with pMHRsbW2 and pMHRsbW3, respectively, and partially purified from their crude extracts by means of affinity chromatography using Ni-Sepharose resin. The purified proteins were subjected to Western blotting analyses using a His-3 monoclonal antibody for detection of His_6-tagged RsbW2 and RsbW3 (upper panel), as well as RsbW1 polyclonal antibodies for detection of RsbW1 (lower panel). The bands representing RsbW1 and His_6-tagged RsbW1, RsbW2, and RsbW3 are indicated by the arrows. M, molecular weight marker lanes.

the expression levels of RsfA and RsfB. The *rsfA* and *rsfB* genes were expressed from an acetamide-inducible promoter on pMHRsfA and pMHRsfB, respectively. The expression level of *MSMEG_1777* was determined in the Δ*rsfA* mutant with pMHRsfA and the Δ*rsfB* mutant with pMHRsfB with increasing concentrations of acetamide in growth medium. As shown in **Figures 7C,D**, the expression level of *MSMEG_1777* in the *M. smegmatis* strains with either pMHRsfA or pMHRsfB were gradually increased with increasing concentrations of acetamide, and lower concentrations of acetamide were required for similar levels of *MSMEG_1777* induction in *M. smegmatis* with pMHRsfA compared to *M. smegmatis* with pMHRsfB. Western blotting analysis showed that the amount of expressed RsfB was proportional to the concentration of acetamide. We did not detect the expressed RsfA by Western blotting analysis in the concentration range of acetamide used in the experiment. Taken together, these results provide the strong evidence that both RsfA and RsfB function as anti-SigF antagonists in *M. smegmatis*.

We further examined whether the significantly reduced expression of *MSMEG_1777* in the Δ*rsfB* mutant is caused by the decreased expression of *sigF* or the increased expression of *rsbW1*. Using Western blotting analysis, the protein levels of expressed SigF and RsbW1 were determined in the WT, Δ*rsfA*, and Δ*rsfB* strains that were grown aerobically to an OD_{600} of 0.45–0.5 (**Supplementary Figure S4**). The Western blotting result showed that the protein levels of SigF and RsbW1 in the Δ*rsfA* and Δ*rsfB* mutant strains are not different from those in the WT strain, indicating that the cellular levels of SigF and RsbW1 in *M. smegmatis* are not decreased under SigF-non-activating conditions. This finding can be explained by the presence of a SigF-independent promoter immediately upstream of the *rsbW1-sigF* operon (Gebhard et al., 2008).

The Functionality of RsfB Is Controlled Through Phosphorylation of Ser-63 by RsbW2

A previous study has reported that purified RsbW2 (MSMEG_6129) phosphorylates RsfB (MSMEG_6127) *in vitro* (Bowman and Ghosh, 2014). Using LC-tandem mass spectrometry, eight Ser/Thr residues (Ser-3, Thr-10, Thr-20, Thr-25, Thr-27, Thr-32, Ser-42, and Ser-63) in RsfB have been identified to be the phosphorylation sites by RsbW2 (Bowman and Ghosh, 2014). To specify the functionally important residue(s) among the identified phosphorylation sites in RsfB, we examined the functionality of the mutant forms of RsfB (T10A, T220A, T25A, T27A, T32A, S42A, S63A, and S63E) by determining *MSMEG_1777* expression in the Δ*rsfB* mutant expressing the mutant forms of RsfB. The Ser-3 was excluded from the experiment since the residue is not present in *Bacillus subtilis* RsbV (**Figure 8A**). The Δ*rsfB* mutant of *M. smegmatis* was complemented by introducing the pMV306 derivatives that carry the WT or mutant *rsfB* genes. The WT and Δ*rsfB* strains carrying the pMV306 empty vector were included in the experiment as positive and negative controls, respectively. As shown in **Figure 8B**, the S63A mutation led to a drastic increase in *MSMEG_1777* expression in *M. smegmatis*, whereas the T32A mutation resulted in abolishment of *MSMEG_1777* expression. The expression level of *MSMEG_1777* was shown to be decreased by about 50% in the *M. smegmatis* strain expressing the phosphomimetic S63E form of RsfB compared to the control *M. smegmatis* strain expressing WT RsfB. Western blotting analysis revealed that the WT and mutant forms of RsfB were expressed at similar levels in the *M. smegmatis* strains used in the experiment. These results suggest the followings: (i) unphosphorylated RsfB is the active form of RsfB as an anti-SigF antagonist, (ii) phosphorylation of RsfB on Ser-63 decreases the functionality of RsfB as in the RsfB homologs such as RsbV and SpoIIAA of *B. subtilis* (Bouillet et al., 2018), (iii) Thr-32 is important for anti-SigF antagonist activity of RsfB.

To examine the phosphorylation state of RsfB in *M. smegmatis* grown under SigF-non-activating conditions and whether Ser-63 is the only residue that is phosphorylated, we expressed

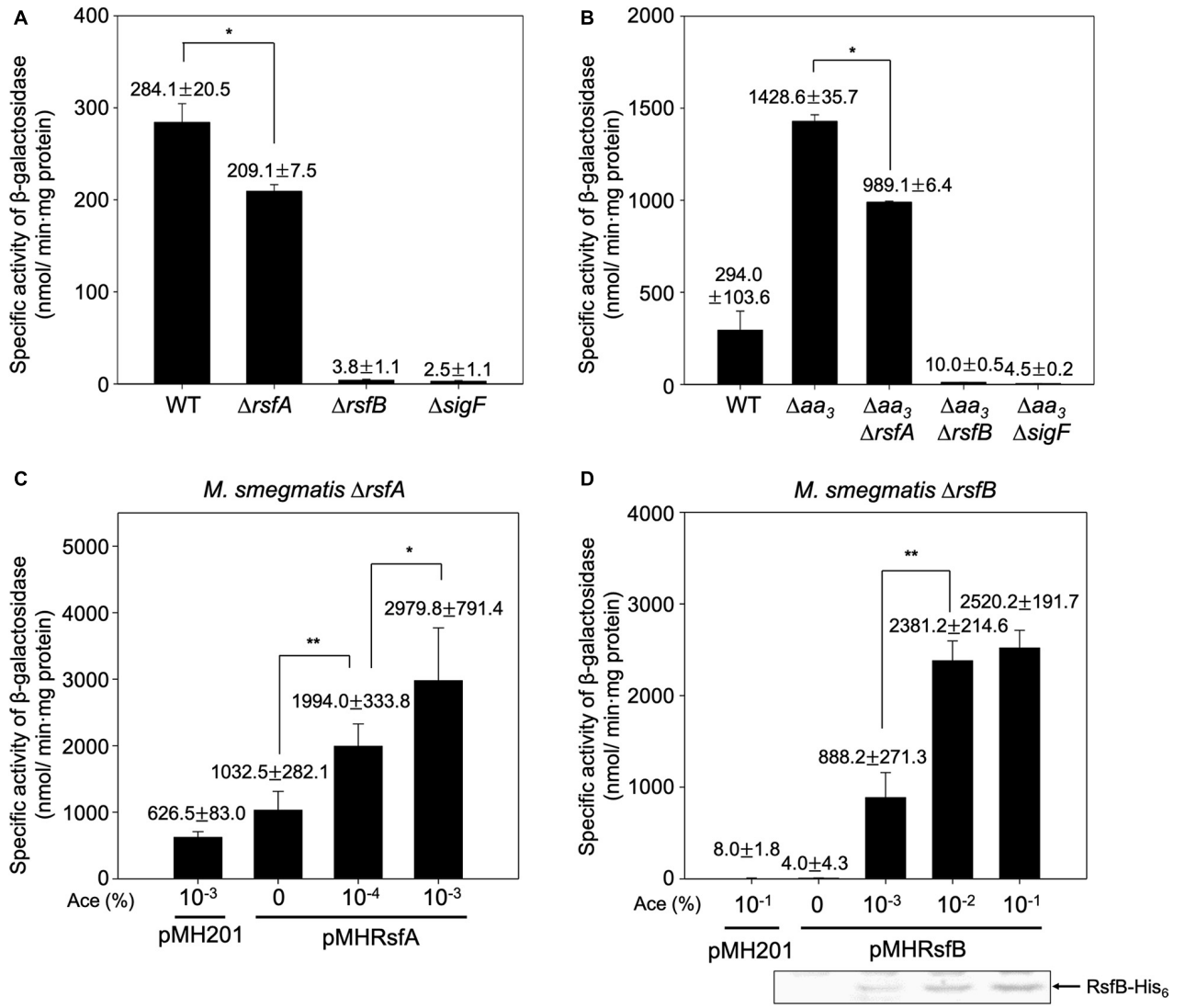


FIGURE 7 | Effects of inactivation and overexpression of *rsfA* and *rsfB* on expression of *MSMEG_1777* in *M. smegmatis*. The expression level of *MSMRG_1777* was determined in the WT, Δ*rsfA*, Δ*rsfB*, Δ*sigF*, Δaa_3, Δaa_3Δ*rsfA*, Δaa_3Δ*rsfB*, and Δaa_3Δ*sigF* strains of *M. smegmatis* harboring pNCII1777 **(A,B)**. The *M. smegmatis* strains were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9 medium. Cell-free crude extracts were used to measure β-galactosidase activity. All values are the means of the results from three biological replicates. The error bars indicate the standard deviations. *$P < 0.05$. Effects of increasing expression of *rsfA* and *rsfB* on *MSMEG_1777* expression were examined in the Δ*rsfA* and Δ*rsfB* strains harboring pNCII1777, respectively **(C,D)**. The pMH201-derived pMHRsfA plasmid was used for the controllable expression of *rsfA* in the Δ*rsfA* strain, and pMHRsfB was used for the controllable expression of *rsfB* in the Δ*rsfB* strain. The strains were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9 medium supplemented with the indicated concentrations of acetamide (Ace). Cell-free crude extracts were used to measure β-galactosidase activity. All values are the means of the results from three biological replicates. The error bars indicate the standard deviations. Western blotting analysis was performed for the detection of expressed His_6-tagged RsfB. Cell-free crude extracts (30 μg) were separated on SDS-PAGE, followed by Western blotting analysis with a His-tag antibody. **$P < 0.01$ and *$P < 0.05$.

the WT and S63E mutant forms of His_6-tagged RsfB in both *M. smegmatis* and *E. coli*, purified the proteins, and determined their phosphorylation state using Phos-tag SDS-PAGE analysis. As shown in **Figure 9**, most fractions of WT RsfB purified from *M. smegmatis* were found to be phosphorylated, while S63E RsfB purified from *M. smegmatis* was not phosphorylated at all. Both WT and S63E mutant forms of RsfB purified from *E. coli* were found to be unphosphorylated. The results indicate that Ser-63 in RsfB is the residue that is phosphorylated in *M. smegmatis*, and that *E. coli* does not have the protein kinase that can phosphorylate RsfB.

We next examined the effect of Ser-63 phosphorylation on protein–protein interactions between RsfB and RsbW1. In place of phosphorylated RsfB, the phosphomimetic S63E mutant form of RsfB was employed for non-denaturing PAGE analysis. The WT and S63A RsfB proteins purified from *E. coli* were used as unphosphorylated RsfB. As shown in **Figure 10**, both WT RsfB and S63A RsfB interacted with purified RsbW1

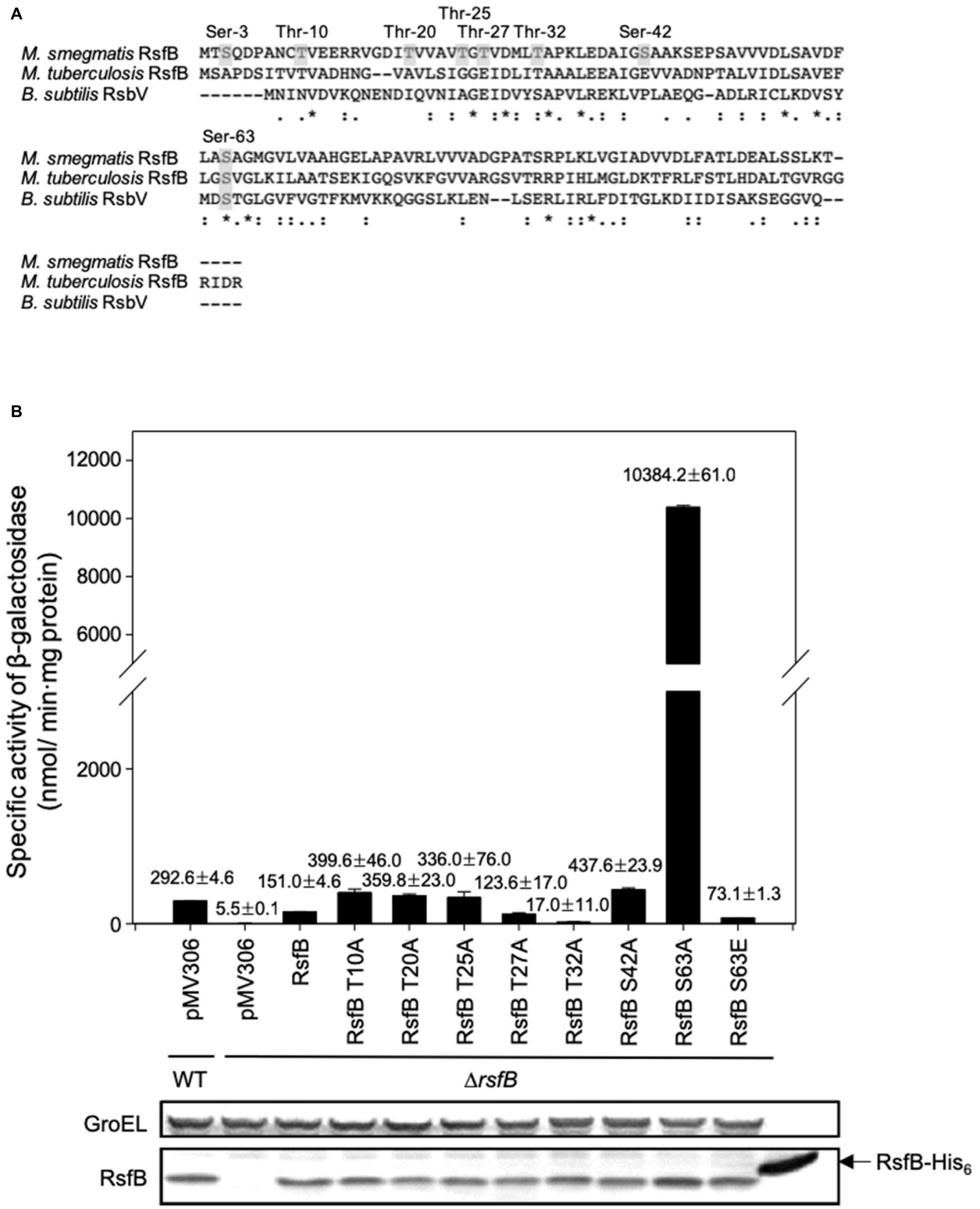


FIGURE 8 | Identification of the amino acid residues that are responsible for the inactivation of RsfB by phosphorylation. **(A)** Multiple sequence alignment of the RsfB homologs of *M. smegmatis*, *M. tuberculosis*, and *B. subtilis* was generated using ClustalW. The asterisks and colons denote the conserved and conservatively substituted amino acid residues, respectively. The residues of *M. smegmatis* RsfB, which were identified to be phosphorylated by MSMEG_6129 *in vitro* (Bowman and Ghosh, 2014), are shown in the gray background. **(B)** Effects of T10A, T20A, T25A, T27A, S42A, S63A, and S63E mutations on the functionality of RsfB *in vivo*.

(Continued)

FIGURE 8 | Continued
The Δ*rsfB* strain harboring pNCII1777 was complemented with pMV306RsfB and its derivatives carrying the mutated *rsfB* gene (pMVRsfBT10A, pMVRsfBT20A, pMVRsfBT25A, pMVRsfBT27A, pMVRsfBT32A, pMVRsfBS42A, pMVRsfBS63A, pMVRsfBS63E). The complementation test was performed by determining the expression level of *MSMEG_1777* in the *M. smegmatis* strains. As controls, the *M. smegmatis* WT and Δ*rsfB* mutant strains with both pNCII1777 and the empty vector pMV306 were included in the experiment. The *M. smegmatis* strains were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. Cell-free crude extracts were used to measure β-galactosidase activity. All values are the means of the results from three biological replicates. The error bars indicate the standard deviations. Protein levels of the WT and mutant forms of RsfB expressed in the strains were detected by Western blotting analysis with RsfB polyclonal antibodies. As a loading control, GroEL was detected by a GroEL monoclonal antibody.

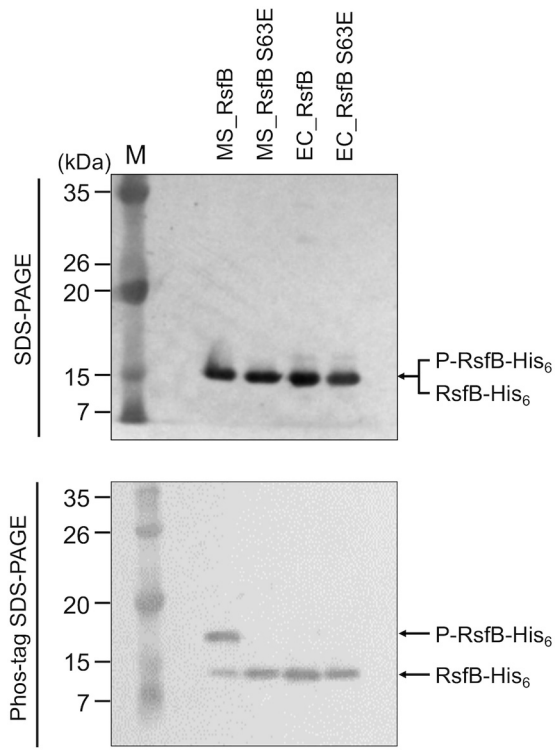


FIGURE 9 | Phosphorylation state of the WT and S63E mutant forms of RsfB in *M. smegmatis* and *E. coli*. The WT and S63E mutant forms of His_6-tagged RsfB were purified from *E. coli* and *M. smegmatis* that were aerobically grown to an OD_{600} of 0.4–0.5. 2 μg each of the WT and S63E mutant forms of RsfB were subjected to SDS-PAGE and Phos-tag SDS-PAGE, followed by Western blotting analysis with RsfB polyclonal antibodies. The bands representing unphosphorylated RsfB (RsfB-His_6) and phosphorylated RsfB (P-RsfB-His_6) are indicated by the arrows. MS_RsfB, RsfB purified from *M. smegmatis*. EC_RsfB, RsfB purified from *E. coli* M, molecular weight marker lanes.

and formed the retarded bands representing the RsfB-RsbW1 complex in non-denaturing PAGE. The intensity of the RsfB-RsbW1 complex bands was increased up to the ratio of RsbW1 to RsfB to be 1:1 with increasing amounts of WT RsfB and S63A RsfB. In contrast, the S63E mutant form of RsfB did not give rise to the RsfB-RsbW1 complex band even at high concentrations of S63E RsfB. The results suggest that phosphorylation of Ser-63 inactivates RsfB to render it unable to interact with RsbW1.

Since the RsbW homologs of *Bacillus* SigB had been demonstrated to function as both anti-SigB and the protein kinase phosphorylating the anti-SigB antagonist (Benson and Haldenwang, 1993; Dufour and Haldenwang, 1994), we wondered whether in addition to RsbW2, RsbW1, and RsbW3 have the protein kinase activity phosphorylating RsfB. To examine this possibility, we performed *in vitro* kinase assay using purified RsbW homologs and RsfB. As shown in **Figure 11A**, only RsbW2 could phosphorylate unphosphorylated RsfB purified from *E. coli*, which is in good agreement with the fact that RsbW1 and RsbW3 are closely clustered with anti-sigma factors lacking the kinase activity, while RsbW2 is clustered with kinase-positive anti-sigma factors (**Supplementary Figure S5**). We also examined the phosphorylation state of RsfB in the WT and Δ*rsbW2* mutant strains grown under SigF-non-activating conditions using Phos-tag SDS-PAGE and Western blotting analysis (**Figure 11B**). The Δ*5437* mutant of *M. smegmatis* was included in the experiment, since it had been suggested that MSMEG_5437 is a Ser/Thr protein kinase that might modulate RsbW2 activity by phosphorylation (Bowman and Ghosh, 2014). Phos-tag SDS-PAGE showed that RsfB in the Δ*rsbW2* mutant was not phosphorylated in contrast to RsfB in the WT and Δ*5437* strains of *M. smegmatis*, indicating that RsbW2 is the only protein kinase that phosphorylates RsfB in *M. smegmatis*, and that RsbW2 is still active in the Δ*5437* mutant.

DISCUSSION

The genome of *M. smegmatis* contains three genes encoding RsbW homologs (RsbW1, RsbW2, and RsbW3). Among them, RsbW1 shows the highest degree of homology to UsfX that is a known anti-SigF in *M. tuberculosis*. Y2H and non-denaturing PAGE analysis revealed the interaction of RsbW1 with SigF, which is in good agreement with the previous result from bacterial two-hybrid assay (Singh et al., 2015). Overexpression of *rsbW1* in *M. smegmatis* led to a significant reduction in expression of *MSMEG_1777* that is under the control of SigF. Furthermore, disruption of the *rsbW1* gene by deletion resulted in an increase in yellow pigmentation of *M. smegmatis* colonies, which appears to be the result of increased isorenieratene biosynthesis. All of these results indicate that RsbW1 is a *bona fide* anti-SigF in *M. smegmatis*.

RsbW2 is most deviated among the RsbW homologs with regard to the reciprocal sequence homology (**Supplementary Figure S5**). In contrast to RsbW1, RsbW2 was shown not to interact with SigF in Y2H and non-denaturing PAGE analysis, implying that RsbW2 does not play a direct role as an anti-SigF. However, overexpression and inactivation of *rsbW2* gave rise to the same phenotype as those of *rsbW1* in terms of

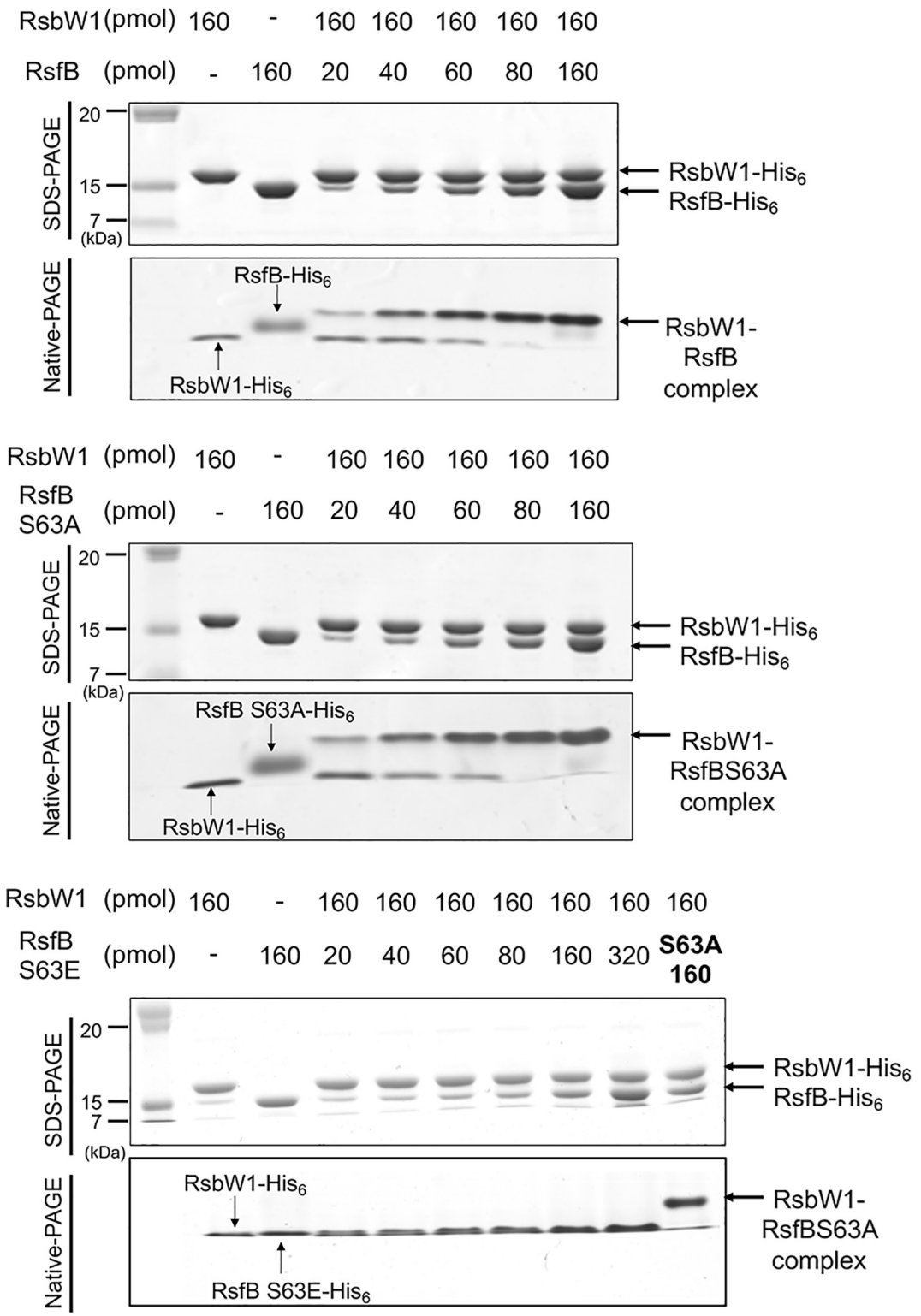


FIGURE 10 | Determination of protein–protein interactions between RsbW1 and several forms of RsfB (WT, S63A, and S63E) by non-denaturing PAGE. 160 pmol of purified RsbW1 was mixed with increasing amounts of the purified WT and mutant forms (S63A and S63E) of RsfB in binding buffer [40 mM Tris–HCl (pH 8.0), 0.01 mM EDTA, 10 mM $MgCl_2$, 20% (v/v) glycerol] and incubated for 30 min at 25°C. The mixtures were subjected to both SDS-PAGE (upper panel) and native PAGE (lower panel). The gels were stained with CBB. The bands representing RsbW1, RsfB (WT and mutant forms), and RsbW1-RsfB complex are indicated by the arrows.

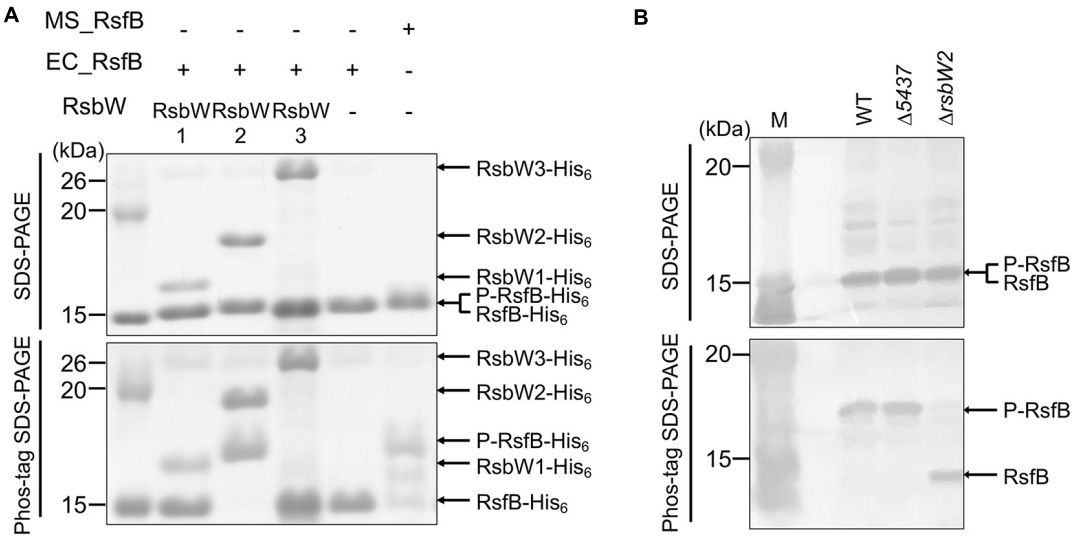


FIGURE 11 | Determination of kinase activities of the RsbW homologs acting on RsfB. **(A)** RsfB purified from *E. coli* was mixed with partially purified RsbW1, purified RsbW2, or partially purified RsbW3 in reaction buffer [20 mM Tris–Cl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, and 10 mM $MnCl_2$]. The phosphorylation reactions were initiated by the addition of 100 μM ATP and performed for 30 min at 30°C. Subsequently, the reaction mixtures were subjected to both SDS-PAGE (upper panel) and Phos-tag SDS-PAGE (lower panel). The gels were stained with CBB. RsfB partially purified from *M. smegmatis* was used as the phosphorylated control of RsfB. The bands of the His_6-tagged RsbW homologs, unphosphorylated RsfB (RsfB-His_6), and phosphorylated RsfB (P-RsfB-His_6) are indicated by the arrows. MS_RsfB, RsfB purified from *M. smegmatis*. EC_RsfB, RsfB purified from *E. coli*. **(B)** The phosphorylation state of the RsfB protein in the WT, Δ*5437*, and Δ*rsbW2* mutant strains of *M. smegmatis*. The strains were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. Cell-free crude extracts (60 μg) of the strains were separated simultaneously on both SDS-PAGE and Phos-tag SDS-PAGE, followed by Western blotting analysis with RsfB polyclonal antibodies. The bands representing unphosphorylated RsfB (RsfB) and phosphorylated RsfB (P-RsfB) are indicated by the arrows. M, molecular weight marker lanes.

both *MSMEG_1777* expression and colony pigmentation. These results suggest that RsbW2 has an activity to decrease SigF functionality without direct binding to SigF. A clue about the anti-SigF activity of RsbW2 came from the kinase motifs (N, G1, and G2) that are conserved in RsbW2. Like *B. subtilis* RsbW that can inactivate the anti-SigB antagonist RsbV through phosphorylation (Dufour and Haldenwang, 1994), RsbW2 has the protein kinase activity that inhibits the functionality of RsfB by phosphorylation of Ser-63. RsfB was shown to exist in an unphosphorylated form in the Δ*rsbW2* mutant in contrast to the isogenic WT strain in which most fractions of RsfB exist in a phosphorylated form, which indicates that RsbW2 is the kinase that can phosphorylate RsfB in *M. smegmatis*. It is worth noting that RsbW of the SigB PSS in *Bacillus* species and *Streptomyces coelicolor* acts as both protein kinase for its cognate anti-SigB antagonist (RsbV) and anti-sigma factor (Benson and Haldenwang, 1993; Dufour and Haldenwang, 1994; Voelker et al., 1996; Lee et al., 2004; van Schaik and Abee, 2005), while RsbW1 and RsbW2 of *M. smegmatis* are specialized to function as anti-SigF and protein kinase, respectively.

Interestingly, both Δ*rsbW1* and Δ*rsbW2* mutants of *M. smegmatis* showed phenotypic instability in terms of yellow pigmentation. When first obtained, both the mutants exhibited yellow pigmentation on solid 7H9-glucose medium. However, when the mutant strains were passed through successive subcultures in solid and especially liquid growth media, the mutant strains lost yellow pigmentation and showed only basal levels of *MSMEG_1777* expression (data not shown). This observation implies that excessive expression of the SigF regulon is detrimental to *M. smegmatis*, leading to secondary mutations that mitigate the expression of the SigF regulon.

RsbW3 is closer to RsbW1 than RsbW2 in terms of sequence homology. In contrast to *rsbW1*, overexpression of *rsbW3* in *M. smegmatis* resulted in a significant increase in *MSMEG_1777* expression. Both the lack of protein–protein interactions between RsbW3 and SigF and the overexpression effect of *rsbW3* suggest the role of RsbW3 as an anti-SigF antagonist. Given both the quaternary structure (homodimers) of RsbW (SpoIIAB)-like anti-sigma factors (Campbell et al., 2002; Masuda et al., 2004) and the results demonstrating protein–protein interactions between RsbW1 and RsbW3 (**Figure 6**), we assume that RsbW3 likely inactivates RsbW1 by forming a heterodimer when overexpressed. The presence of a SigF–recognizing promoter (GTTT-N_{17}-GGGTAA) upstream of *rsbW3* (Table S4) and abolishment of *rsbW3* expression by the inactivation of the *sigF* gene (**Supplementary Figure S6**) indicate that *rsbW3* belongs to the SigF regulon. Based on these findings, we suggest that RsbW3 serves as a booster for expression of the SigF regulon under SigF-activating conditions via the positive feedback loop.

So far, the roles of RsfA and RsfB as anti-SigF antagonists in mycobacteria have been predicted from both their protein interactions with anti-SigF (RsbW and UsfX) and the result from *in vitro* transcription analysis (Beaucher et al., 2002; Malik et al., 2008, 2009; Singh et al., 2015). Through both deletion and overexpression of *rsfA* and *rsfB*, we first demonstrated the physiological roles of RsfA and RsfB as anti-SigF antagonists in

M. smegmatis in vivo. As judged by the RPKM values obtained from RNA sequencing analysis (Lee et al., 2018), the transcript level of *rsfB* was estimated to be ∼6-fold higher than that of *rsfA* in *M. smegmatis* grown aerobically to an OD_{600} of 0.4–0.5 in 7H9-glucose medium (**Supplementary Figure S7**). The difference in the expression levels of *rsfA* and *rsfB* might give a clue explaining the dominant role of RsfB as an anti-SigF antagonist. The observation that expression of *MSMEG_1777* was nearly abolished in the Δ*rsfB* mutant despite the presence of RsfA implies that the cellular level of active RsfA might not be sufficient to quarantine RsbW1 to such an extent as to induce the SigF regulon in the absence of RsfB. The result of *rsfA* and *rsfB* overexpression using an acetamide-inducible promoter clearly showed that when RsfA is sufficiently expressed, it acts as an anti-SigF antagonist more efficiently than RsfB. This observation is in good agreement with the previous report demonstrating that RsfA interacts more strongly with RsbW1 (UsfX) than RsfB (Singh et al., 2015). When the Δ*rsfA* mutant was complemented by introducing pMHRsfA (pMH201:*rsfA*), the strain did not grow at 0.01% acetamide in contrast to the Δ*rsfB* mutant carrying pMHRsfB (pMH201:*rsfB*) that grew, albeit slowly, even in the presence of 0.1% acetamide (data not shown). The inability of the Δ*rsfA* mutant with pMHRsfA to grow in the presence of 0.01% acetamide reinforces our assumption that expression of the SigF regulon in excess is toxic to *M. smegmatis*.

We found that Ser-63 is the amino acid residue of RsfB that is phosphorylated by RsbW2. It was also demonstrated that the S63A mutant form of RsfB interacts with RsbW1 with a similar affinity as the unphosphorylated form of WT RsfB, while the phosphomimetic (S63E) form of RsfB does not interact with RsbW1. These results confirm that the phosphorylation state of Ser-63 determines the functionality of RsfB as an anti-SigF antagonist. The importance of the corresponding serine residue in the functionality of anti-sigma factor antagonists has been reported for several RsfB homologs (RsfB of *M. tuberculosis*, RsbV and SpoIIAB of *B. subtilis*) (Diederich et al., 1994; Najafi et al., 1995; Yang et al., 1996; Beaucher et al., 2002).

The Δaa_3 mutant of *M. smegmatis* lacking the aa_3 cytochrome *c* oxidase of the respiratory ETC has been reported to exhibit 53% of the oxygen consumption rate observed for the isogenic WT strain (Jeong et al., 2018), indicating that electron flow through the ETC is inhibited in the mutant by ∼50% relative to the WT strain. The finding that expression of the SigF regulon is significantly induced in the Δaa_3 mutant relative to the WT stain implies that the availability of free active SigF is increased in response to inhibition of the respiratory ETC. This observation is in good agreement with the suggestion that SigF makes direct contributions to transcriptomic remodeling in *M. smegmatis* under hypoxic growth conditions (Martini et al., 2019). The activation of SigF under respiration-inhibitory conditions might result from energy limitation as in the case of the SigB PSS in *Bacillus* species (Hecker and Volker, 2001; Marles-Wright and Lewis, 2007; de Been et al., 2011; Paget, 2015), or from other factors associated with ETC functions such as changes in the redox state of electron carriers, membrane potential, and proton motive force, etc. The inactivation of the aa_3 cytochrome *c* oxidase in the background of the Δ*rsfB* mutant was shown not to lead to induction of *MSMEG_1777* expression. This result implies that RsfB mediates the induction of the SigF regulon under respiration-inhibitory conditions.

We demonstrated that the sole protein kinase that phosphorylates RsfB in *M. smegmatis* is RsbW2. RsbW2 has been suggested to be phosphorylated by a Ser/Thr protein kinase, MSMEG_5437 (Bowman and Ghosh, 2014), although the role of MSMEG_5437 in the SigF PSS remains elusive. In the vicinity of the *rsfB-rsbW2* operon occur the genes encoding a histidine kinase (MSMEG_6130) and a receiver domain-containing PP2C phosphatase (MSMEG_6131). The PP2C-family phosphatases are known to be responsible for dephosphorylation of the anti-SigB antagonist RsbV in *Bacillus* species and *S. coelicolor* (de Been et al., 2011). Indeed, our preliminary result showed that MSMEG_6131 could dephosphorylate the phosphorylated RsfB protein (data not shown). It is conceivable that the phosphorylation state of RsfB might be modulated by the combined control of the kinase activity of RsbW2 and the phosphatase activity of MSMEG_6131 that might be regulated by MSMEG_5437 Ser/The protein kinase and MSMEG_6130 histidine kinase, respectively. Further study is required to reveal the mechanism by which inhibition of the respiratory ETC leads to the activation of SigF in *M. smegmatis*.

DATA AVAILABILITY STATEMENT

The RNA sequencing data described in this study have been deposited in NCBI's Gene Expression Omnibus and are accessible through the GEO Series accession number GSE155251.

AUTHOR CONTRIBUTIONS

J-IO, S-YS, and YO: conception or design of the study. YO, S-YS, H-JK, GH, and H-YK: acquisition of the data. YO, S-YS, JH, and J-IO: analysis or interpretation of the data and writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2020R1A2C1005305).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.588487/full#supplementary-material

REFERENCES

Barbieri, C. M., and Stock, A. M. (2008). Universally applicable methods for monitoring response regulator aspartate phosphorylation both *in vitro* and *in vivo* using Phos-tag-based reagents. *Anal. Biochem.* 376, 73–82. doi: 10.1016/j.ab.2008.02.004

Beaucher, J., Rodrigue, S., Jacques, P. E., Smith, I., Brzezinski, R., and Gaudreau, L. (2002). Novel *Mycobacterium tuberculosis* anti-σ factor antagonists control σF activity by distinct mechanisms. *Mol. Microbiol.* 45, 1527–1540. doi: 10.1046/j.1365-2958.2002.03135.x

Benson, A. K., and Haldenwang, W. G. (1993). *Bacillus subtilis* σB is regulated by a binding protein (RsbW) that blocks its association with core RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 90, 2330–2334. doi: 10.1073/pnas.90.6.2330

Betts, J. C., Lukey, P. T., Robb, L. C., Mcadam, R. A., and Duncan, K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.* 43, 717–731. doi: 10.1046/j.1365-2958.2002.02779.x

Bouillet, S., Arabet, D., Jourlin-Castelli, C., Mejean, V., and Iobbi-Nivol, C. (2018). Regulation of σ factors by conserved partner switches controlled by divergent signalling systems. *Environ. Microbiol. Rep.* 10, 127–139. doi: 10.1111/1758-2229.12620

Bowman, J., and Ghosh, P. (2014). A complex regulatory network controlling intrinsic multidrug resistance in *Mycobacterium smegmatis*. *Mol. Microbiol.* 91, 121–134. doi: 10.1111/mmi.12448

Campbell, E. A., Masuda, S., Sun, J. L., Muzzin, O., Olson, C. A., Wang, S., et al. (2002). Crystal structure of the *Bacillus stearothermophilus* anti-σ factor SpoIIAB with the sporulation σ factor σF. *Cell* 108, 795–807. doi: 10.1016/s0092-8674(02)00662-1

Chen, C. C., Lewis, R. J., Harris, R., Yudkin, M. D., and Delumeau, O. (2003). A supramolecular complex in the environmental stress signalling pathway of *Bacillus subtilis*. *Mol. Microbiol.* 49, 1657–1669. doi: 10.1046/j.1365-2958.2003.03663.x

Chen, P., Ruiz, R. E., Li, Q., Silver, R. F., and Bishai, W. R. (2000). Construction and characterization of a *Mycobacterium tuberculosis* mutant lacking the alternate sigma factor gene, *sigF*. *Infect. Immun.* 68, 5575–5580. doi: 10.1128/iai.68.10.5575-5580.2000

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544. doi: 10.1038/31159

de Been, M., Francke, C., Siezen, R. J., and Abee, T. (2011). Novel σB regulation modules of Gram-positive bacteria involve the use of complex hybrid histidine kinases. *Microbiology* 157, 3–12. doi: 10.1099/mic.0.045740-0

DeMaio, J., Zhang, Y., Ko, C., and Bishai, W. R. (1997). *Mycobacterium tuberculosis sigF* is part of a gene cluster with similarities to the *Bacillus subtilis sigF* and *sigB* operons. *Int. J. Tuberc. Lung Dis.* 78, 3–12. doi: 10.1016/s0962-8479(97)90010-1

DeMaio, J., Zhang, Y., Ko, C., Young, D. B., and Bishai, W. R. (1996). A stationary-phase stress-response sigma factor from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 93, 2790–2794. doi: 10.1073/pnas.93.7.2790

Diederich, B., Wilkinson, J. F., Magnin, T., Najafi, M., Errington, J., and Yudkin, M. D. (1994). Role of interactions between SpoIIAA and SpoIIAB in regulating cell-specific transcription factor σF of *Bacillus subtilis*. *Genes Dev.* 8, 2653–2663. doi: 10.1101/gad.8.21.2653

Dufour, A., and Haldenwang, W. G. (1994). Interactions between a *Bacillus subtilis* anti-σ factor (RsbW) and its antagonist (RsbV). *J. Bacteriol.* 176, 1813–1820. doi: 10.1128/jb.176.7.1813-1820.1994

Dutta, D., Srivastava, V., Tripathi, A., Singh, V., Ansari, M. M., Pant, G., et al. (2019). *Mycobacterium bovis sigF* mutant exhibits altered surface phenotype and compromised pathogenesis. *Tuberculosis* 118, 1–13. doi: 10.1016/j.tube.2019.07.006

Forrellad, M. A., Klepp, L. I., Gioffre, A., Sabio, Y., Garcia, J., Morbidoni, H. R., et al. (2013). Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence* 4, 3–66. doi: 10.4161/viru.22329

Gebhard, S., Humpel, A., Mclellan, A. D., and Cook, G. M. (2008). The alternative sigma factor SigF of *Mycobacterium smegmatis* is required for survival of heat shock, acidic pH and oxidative stress. *Microbiology* 154, 2786–2795. doi: 10.1099/mic.0.2008/018044-0

Geiman, D. E., Kaushal, D., Ko, C., Tyagi, S., Manabe, Y. C., Schroeder, B. G., et al. (2004). Attenuation of late-stage disease in mice infected by the *Mycobacterium tuberculosis* mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis. *Infect. Immun.* 72, 1733–1745. doi: 10.1128/iai.72.3.1733-1745.2004

Graham, J. E., and Clark-Curtiss, J. E. (1999). Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). *Proc. Natl. Acad. Sci. U.S.A.* 96, 11554–11559. doi: 10.1073/pnas.96.20.11554

Green, M., and Sambrook, J. (2012). *Molecular Cloning: A Laboratory Manual*, 4th Edn. New York, NY: Cold Spring Harbor Laboratory Press.

Greenstein, A. E., Hammel, M., Cavazos, A., and Alber, T. (2009). Interdomain communication in the *Mycobacterium tuberculosis* environmental phosphatase Rv1364c. *J. Biol. Chem.* 284, 29828–29835. doi: 10.1074/jbc.M109.056168

Guthrie, C., and Fink, G. R. (1991). Guide to yeast genetics and molecular biology. *Methods Enzymol.* 194, 1–932.

Hartkoorn, R. C., Sala, C., Uplekar, S., Busso, P., Rougemont, J., and Cole, S. T. (2012). Genome-wide definition of the SigF regulon in *Mycobacterium tuberculosis*. *J. Bacteriol.* 194, 2001–2009. doi: 10.1128/JB.06692-11

Hecker, M., Pane-Farre, J., and Volker, U. (2007). SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu. Rev. Microbiol.* 61, 215–236. doi: 10.1146/annurev.micro.61.080706.093445

Hecker, M., and Volker, U. (2001). General stress response of *Bacillus subtilis* and other bacteria. *Adv. Microb. Physiol.* 44, 35–91. doi: 10.1016/s0065-2911(01)44011-2

Humpel, A., Gebhard, S., Cook, G. M., and Berney, M. (2010). The SigF regulon in *Mycobacterium smegmatis* reveals roles in adaptation to stationary phase, heat, and oxidative stress. *J. Bacteriol.* 192, 2491–2502. doi: 10.1128/JB.00035-10

Jeong, J. A., Park, S. W., Yoon, D., Kim, S., Kang, H. Y., and Oh, J. I. (2018). Roles of alanine dehydrogenase and Induction of Its gene in *Mycobacterium smegmatis* under respiration-inhibitory conditions. *J. Bacteriol.* 200:e00152-18. doi: 10.1128/JB.00152-18

Karls, R. K., Guarner, J., Mcmurray, D. N., Birkness, K. A., and Quinn, F. D. (2006). Examination of *Mycobacterium tuberculosis* sigma factor mutants using low-dose aerosol infection of guinea pigs suggests a role for SigC in pathogenesis. *Microbiology* 152, 1591–1600. doi: 10.1099/mic.0.28591-0

Keren, I., Minami, S., Rubin, E., and Lewis, K. (2011). Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters. *mBio* 2:e00100-11. doi: 10.1128/mBio.00100-11

Kim, M. J., Park, K. J., Ko, I. J., Kim, Y. M., and Oh, J. I. (2010). Different roles of DosS and DosT in the hypoxic adaptation of Mycobacteria. *J. Bacteriol.* 192, 4868–4875. doi: 10.1128/JB.00550-10

Lee, E. J., Cho, Y. H., Kim, H. S., Ahn, B. E., and Roe, J. H. (2004). Regulation of σB by an anti- and an anti-anti-sigma factor in *Streptomyces coelicolor* in response to osmotic stress. *J. Bacteriol.* 186, 8490–8498. doi: 10.1128/JB.186.24.8490-8498.2004

Lee, H. N., Ji, C. J., Lee, H. H., Park, J., Seo, Y. S., Lee, J. W., et al. (2018). Roles of three FurA paralogs in the regulation of genes pertaining to peroxide defense in *Mycobacterium smegmatis* mc2 155. *Mol. Microbiol.* 108, 661–682. doi: 10.1111/mmi.13956

Malik, S. S., Luthra, A., and Ramachandran, R. (2009). Interactions of the *M. tuberculosis* UsfX with the cognate sigma factor SigF and the anti-anti sigma factor RsfA. *Biochim. Biophys. Acta* 1794, 541–553. doi: 10.1016/j.bbapap.2008.11.007

Malik, S. S., Luthra, A., Srivastava, S. K., and Ramachandran, R. (2008). *Mycobacterium tuberculosis* UsfX (Rv3287c) exhibits novel nucleotide binding and hydrolysis properties. *Biochem. Biophys. Res. Commun.* 375, 465–470. doi: 10.1016/j.bbrc.2008.08.043

Manganelli, R. (2014). Sigma factors: key molecules in *Mycobacterium tuberculosis* physiology and virulence. *Microbiol. Spectr.* 2, 1–23. doi: 10.1128/microbiolspec.MGM2-0007-2013

Manganelli, R., Provvedi, R., Rodrigue, S., Beaucher, J., Gaudreau, L., and Smith, I. (2004). σ factors and global gene regulation in *Mycobacterium tuberculosis*. *J. Bacteriol.* 186, 895–902. doi: 10.1128/jb.186.4.895-902.2004

Mariani, F., Cappelli, G., Riccardi, G., and Colizzi, V. (2000). *Mycobacterium tuberculosis* H37Rv comparative gene-expression analysis in synthetic medium and human macrophage. *Gene* 253, 281–291. doi: 10.1016/s0378-1119(00)00249-3

Marles-Wright, J., and Lewis, R. J. (2007). Stress responses of bacteria. *Curr. Opin. Struct. Biol.* 17, 755–760. doi: 10.1016/j.sbi.2007.08.004

Martini, M. C., Zhou, Y., Sun, H., and Shell, S. S. (2019). Defining the Transcriptional and Post-transcriptional Landscapes of *Mycobacterium smegmatis* in Aerobic Growth and Hypoxia. *Front. Microbiol.* 10:591. doi: 10.3389/fmicb.2019.00591

Masuda, S., Murakami, K. S., Wang, S., Anders Olson, C., Donigian, J., Leon, F., et al. (2004). Crystal structures of the ADP and ATP bound forms of the *Bacillus* anti-σ factor SpoIIAB in complex with the anti-anti-σ SpoIIAA. *J. Mol. Biol.* 340, 941–956. doi: 10.1016/j.jmb.2004.05.040

Michele, T. M., Ko, C., and Bishai, W. R. (1999). Exposure to antibiotics induces expression of the *Mycobacterium tuberculosis sigF* gene: implications for chemotherapy against mycobacterial persistors. *Antimicrob. Agents Chemother.* 43, 218–225. doi: 10.1128/aac.43.2.218

Misra, R., Menon, D., Arora, G., Virmani, R., Gaur, M., Naz, S., et al. (2019). Tuning the *Mycobacterium tuberculosis* alternative sigma factor sigf through the multidomain regulator Rv1364c and osmosensory kinase protein Kinase D. *J. Bacteriol.* 201, 1–16. doi: 10.1128/JB.00725-18

Mouncey, N. J., and Kaplan, S. (1998). Redox-dependent gene regulation in *Rhodobacter sphaeroides* 2.4.1T: effects on dimethyl sulfoxide reductase (*dor*) gene expression. *J. Bacteriol.* 180, 5612–5618. doi: 10.1128/jb.180.21.5612-5618.1998

Najafi, S. M., Willis, A. C., and Yudkin, M. D. (1995). Site of phosphorylation of SpoIIAA, the anti-anti-sigma factor for sporulation-specific σF of *Bacillus subtilis*. *J. Bacteriol.* 177, 2912–2913. doi: 10.1128/jb.177.10.2912-2913.1995

Oh, J. I., and Kaplan, S. (1999). The *cbb3* terminal oxidase of *Rhodobacter sphaeroides* 2.4.1: structural and functional implications for the regulation of spectral complex formation. *Biochemistry* 38, 2688–2696. doi: 10.1021/bi9825100

Paget, M. S. (2015). Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 5, 1245–1265. doi: 10.3390/biom5031245

Parida, B. K., Douglas, T., Nino, C., and Dhandayuthapani, S. (2005). Interactions of anti-sigma factor antagonists of *Mycobacterium tuberculosis* in the yeast two-hybrid system. *Tuberculosis* 85, 347–355. doi: 10.1016/j.tube.2005.08.001

Provvedi, R., Kocincova, D., Dona, V., Euphrasie, D., Daffe, M., Etienne, G., et al. (2008). SigF controls carotenoid pigment production and affects transformation efficiency and hydrogen peroxide sensitivity in *Mycobacterium smegmatis*. *J. Bacteriol.* 190, 7859–7863. doi: 10.1128/JB.00714-08

Robinson, M. D., Mccarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rodrigue, S., Provvedi, R., Jacques, P. E., Gaudreau, L., and Manganelli, R. (2006). The σ factors of *Mycobacterium tuberculosis*. *FEMS Microbiol. Rev.* 30, 926–941. doi: 10.1111/j.1574-6976.2006.00040.x

Sachdeva, P., Misra, R., Tyagi, A. K., and Singh, Y. (2010). The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. *FEBS J.* 277, 605–626. doi: 10.1111/j.1742-4658.2009.07479.x

Sachdeva, P., Narayan, A., Misra, R., Brahmachari, V., and Singh, Y. (2008). Loss of kinase activity in *Mycobacterium tuberculosis* multidomain protein Rv1364c. *FEBS J.* 275, 6295–6308. doi: 10.1111/j.1742-4658.2008.06753.x

Singh, A. K., Dutta, D., Singh, V., Srivastava, V., Biswas, R. K., and Singh, B. N. (2015). Characterization of *Mycobacterium smegmatis sigF* mutant and its regulon: overexpression of SigF antagonist (MSMEG_1803) in *M. smegmatis* mimics sigF mutant phenotype, loss of pigmentation, and sensitivity to oxidative stress. *Microbiology* 4, 896–916. doi: 10.1002/mbo3.288

Singh, A. K., and Singh, B. N. (2008). Conservation of Sigma F in mycobacteria and its expression in *Mycobacterium smegmatis*. *Curr. Microbiol.* 56, 574–580. doi: 10.1007/s00284-008-9126-8

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., and Jacobs, W. R. Jr. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* 4, 1911–1919. doi: 10.1111/j.1365-2958.1990.tb02040.x

van Schaik, W., and Abee, T. (2005). The role of σB in the stress response of Gram-positive bacteria - targets for food preservation and safety. *Curr. Opin. Biotechnol.* 16, 218–224. doi: 10.1016/j.copbio.2005.01.008

Vijay, K., Brody, M. S., Fredlund, E., and Price, C. W. (2000). A PP2C phosphatase containing a PAS domain is required to convey signals of energy stress to the σB transcription factor of *Bacillus subtilis*. *Mol. Microbiol.* 35, 180–188. doi: 10.1046/j.1365-2958.2000.01697.x

Voelker, U., Voelker, A., and Haldenwang, W. G. (1996). Reactivation of the *Bacillus subtilis* anti-σB antagonist, RsbV, by stress- or starvation-induced phosphatase activities. *J. Bacteriol.* 178, 5456–5463. doi: 10.1128/jb.178.18.5456-5463.1996

Waagmeester, A., Thompson, J., and Reyrat, J. M. (2005). Identifying sigma factors in *Mycobacterium smegmatis* by comparative genomic analysis. *Trends Microbiol.* 13, 505–509. doi: 10.1016/j.tim.2005.08.009

Williams, E. P., Lee, J. H., Bishai, W. R., Colantuoni, C., and Karakousis, P. C. (2007). Mycobacterium tuberculosis SigF regulates genes encoding cell wall-associated proteins and directly regulates the transcriptional regulatory gene *phoY1*. *J. Bacteriol.* 189, 4234–4242. doi: 10.1128/JB.00201-07

Yang, X., Kang, C. M., Brody, M. S., and Price, C. W. (1996). Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. *Genes Dev.* 10, 2265–2275. doi: 10.1101/gad.10.18.2265

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: 30 November 2020
doi: 10.3389/fmicb.2020.608624

Induction of the *cydAB* Operon Encoding the *bd* Quinol Oxidase Under Respiration-Inhibitory Conditions by the Major cAMP Receptor Protein MSMEG_6189 in *Mycobacterium smegmatis*

*Eon-Min Ko and Jeong-Il Oh**

Department of Integrated Biological Science, Pusan National University, Busan, South Korea

OPEN ACCESS

Edited by:
Morigen Morigen,
Inner Mongolia University, China

Reviewed by:
Greg Cook,
University of Otago, New Zealand
Joao B. Vicente,
NOVA University of Lisbon, Portugal
Meike Baumgart,
Institute of Bio- and Geosciences, IBG-1: Biotechnology, Germany

***Correspondence:**
Jeong-Il Oh
joh@pusan.ac.kr

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *21 September 2020*
Accepted: *06 November 2020*
Published: *30 November 2020*

Citation:
Ko E-M and Oh J-I (2020) Induction of the cydAB Operon Encoding the bd Quinol Oxidase Under Respiration-Inhibitory Conditions by the Major cAMP Receptor Protein MSMEG_6189 in Mycobacterium smegmatis. Front. Microbiol. 11:608624. doi: 10.3389/fmicb.2020.608624

The respiratory electron transport chain (ETC) of *Mycobacterium smegmatis* is terminated with two terminal oxidases, the aa_3 cytochrome *c* oxidase and the cytochrome *bd* quinol oxidase. The *bd* quinol oxidase with a higher binding affinity for O_2 than the aa_3 oxidase is known to play an important role in aerobic respiration under oxygen-limiting conditions. Using relevant *crp1* (*MSMEG_6189*) and *crp2* (*MSMEG_0539*) mutant strains of *M. smegmatis*, we demonstrated that Crp1 plays a predominant role in induction of the *cydAB* operon under ETC-inhibitory conditions. Two Crp-binding sequences were identified upstream of the *cydA* gene, both of which are necessary for induction of *cydAB* expression under ETC-inhibitory conditions. The intracellular level of cAMP in *M. smegmatis* was found to be increased under ETC-inhibitory conditions. The *crp2* gene was found to be negatively regulated by Crp1 and Crp2, which appears to lead to significantly low cellular abundance of Crp2 relative to Crp1 in *M. smegmatis*. Our RNA sequencing analyses suggest that in addition to the SigF partner switching system, Crp1 is involved in induction of gene expression in *M. smegmatis* exposed to ETC-inhibitory conditions.

Keywords: *aa3* cytochrome *c* oxidase, cAMP, Crp, electron transport chain, *Mycobacterium*, regulation of gene expression, respiration

INTRODUCTION

The respiratory electron transport chain (ETC) of mycobacteria consists of the membrane-associated electron carriers (menaquinone/menaquinol and cytochrome *c*) and enzymes that catalyze electron-transfer reactions with the concomitant generation of proton motive force (Cook et al., 2014). The respiratory ETC of *Mycobacterium smegmatis* is terminated with two terminal oxidases like that of *Mycobacterium tuberculosis*, the aa_3 cytochrome *c* oxidase and the cytochrome *bd* quinol oxidase, which catalyze the reduction of O_2 to water molecules using the electrons from reduced cytochrome *c* and menaquinol, respectively (Kana et al., 2001; Matsoso et al., 2005). The aa_3 cytochrome *c* oxidase, which belongs to the heme-copper superfamily of oxidases (HCOs) and serves as the major oxidase under aerobic conditions, forms a supercomplex with the

cytochrome bc_1 complex and cytochrome *c* (Matsoso et al., 2005; Megehee et al., 2006). Although the cytochrome *bd* quinol oxidase is not capable of pumping protons across the membrane during the reduction of O_2, it has a higher affinity for O_2 and is much less sensitive to inhibition by cyanide (CN^-) than the aa_3 oxidase (Puustinen et al., 1991; Cunningham et al., 1997; Kana et al., 2001; Belevich et al., 2005, 2007). It has been also demonstrated that the *bd* quinol oxidase is relatively insensitive to the physiologically relevant respiration-inhibiting molecules nitric oxide (NO) and hydrogen sulfide (H_2S) that are produced by activated or infected host immune cells and serve as inhibitors for the bacterial and mitochondrial HCOs (Mason et al., 2009; Giuffre et al., 2012; Forte et al., 2016). From these findings, it has been suggested that the *bd* quinol oxidase is involved in adaptation of pathogenic bacteria such as *M. tuberculosis* to hostile environments created by host immunity during the infection process (Giuffre et al., 2012; Forte et al., 2016). Although the *bd* quinol oxidase is not essential to *M. smegmatis* at ambient oxygen tensions, it plays an important role in aerobic respiration under hypoxic conditions, as well as under inhibitory conditions of the bcc_1-aa_3 branch (Kana et al., 2001; Matsoso et al., 2005; Aung et al., 2014; Jeong et al., 2018). The *bd* quinol oxidase of *M. smegmatis* is encoded by the *cydAB* operon (*cydA*: *MSMEG_3233*, *cydB*: *MSMEG_3232*) (Kana et al., 2001; Aung et al., 2014). Inactivation of the bcc_1-aa_3 branch and hypoxic conditions were shown to result in strong upregulation of *cydAB* expression in *M. smegmatis* (Kana et al., 2001; Matsoso et al., 2005; Aung et al., 2014; Jeong et al., 2018). It was also demonstrated that albeit moderately, expression of *cydA* was induced in *M. tuberculosis* exposed to hypoxia and NO, as well as in *M. tuberculosis*, *Mycobacterium bovis* BCG, and *Mycobacterium marinum* treated with ETC inhibitors such as Q203, bedaquiline, and clofazimine (Shi et al., 2005; Koul et al., 2014; Boot et al., 2017; Kalia et al., 2019).

cAMP is a critical secondary messenger that controls a wide variety of cellular functions in many organisms. The cAMP receptor protein (Crp) is a transcriptional regulator that controls gene expression by recognizing altered cAMP levels in prokaryotic cells. The genome of *M. smegmatis* has two genes (*crp1*: *MSMEG_6189*, *crp2*: *MSMEG_0539*) encoding the Crp paralogs that show 78% sequence identity at the amino acid level (Sharma et al., 2014; Aung et al., 2015). Sequence homology and biochemical analyses revealed that Crp1 corresponds to the Crp protein (Rv3676) occurring in *M. tuberculosis* (Bai et al., 2005; Stapleton et al., 2010; Sharma et al., 2014; Aung et al., 2015). On the basis of high sequence similarity between the helix-turn-helix (HTH) motifs of Crp1 and Crp2, together with the results of DNA-binding analyses, it was suggested that both Crp proteins recognize and bind to the same consensus sequence (TGTGA-N_6-TCACA) (Sharma et al., 2014). However, there are significant differences in amino acids forming the cAMP binding pockets of the two proteins, which accounts for their different biochemical properties such as their binding affinity for cAMP and DNA, as well as cAMP-dependent enhancement of the DNA-binding affinity (Sharma et al., 2014; Aung et al., 2015).

The expression and regulation patterns of the *bd* quinol oxidase genes in diverse bacteria are similar to those in mycobacteria. Expression of the genes is commonly induced under hypoxic or anaerobic conditions (Kana et al., 2001; Borisov et al., 2011; Small et al., 2013; Aung et al., 2014; Jeong et al., 2018; Mascolo and Bald, 2020). The regulation mechanism of the *bd* quinol oxidase genes in response to changes in oxygen availability was well-established for several bacteria. In *Escherichia coli*, expression of the *cydAB* operon is controlled by the ArcBA two-component system (TCS) and Fnr (fumarate and nitrate reduction regulatory protein) in response to changes in oxygen tensions (Cotter et al., 1990, 1997; Fu et al., 1991; Cotter and Gunsalus, 1992; Tseng et al., 1996). In *Streptomyces coelicolor* A3, expression of *cydAB* is regulated by the Rex repressor that exerts transcriptional regulation in response to changes in the cellular NADH/NAD^+ ratio (Brekasis and Paget, 2003). Using both site-directed mutagenesis of a Crp-binding sequence upstream of *cydA* and Electrophoretic mobility shift analysis (EMSA) with purified Crp2 (MSMEG_0539), expression of the *cydAB* operon in *M. smegmatis* was suggested to be positively regulated by Crp (Aung et al., 2014). However, detailed study has not been reported regarding the regulatory mechanism for the induction of *cydAB* expression by inactivation of the bcc_1-aa_3 pathway and whether two CRP paralogs differentially contribute to the regulation of *cydAB* expression. Using relevant *crp1* and *crp2* mutant strains of *M. smegmatis*, we here report the roles of Crp1 and Crp2 in upregulation of *cydAB* expression under respiration-inhibitory conditions.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Culture Conditions

The bacterial strains and plasmids used in this study are listed in **Supplementary Table 1**. *E. coli* strains were grown in Luria-Bertani (LB) medium at 37°C. *M. smegmatis* strains were grown in 7H9-glucose medium [Middlebrook 7H9 medium (Difico, Sparks, MD) supplemented with 0.2% (w/v) glucose as a carbon source and 0.02% (v/v) Tween 80 as an anticlumping agent] at 37°C. *M. smegmatis* strains were grown aerobically in a 500-ml flask filled with 100 ml of 7H9-glucose medium on a gyratory shaker (200 rpm). For treatment of *M. smegmatis* cultures with potassium cyanide (KCN), *M. smegmatis* strains were grown until the optical density at 600 nm (OD_{600}) reached 0.45–0.5. Following the addition of KCN to the cultures to a final concentration of 100 μM, the cultures were further grown for 15 min. For treatment of *M. smegmatis* cultures with sodium nitroprusside (SNP; an NO generator) or NaHS (an H_2S generator), SNP and NaHS were added to the *M. smegmatis* cultures grown to an OD_{600} of 0.45–0.5 to final concentrations of 5 mM and 200 μM, respectively, and the cultures were further grown for 30 min. The SNP-treated cultures were grown under illumination of light (100 W/m^2). Ampicillin (100 μg/ml for *E. coli*), kanamycin (50 μg/ml for *E. coli* and 15 or 30 μg/ml for *M. smegmatis*), and hygromycin (200 μg/ml for *E. coli* and 25 or 50 μg/ml for *M. smegmatis*) were added to the growth medium when required. The construction of the mutants and plasmids used in this study is described in **Supplemental Material**.

DNA Manipulation and Transformation

Standard protocols and manufacturers' instructions were followed for recombinant DNA manipulations (Sambrook and Green, 2012). Transformation of *M. smegmatis* with plasmids was carried out by electroporation as described elsewhere (Snapper et al., 1990). The primers used for PCR are listed in **Supplementary Table 2**.

Site-Directed Mutagenesis

To introduce point mutations into the Crp-binding sites (CBS1 and CBS2), PCR-based mutagenesis was performed using the Quick Change site-directed mutagenesis procedure (Stratagene, La Jolla, CA). Synthetic oligonucleotides 31 bases long containing the substituted nucleotides in the middle of their sequences were used to mutagenize the sequences. The primers used for mutagenesis are listed in **Supplementary Table 2**. Mutations were verified by DNA sequencing.

Quantitative Real-Time PCR

RNA isolation from *M. smegmatis* strains and cDNA synthesis were performed as described elsewhere (Kim et al., 2010) except for the use of a random hexamer primer (ThermoFisher, Waltham, MA) in place of the gene-specific primers in cDNA synthesis. The contamination of DNA in the isolated RNA was checked by PCR with the primers to be used in quantitative real-time PCR (qRT-PCR). To determine the transcript levels of *cydA*, *crp2*, *MSMEG_3680*, and *sigA*, qRT-PCR was performed in a 20-μl mixture containing 5 μl of the template cDNA, 15 pmol of each of two gene-specific primers, 10 μl of TB GreenTM Premix Ex TaqTM (Tli RNase Plus) (Takara, Tokyo, Japan), 0.4 μl of the ROX passive fluorescent dye, and 2.6 μl of distilled water. Thermal cycling was initiated with 1 cycle at 95°C for 2 min, followed by 40 cycles of 95°C for 5 s and 64°C for 30 s. The *sigA* gene encoding the principal sigma factor was used as a reference gene for qRT-PCR to normalize the expression levels of *cydA*, *crp2*, and *MSMEG_3680* since our RNA sequencing analyses revealed that the *sigA* gene is constitutively expressed at similar levels in the wild-type (WT), Δaa_3, Δ*crp1*, and Δ*crp2* mutant strains (**Supplementary Figure 1**). Melting curve analysis was performed for each reaction to examine whether a single PCR product was amplified during qRT-PCR. The primers used for qRT-PCR are listed in **Supplementary Table 2**.

β-Galactosidase Assay and Determination of the Protein Concentration

The β-Galactosidase activity was measured spectrophotometrically as described previously (Oh and Kaplan, 1999). The protein concentration was determined using a Bio-Rad protein assay kit (Bio-Rad, Hercules, CA) with bovine serum albumin (BSA) as a standard protein.

Western Blotting Analysis

Cell-free crude extracts were subjected to SDS-PAGE, and proteins on the gel were transferred to polyvinylidene fluoride membranes (Millipore, Burlington, MA). Western blotting analysis using an anti-2B8 antibody (Biojane, Pyeongtaek-si, South Korea) was performed as described previously (Mouncey and Kaplan, 1998). The anti-2B8 antibody was used at a dilution of 1:20,000. To detect GroEL, a mouse monoclonal antibody against Hsp65 (Santa Cruz Biotechnology, Dallas, TX; sc58170) was used at a 1:2,000 dilution. Alkaline phosphatase-conjugated anti-mouse IgG produced in rabbit (Sigma, St. Louis, MO; A4312) was used at a 1:10,000 dilution for the detection of the primary antibodies.

Protein Purification

C-terminally His_6-tagged Crp1 was expressed in the *E. coli* BL21 (DE3) strain harboring pT7-7crp1. The *E. coli* strain was cultivated aerobically to an OD_{600} of 0.4–0.6 at 37°C in LB medium containing 100 μg/ml ampicillin. Expression of the *crp1* gene was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to the cultures to a final concentration of 0.5 mM, and then cells were further grown for 4 h at 30°C. Cells were harvested from 300 ml cultures and resuspended in 10 ml of buffer A [20 mM Tris-HCl (pH 8.0) and 200 mM NaCl] containing DNase I (10 U/ml) and 10 mM $MgCl_2$. The resuspended cells were disrupted twice using a French pressure cell, and cell-free crude extracts were obtained by centrifugation twice at 20,000 × *g* for 15 min. The crude extracts were loaded into a column packed with 500 μl of the 80% (v/v) slurry of Ni-Sepharose high-performance resin (GE Healthcare, Piscataway, NJ). The resin was washed with 40 bed volumes of buffer A containing 5 mM imidazole and washed further with 20 bed volumes of buffer A containing 60 mM imidazole. His_6-tagged Crp1 was eluted from the resin with 6 bed volumes of buffer A containing 250 mM imidazole. The eluted protein was desalted using a PD-10 desalting column (GE Healthcare) equilibrated with appropriate buffer. The purity of Crp1 was checked by SDS-PAGE (**Supplementary Figure 2**).

EMSA

A 99-bp DNA fragment containing the upstream region of *cydA* and an 80-bp control DNA fragment without the Crp-binding site were used in EMSA. The 99-bp DNA fragment was amplified by PCR using pBSIIcydA as a template and the primers F_cydAEMSA and R_cydAEMSA. The 80-bp control DNA fragment was generated by PCR using pUC19 as a template and the primers F_80_EMSA and R_80_EMSA. Purified Crp1 protein was incubated with 70 fmol of the DNA fragments containing the *cydA* upstream region and 100 fmol of the control DNA fragments in binding buffer [20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 2.5 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol (DTT), 50 μg/ml BSA and 10% (v/v) glycerol] in a reaction volume of 10 μl for 20 min at 25°C. After addition of 2 μl of 6x loading buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol and 40% (w/v) sucrose), the mixtures were subjected to non-denaturing PAGE [8% (w/v) acrylamide] using 0.5x TBE buffer (41.5 mM Tris-borate and 0.5 mM EDTA, pH8.3) at 70 V for 2 h 20 min at 4°C. The gels were stained with SYBR Green staining solution for 1 h.

DNase I Footprinting Analysis

DNase I Footprinting was carried out using fluorescence (TAMRA)-labeled DNA fragments and purified Crp1.

TAMRA-labeled DNA fragments (271 bp) containing the *cydA* upstream region were generated by PCR using the primers (F_TAMRA_pUC19 and F_cydAFootR) and pUC19cydAFootR as a template. The PCR products were purified after agarose gel electrophoresis. DNA binding reaction mixtures were composed of 5 pmol of labeled DNA probes, purified Crp1 (0.15, 0.3, or 0.6 μM), 20 mM Tris-HCl (pH 8.0), 0.2 mM $MgCl_2$, 2.1 mM KCl, 0.04 mM DTT, and 11.1% (v/v) glycerol in a final volume of 190 μl. When necessary, the Crp1 protein was mixed with 200 μM cAMP, and the mixture was incubated for 10 min at 25°C prior to binding reactions for 10 min at 25°C. DNase I (Takara) was diluted in buffer containing 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM DTT, and 10% (v/v) glycerol to reach a final concentration of 0.675 mU/μl. DNase I digestion was initiated with the addition of 10 μl of diluted DNase I to the binding reaction mixtures, conducted for 1 min at 25°C, and stopped with the addition of 400 μl of stop solution [20 mM Tris-HCl (pH 8.0) and 40 mM EDTA]. DNA was purified by phenol/chloroform/isoamyl alcohol (25:24:1) extraction and isopropyl alcohol precipitation. The pellets were dissolved in TE buffer [10 mM Tris-HCl (pH 8.0) and 1 mM EDTA]. After the addition of loading buffer [95% deionized formamide, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol FF, and 5 mM EDTA (pH 8.0)], the samples were analyzed by electrophoresis on 6% (w/v) denaturing polyacrylamide gels with 7 M urea in 0.8x Tris-taurine-EDTA (TTE) buffer using an ABI PRISM 377 DNA sequencer (Applied Biosystems, Foster City, CA). Reference sequencing was performed by using a Thermo sequenase dye primer manual cycle sequencing kit (ThermoFisher) with the primer F_TAMRA_pUC19 and the template plasmid pUC19cydAFootR.

Determination of the Intracellular cAMP Concentration

M. smegmatis cells corresponding to 1 ml of cultures at OD_{600} of 0.4 were harvested. Cell pellets were resuspended in 1 ml of 0.1 M HCl and then incubated for 10 min. Cells were disrupted once by using a Fastprep 120 beadbeater (ThermoFisher) at 5.0 m/s for 45 s. Cell-free supernatants were obtained by centrifugation at 20,000 × *g* for 10 min. The concentration of cAMP in the prepared supernatants was determined by using a DetectX Direct Cyclic AMP Enzyme Immunoassay kit (Arbor Assays, Ann Arbor, MI) and a microplate reader (Bio-Rad) following the manufacture's instruction.

RNA Sequencing and Gene Expression Profiling

Three biological replicate cultures of the WT and Δ*crp1* strains were grown aerobically to an OD_{600} of 2.0–2.1 (late exponential phase). Total RNA of each culture was isolated as described previously (Kim et al., 2010). rRNA was removed from the isolated total RNA using a Ribo-Zero rRNA Removal Kit (Bacteria) (Illumina, San Diego, CA). The RNA sequencing libraries were created using a TruSeq RNA Sample Prep Kit v2 (Illumina) with the standard low-throughput protocol. Sequencing of the six libraries was conducted on an Illumina HiSeq 4000 platform at Macrogen Inc. (Seoul, South Korea) using the Hiseq 3000–4000 sequencing protocol and TruSeq 3000–4000 SBS Kit v3 reagent (Illumina). Paired-end reads (101 bp) were then mapped to the reference genome sequence of *M. smegmatis* mc^2155 (GCF_000015005.1_ASM1500v1) with the program Bowtie 1.1.2 using default settings. Summarized statistics of RNA sequencing alignment are listed in **Supplementary Table 3**. The differentially expressed genes (DEGs) were subsequently identified pair-wise by the edgeR package in R language (Robinson et al., 2010). In this analysis, the genes with $p < 0.05$ and $|FC| > 1.5$ were regarded as DEGs. The RNA sequencing data have been deposited in NCBI's Gene Expression Omnibus and are accessible through the GEO Series accession number GSE158137.

RESULTS

Induction of *cydA* Expression Under Respiration-Inhibitory Conditions Is Independent of SigF

Under aerobic culture conditions, the inactivation of the aa_3 cytochrome *c* oxidase in *M. smegmatis* by mutation was previously shown to lead to an overall reduction in the respiration rate by ∼50% and a significant increase in expression of the *cydAB* operon encoding the *bd* quinol oxidase (Jeong et al., 2018). Furthermore, we found that the genes, which belong to the SigF (an alternative sigma factor) regulon, are strongly upregulated in a mutant strain of *M. smegmatis* lacking the aa_3 oxidase (Oh et al., 2020). To examine whether the induction of *cydAB* expression in the Δaa_3 mutant with a deletion in *ctaC* encoding subunit III of the aa_3 cytochrome *c* oxidase is a result of SigF activation, we determined expression of *cydA* in the WT, Δaa_3 mutant, and Δaa_3Δ*sigF* double mutant strains using the strains harboring the *cydA*::*lacZ* translational fusion plasmid pNCIIcydA. In good agreement with the previous report (Matsoso et al., 2005; Jeong et al., 2018), the expression level of *cydA* was increased in the Δaa_3 mutant by ∼107-fold relative to the WT strain (**Figure 1**). The expression level of *cydA* was not decreased in the Δaa_3Δ*sigF* mutant compared to the Δaa_3 mutant, indicating that the *cydAB* operon does not belong to the SigF regulon, and that the strong upregulation of the *cydAB* operon under respiration-inhibitory conditions is not caused by the activation of SigF.

MSMEG_6189 Is the Major Crp in *M. smegmatis*

Previously it has been reported that Crp is involved in the positive regulation and hypoxic induction of the *cydAB* operon in *M. smegmatis* (Aung et al., 2014). However, it remained unanswered whether two Crp paralogs (Crp1: MSMEG_6189, Crp2: MSMEG_0539) play a distinct role in induction of the *cydAB* operon under respiration-inhibitory conditions. To examine the involvement of Crp1 and Crp2 in the regulation of *cydA* expression, we determined the expression level of *cydA* in the Δ*crp1* and Δ*crp2* mutant strains. Since we failed to obtain a Δaa_3Δ*crp1* double mutant strain, treatment of *M. smegmatis* cultures with KCN, an inhibitor of aa_3 cytochrome *c* oxidase, was applied to mimic the Δaa_3 mutant conditions.

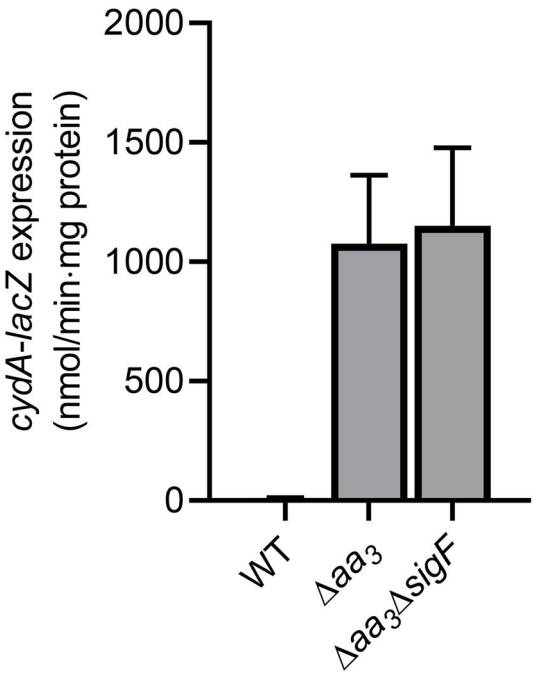


FIGURE 1 | Expression of the *cydAB* operon in the WT, Δaa_3, and $\Delta aa_3\Delta sigF$ mutant strains of *M. smegmatis*. The WT, Δaa_3, and $\Delta aa_3\Delta sigF$ mutant strains of *M. smegmatis* harboring the *cydA::lacZ* translational fusion plasmid pNCIIcydA were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. Cell crude extracts were used to determined β-galactosidase activity. All values provided are the averages of the results from three biological replicates. The error bars indicate the standard deviations.

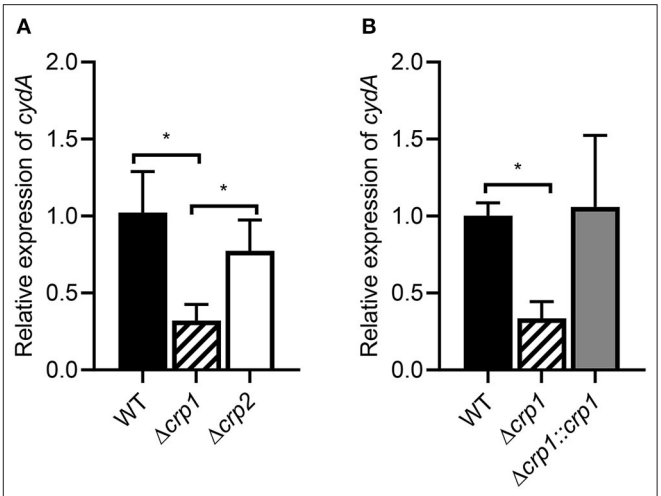


FIGURE 2 | Expression of the *cydAB* operon in the WT, Δ*crp1*, and Δ*crp2* mutant strains of *M. smegmatis*. **(A)** The relative transcript level of *cydA* in the WT, Δ*crp1*, and Δ*crp2* mutant strains. **(B)** Complementation of the Δ*crp1* mutant. For complementation of the Δ*crp1* mutant, pMV306crp (a pMV306-derived plasmid carrying the intact *crp1* gene and its own promoter) was introduced into the mutant. As control strains, the WT and Δ*crp1* strains with the empty vector pMV306 were used in the experiment. All the strains were grown aerobically to an OD_{600} of 0.45–0.5 and treated with 100 μM KCN for 15 min. The expression level of *cydA* was quantitatively determined by qRT-PCR and normalized to *sigA* (the gene for the principal sigma factor) expression. The expression level of *cydA* in the KCN-treated WT strain is set at 1, and the relative values are expressed for the mutant strains. All values provide are the average of the results from three biological replicates. The error bars indicate the standard deviations. Statistical significance was determined by two-tailed Student's *t*-test. $^*p < 0.05$.

Effects of KCN treatment on *cydA* expression were quantitatively determined by qRT-PCR in the WT, Δ*crp1*, and Δ*crp2* mutant strains that were grown aerobically. We could not use the *cydA*::*lacZ* transcriptional fusion pNCIIcydA to determine the expression level of *cydA* in the *M. smegmatis* strains treated with KCN, since the addition of KCN interfered with expression or assay of β-galactosidase for unknown reasons. The treatment of the WT strain with 100 μM KCN led to induction of *cydA* expression by 388-fold compared to the untreated WT control strain (data not shown). When the aerobically grown WT, Δ*crp1*, and Δ*crp2* mutant strains were treated with KCN, the Δ*crp1* mutant showed only 30% of *cydA* expression observed in the WT strain, while the expression level of *cydA* was only slightly reduced in the Δ*crp2* mutant relative to the WT strain (**Figure 2A**). This result indicates that Crp1 plays a predominant role in induction of *cydA* expression under respiration-inhibitory conditions. The ectopic expression of the intact *crp1* gene in the Δ*crp1* mutant using pMV306crp restored the expression level of *cydA* to that in the WT strain with the empty integration vector pMV306 (**Figure 2B**), confirming that the reduction of *cydA* expression in the Δ*crp1* mutant is the result of *crp1* inactivation. To confirm the Crp1-mediated induction of *cydA* expression by inactivation of the aa_3 cytochrome *c* oxidase, effects of NO and H_2S, which are the physiologically relevant inhibitors of the aa_3 oxidase, on *cydA* expression were assessed in the WT and Δ*crp1* mutant strains (**Supplementary Figure 3**). As in the Δaa_3 mutant and the WT strain treated with KCN, expression of *cydA* was significantly (325-fold) increased in the WT strain treated with 5 mM SNP (NO generator) relative to that in the SNP-untreated control WT strain. Induction of *cydA* expression was significantly compromised in the SNP-treated Δ*crp1* mutant compared to the WT strain treated with SNP. Similarly, expression of *cydA* was increased by 114-fold in the WT strain treated with 200 μM NaHS (H_2S generator) relative to that in the NaHS-untreated control WT strain. The expression level of *cydA* in the NaHS-treated Δ*crp1* mutant was shown to amount to ∼47% of that observed for the NaHS-treated WT strain. Using KCN, we examined the aerobic growth of the WT, Δ*crp1*, and Δ*crp2* mutant strains when the aa_3 oxidase is inhibited (**Supplementary Figure 4**). The WT, Δ*crp1*, and Δ*crp2* strains grew at the similar rate in 7H9-glucose medium in the absence of KCN. In contrast, the growth of the Δ*crp1* mutant was severely compromised in the presence of 100 μM KCN, and that of the Δ*crp2* mutant was moderately affected compared to the WT strain. The high susceptibility of Δ*crp1* mutant to SNP regarding inhibition of aerobic growth was also demonstrated previously (Lee et al., 2014). Altogether, these results reinforce that inhibition of the respiratory ETC by inactivation of the aa_3 oxidase leads to Crp1-mediated induction of the *cydAB* operon in *M. smegmatis*.

The expression levels of *crp1* and *crp2* in *M. smegmatis* were extrapolated from the reads per kilo base pair per million mapped reads (RPKM) values obtained from RNA sequencing analysis on the WT strain of *M. smegmatis* that was aerobically grown to an OD_{600} of 0.45–0.5 (Lee et al., 2018). The RPKM values of *crp1* and *crp2* suggested that the transcript level of *crp1* is 8-fold higher than that of *crp2* in the WT strain (**Figure 3A**). To assess whether the estimated transcript levels of *crp1* and *crp2* correlate with their cellular protein levels, Western blotting analysis was performed using the Δ*crp1* and Δ*crp2* mutant strains expressing the C-terminally 2B8 epitope-tagged Crp1 and Crp2 proteins, respectively. For the construction of the strains, the *crp1* and *crp2* genes with the upstream regions encompassing their own promoters and regulatory sequences were cloned into the integration vector pMV306, and the resulting pMV306crp1_2B8 and pMV306crp2_2B8 plasmids were integrated into the chromosomes of the Δ*crp1* and Δ*crp2* mutant strains, respectively. Western blotting analysis revealed that Crp1 was expressed at much higher levels than Crp2 in *M. smegmatis* grown to various stages of exponential growth phase, while the protein level of GroEL, which was used as a loading control, was relatively constant in both strains grown to various stages of exponential growth phase (**Figure 3B**). This finding that Crp1 is the predominantly expressed Crp in *M. smegmatis* might explain the dominant role of Crp1 in the positive regulation of the *cydAB* operon.

Identification of Two Crp-Binding Sites in the Upstream Region of *cydA* and Their Roles in *cydA* Expression

To identify the Crp-binding sequence(s) in the upstream region of *cydA*, DNase I footprinting analysis was performed with purified Crp1 and 271-bp TAMRA-labeled DNA fragments containing the *cydA* promoter region. Since Crp1 was shown to play a predominant role in the regulation of the *cydAB* operon and both Crp paralogs were suggested to bind to the same DNA sequence (Sharma et al., 2014), we used only Crp1 for DNA-binding analyses. As shown in **Figure 4A**, binding of Crp1 protected DNA from DNase I cleavage at positions between −52 and −102 with regard to the transcription start point (TSP) of *cydA*. The protected region contains two Crp-binding sequences (CBS1 and CBS2) that are similar to the known Crp-binding consensus sequence (TGTGA-N_6-TCACA) (**Figure 4B**). CBS1 is located at positions between −86 and −101 relative to the TSP, and CBS2 was located between −58 and −73. The addition of cAMP to the reaction mixtures resulted in wider and more clearly protected windows for both CBS1 and CBS2, indicating that cAMP enhances the binding of Crp1 to both Crp-binding sites. At low Crp1 concentrations, CBS2 was protected better than CBS1 in the presence and absence of cAMP, indicating that Crp1 binds better to CBS2 than it does to CBS1. To confirm the result of DNase I footprinting, EMSAs were performed with purified Crp1 and 99-bp DNA fragments encompassing the *cydA* upstream sequence (**Figure 5**). An 80-bp DNA fragment without the Crp-binding sequence (non-specific DNA) was used as a negative control DNA. The binding ability of Crp1 for the DNA fragments was estimated from the band intensity of the unretarded free DNA. As shown in **Figure 5A**, the binding of Crp1 to the *cydA* regulatory region was enhanced in the presence of 200 μM cAMP, which is consistent with the DNase I footprinting result. In order to determine to what extent mutations of each Crp-binding site affect the binding of Crp1 to the *cydA* regulatory region, we performed EMSAs using purified Crp1 and 99-bp DNA fragments containing the WT or mutated Crp-binding sites (M1, M2, and M3) in the presence of cAMP. As shown in **Figure 5B**, the M1 and M2 DNA fragments containing mutations within CBS1 and CBS2, respectively, were retarded by Crp1 to a lesser extent than the WT DNA fragment as judged by the levels of free DNA. Especially the M2 mutation significantly affected the binding of Crp1 to the *cydA* regulatory region. Mutations of both CBS1 and CBS2 virtually abolished the binding of Crp1 to the M3 DNA fragment. Taken together, the EMSA and DNase I footprinting results indicate that Crp1 can bind to both CBS1 and CBS2 with a higher binding affinity for CBS2.

To investigate the role of CBS1 and CBS2 in induction of *cydA* expression under respiration-inhibitory conditions, a series of *cydA*::*lacZ* translational fusions with 5'-serial deletions of the *cydA* upstream region [pNCIISD1 (SD1), pNCIISD2 (SD2), pNCIISD3 (SD3), and pNCIISD4 (SD4)] were used to determine expression of *cydA* in the Δaa_3 mutant of *M. smegmatis* (**Figure 6A**). Consistent with the result presented in **Figure 1**, expression of *cydA* was strongly induced in the Δaa_3 mutant carrying pNCIIcydA (Con) relative to the control WT strain with pNCIIcydA. The 5'-deletion up to the position−163 relative to the TSP (SD1) did not affect *cydA* expression in the Δaa_3 mutant, and the additional 20-bp deletion of the *cydA* upstream region (SD2) led to a ∼35% decrease in *cydA* expression in the Δaa_3 mutant. When the *cydA* upstream region was further deleted to remove the 5′ half of CBS1 (SD3), the induction of *cydA* expression in the Δaa_3 mutant was almost abolished. The deletion of both CBS1 and CBS2 (SD4) resulted in complete abolishment of *cydA* expression in the Δaa_3 mutant. To more precisely assess the role of CBS1 and CBS2 in the regulation of *cydA* expression, point mutations were introduced into CBS1, CBS2, or both CBS1 and CBS2 on pNCIIcydA, and the expression level of *cydA* was measured using the Δaa_3 mutants carrying the corresponding *cydA*::*lacZ* translational fusion plasmids [pNCIIcydA (Con), pNCIIM1 (M1), pNCIIM2 (M2), and pNCIIM3 (M3)] (**Figure 6B**). As observed for the Δaa_3 mutants with M1 and M2, the introduction of point mutations into CBS1 (TGTCG-N_6-TCACC to TCCCG-N_6-TTTTC) and CBS2 (CGTGA-N_6-CCACC to CCCCA -N_6-CTTTC) resulted in only 15 and 23% of *cydA* expression observed in the Δaa_3 mutant with pNCIIcydA, respectively. The effect of mutations within both CBS1 and CBS2 appeared to be cumulative as judged by the expression level of *cydA* in the Δaa_3 strain with M3. Altogether, the results presented in **Figure 6** suggest that both CBS1 and CBS2 are required for induction of *cydA* expression under respiration-inhibitory conditions. It is noteworthy that expression of *cydA* was still 7.3-fold induced in the Δaa_3 mutant with M3 compared to the WT strain with pNCIIcydA (**Figure 6B**). From this finding together with the observed reduction in *cydA* expression from SD2 relative to

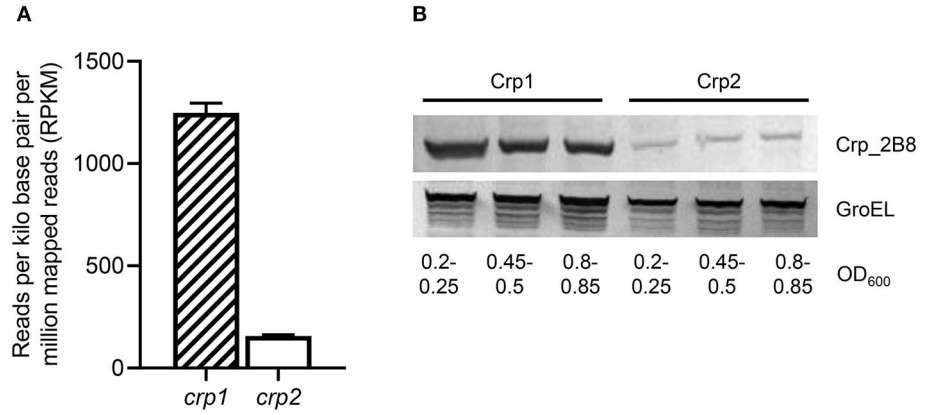


FIGURE 3 | Expression levels of two Crp paralogs in *M. smegmatis*. **(A)** Transcript levels of *crp1* and *crp2* were extrapolated from the RPKM values obtained from RNA sequencing analysis on the WT strain of *M. smegmatis* (Lee et al., 2018). **(B)** Expressed protein levels of Crp1 and Crp2. To determine the protein levels of Crp1 and Crp2 expressed in *M. smegmatis*, the Δ*crp1* and Δ*crp2* mutant strains complemented with pMV306crp1_2B8 and pMV306crp2_2B8, respectively, were grown aerobically to the indicated OD_{600} in 7H9-glucose medium, and their crude extracts (30 μg for detection of Crp1 and Crp2; 5 μg for GroEL detection) were subjected to Western blotting analyses using a 2B8 antibody and an Hsp65 antibody to detect the C-terminally 2B8-tagged Crp and GroEL, respectively. The protein level of GroEL was determined as a loading control.

that from SD1 (**Figure 6A**), we cannot rule out the possibility that there might be an additional *cis*-acting element within or overlapping the 20-bp sequence between SD1 and SD2, which is implicated in the induction of *cydA* expression in *M. smegmatis* under respiration-inhibitory conditions.

SigF and Crp1 Are the Major Contributors in Induction of Gene Expression in *M. smegmatis* Under Respiration-Inhibitory Conditions

Based on the findings that the *cydAB* operon is strongly upregulated in the Δaa_3 mutant in a Crp-dependent way and does not require SigF for its transcription, we searched for the genes that are regulated in a similar way as the *cydAB* operon to further exemplify the Crp-mediated induction of gene expression under respiration-inhibitory conditions. As shown in **Figure 7A**, our comparative RNA sequencing analysis on the WT and Δaa_3 mutant strains revealed 103 DEGs whose expression is increased in the Δaa_3 mutant by more than 4-fold with a $p < 0.05$ relative to the WT strain. Sixty-one genes among the 103 DEGs were found to belong to the known SigF regulon (Singh et al., 2015). RNA sequencing analysis on the WT and Δ*crp1* mutant strains showed that among the remaining 42 genes, 25 genes was found to be less expressed in the Δ*crp1* mutant by more than 1.5-fold with a $p < 0.05$ relative to the WT strain (**Supplementary Table 4**). Due to the presence of intact *crp2* in the Δ*crp1* mutant, the less stringent cutoff value (1.5) was applied to select the genes that are positively regulated by Crp1. Among the identified 25 DEGs showing the similar expression patterns as *cydAB* in terms of the strong induction of gene expression in the Δaa_3 mutant in a SigF-independent way and a decrease in gene expression in the Δ*crp1* mutant, *MSMEG_3680* annotated as a hypothetical protein gene is such a gene that has two putative Crp-binding sites in its upstream region which are arranged similarly as those found upstream of *cydA* (**Figure 7B**). Using qRT-PCR, the expression pattern of *MSMEG_3680* was verified in the WT and Δ*crp1* mutant strains that were grown aerobically with or without treatment of KCN. In the same way as *cydA* expression, expression of *MSMEG_3680* was strongly induced in the WT strain by treatment of KCN, and induction of its expression by KCN was significantly compromised in the Δ*crp1* mutant. Altogether, the results from RNA sequencing analyses suggest the possibility that Crp1 is likely involved in induction of gene expression in *M. smegmatis* under respiration-inhibitory conditions.

Since Crp is a regulatory protein that regulates gene expression in response to changes in cAMP levels, we determined whether the level of cAMP in *M. smegmatis* is changed in response to respiration inhibition. As shown in **Figure 8A**, the intracellular level of cAMP in the Δaa_3 mutant grown aerobically was found to be 3.2-fold higher than that detected in the WT strain grown under the same conditions. The level of cAMP was even more (14-fold) increased, when the aerobically grown WT strain of *M. smegmatis* was treated with 100 μM KCN for 15 min relative to the KCN-untreated control WT strain (**Figure 8B**). The respiratory ETC in the WT strain subjected to a short period (15 min) of KCN treatment is assumed to be more inhibited, at least temporarily before the induced synthesis of the *bd* quinol oxidase, than that in the Δaa_3 mutant in which expression of the *cydAB* operon is constitutively induced. The higher cAMP level in the KCN-treated WT strain than in the Δaa_3 mutant likely results from the more severe inhibition of the ETC in the KCN-treated WT strain. These results suggest that inhibition of the respiratory ETC entails an increase in intracellular cAMP levels, which appears to contribute to Crp-mediated induction of *cydA* expression in *M. smegmatis* under respiration-inhibitory conditions.

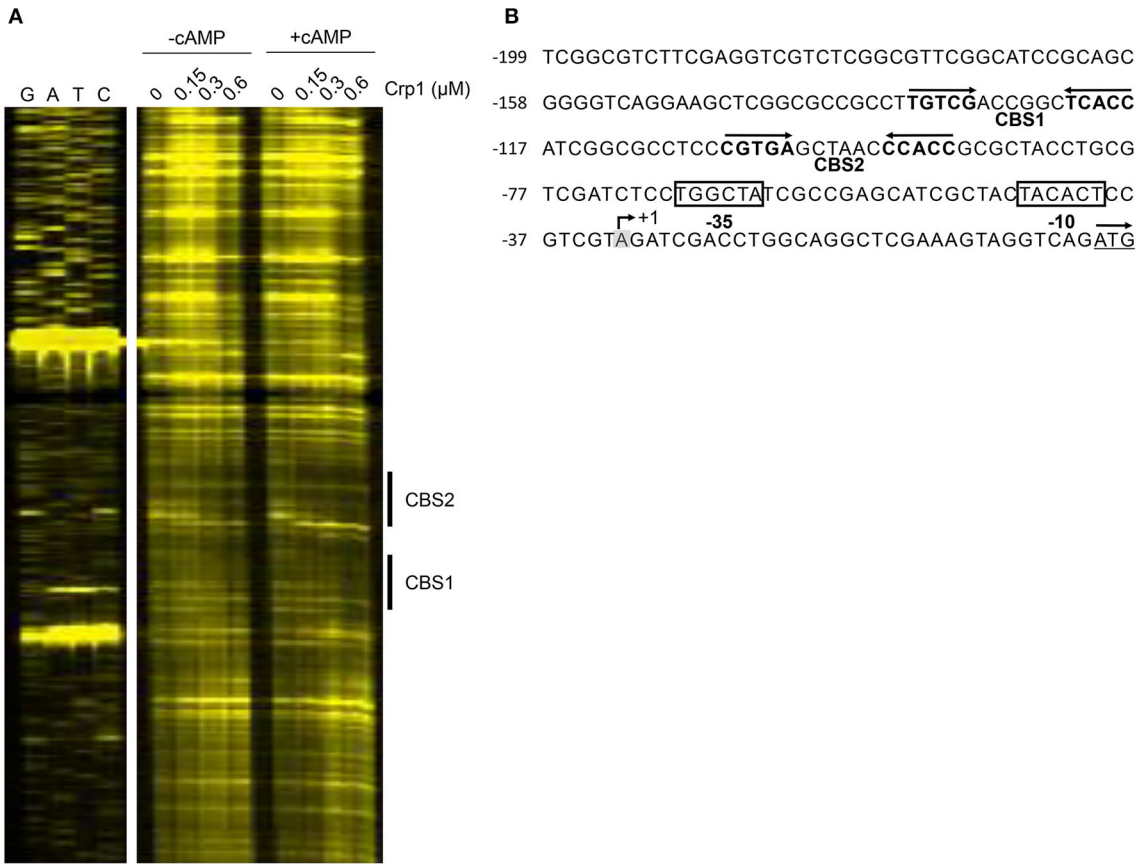


FIGURE 4 | Identification of the Crp-binding sites in the upstream region of *cydA*. **(A)** DNase I Footprinting analysis of the *cydA* upstream region bound by Crp1. The DNA fragments containing the non-coding strands labeled with TAMRA at their 5′ ends were incubated with increasing concentrations of purified Crp1 in the absence or presence of 200 μM cAMP and then subjected to DNase I footprinting reactions. The concentrations of Crp1 used are given above the lanes. The Crp-binding sites (CBS1 and CBS2) are marked by the thick lines on the right. Lanes G, A, T, and C represent the sequence ladders. **(B)** The upstream sequence of the *cydA* gene. The Crp-binding sites (CBS1 and CBS2) are marked by the two head-facing arrows above the sequence. The nucleotide reported to be the TSP (+1) of *cydA* is shaded in gray, and the putative −10 and −35 elements extrapolated from +1 are boxed (Aung et al., 2014). The start codon of *cydA* is underlined and the arrow above it indicates the transcriptional direction. The numbers on the left of the sequences indicate the positions of the leftmost nucleotides relative to the *cydA* gene. The concentration of Crp1 refers to that of the Crp1 monomer.

RNA sequencing analysis showed that expression of *crp2* was increased by 3.4-fold in the Δ*crp1* mutant relative to the WT strain, when both strains were grown aerobically (data not shown). We found that a Crp-binding consensus sequence overlaps with the TSP of *crp2* (**Figure 9A**). Using qRT-PCR, we determined the expression level of *crp2* in the WT, Δ*crp1*, and Δ*crp2* mutant strains after the strains had been treated with KCN (cAMP-increasing conditions). As shown in **Figure 9A**, expression of *crp2* was increased in the Δ*crp1* and Δ*crp2* mutant strains by 4.1- and 1.5-fold, respectively, when compared to that in the WT strain. This result indicates that *crp2* is under the negative control of Crp1 and Crp2, and that Crp1 predominates in the negative regulation of *crp2* under respiration-inhibitory conditions. We also determined the expression level of *crp2* in the WT and Δaa_3 mutant strains that were grown aerobically (**Figure 9B**). The expression level of *crp2* was reduced by 50% in the Δaa_3 mutant compared to that in the WT strain, suggesting that expression of *crp2* in *M. smegmatis* is even more repressed under respiration-inhibitory conditions probably via the negative regulation by Crp1.

DISCUSSION

The prokaryotic cytochrome *bd* quinol oxidase is structurally and functionally distinct from the HCOs including the aa_3 cytochrome *c* oxidase (Belevich et al., 2005; Megehee et al., 2006; Borisov et al., 2011). The *bd* quinol oxidase has a higher affinity for O_2 than HCOs (Puustinen et al., 1991; Belevich et al., 2005, 2007) and has been suggested to possess the additional activities that decompose H_2O_2 and peroxynitrite which is a product of the spontaneous reaction of superoxide radical with NO (Lindqvist et al., 2000; Borisov et al., 2004, 2013; Mason et al., 2009; Giuffre et al., 2014). These properties, together with insensitivity of the *bd* quinol oxidase to the physiologically relevant HCO inhibitors NO and H_2S, make the *bd* quinol oxidase beneficial for bacteria to adapt to and survive in hostile

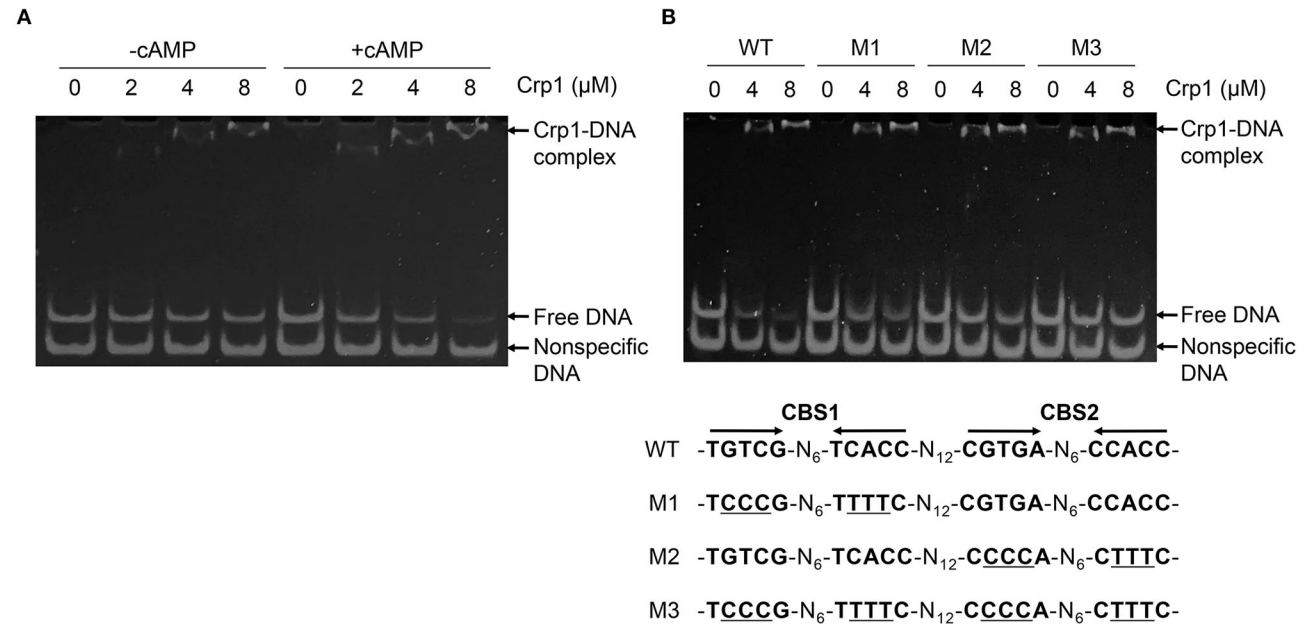


FIGURE 5 | Binding of Crp1 to the *cydA* control region. **(A)** EMSA showing the effect of cAMP on the binding of Crp1 to the upstream region of *cydA*. The mixtures of 99-bp DNA fragments (70 fmol) containing two Crp-binding sites (CBS1 and CBS2) upstream of *cydA* and 80-bp non-specific DNA fragments without the Crp-binding site (100 fmol) were incubated with increasing concentrations of purified Crp1 in the presence and absence of 200 μM cAMP. **(B)** 99-bp DNA fragments (70 fmol) containing either WT or mutated Crp-binding sites (M1, M2, and M3), as well as 80-bp non-specific DNA fragments without the Crp-binding site (100 fmol), were mixed with increasing concentrations of purified Crp1 in the presence of 200 μM cAMP. The Crp1-DNA reaction mixtures were subjected to native PAGE. The concentrations of Crp1 used in EMSA are given above the lanes. The concentration of Crp1 refers to that of the Crp1 monomer.

host conditions such as hypoxia and conditions exposed to reactive oxygen species, NO, and H_2S (Kana et al., 2001; Matsoso et al., 2005; Giuffre et al., 2012; Small et al., 2013; Forte et al., 2016; Rahman et al., 2020; Saini et al., 2020). Accordingly, expression of the *bd* quinol oxidase genes is regulated to be induced under respiration-inhibitory conditions such as hypoxia and in the presence of NO and H_2S that act as inhibitors of HCOs (Kana et al., 2001; Shi et al., 2005; Giuffre et al., 2012; Small et al., 2013; Aung et al., 2014; Jones-Carson et al., 2016; Jeong et al., 2018). The known regulatory systems, which are involved in the induction of gene expression under hypoxic or anaerobic conditions, generally sense either the molecular oxygen itself or cellular changes caused by the inhibition of the respiratory ETC. The DevSR TCS in *M. smegmatis* senses directly the molecular oxygen through the heme *b* in the DevS histidine kinase and upregulates its target genes under hypoxic and anaerobic conditions (Mayuri et al., 2002; O'Toole et al., 2003; Lee et al., 2008; Kim et al., 2010). The Fnr regulators found in many bacteria also sense O_2 levels through their O_2-labile [4Fe-4S] center and regulate gene expression in response to changes in oxygen availability (Kiley and Beinert, 1998). The inhibition of the respiratory ETC by O_2 depletion or the inactivation (or inhibition) of the ETC components such as the terminal oxidases is expected to entail changes in the redox state of electron carriers to a more reduced state. The ArcB histidine kinase of the ArcBA TCS in *E. coli* and the RegB histidine kinase of the RegBA TCS in *Rhodobacter capsulatus* have been suggested to be activated by increased ubiquinol of the quinol/quinone pool under anaerobic conditions (Georgellis et al., 2001; Malpica et al., 2004; Swem et al., 2006; Bekker et al., 2010; Wu and Bauer, 2010). The Rex repressors in Gram-positive bacteria regulate gene expression in response to changes in the redox poise of the NADH/NAD^+ pool (Brekasis and Paget, 2003; Schau et al., 2004; Larsson et al., 2005; Gyan et al., 2006). Under respiration-inhibitory conditions, the increased ratio of NADH to NAD^+ inhibits the DNA-binding activity of Rex via the binding of NADH to the conserved Rossman fold of Rex (Brekasis and Paget, 2003). Since expression of the *cydAB* operon is strongly induced in the Δaa_3 mutant grown under ambient air conditions, the regulatory system responsible for hypoxic induction of the *cydAB* operon is assumed not to be the regulator that can directly sense the molecular oxygen, but to be the regulatory system that controls gene expression through reflecting the cellular redox state or the functionality of the respiratory ETC. In this respect, the DevSR TCS appears not to be the regulatory system that is responsible for induction of *cydA* in *M. smegmatis* under respiration-inhibitory conditions. Consistent with our assumption, expression of *cydA* in the Δ*devR* (*MSMEG_5244*) mutant of *M. smegmatis* was induced to a similar level as that in the WT strain, when growth of both strains was shifted from aerobic to hypoxic conditions (data not shown). This observation is further supported by the previous report that the *cydAB* operon is not included in 49 genes identified as the DevR regulon (Berney et al., 2014). Our search for the Rex, Fnr, RegB, and

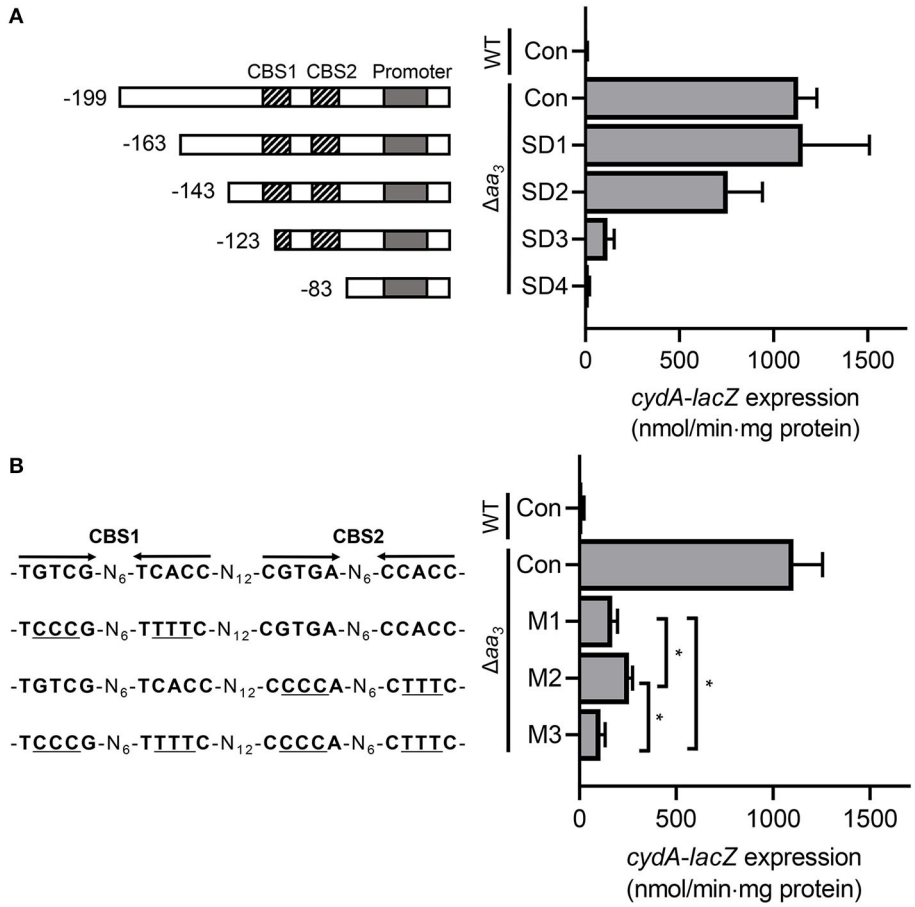


FIGURE 6 | Effects of deletions and mutations in the Crp-binding sites (CBS1 and CBS2) on expression of the *cydAB* operon. **(A)** The *cydA* promoter activity was determined using the pNCIIcydA-derived *cydA::lacZ* translational fusions containing serial deletions of the *cydA* upstream region (pNCIISD1, pNCIISD2, pNCIISD3, and pNCIISD4). **(B)** The *cydA* promoter activity was determined using several pNCIIcydA derivatives containing mutations in the Crp-binding sites (pNCIIM1, pNCIIM2, and pNCIIM3). The mutations in the Crp-binding sites of the *cydA::lacZ* translational fusions are presented on the left. The mutations within CBS1 and CBS2 are indicated by underlines. The WT and Δaa_3 mutant strains of *M. smegmatis* harboring the *cydA::lacZ* translational fusion plasmids were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. Cell crude extracts were used to determined β-galactosidase activity. All values provided are the averages of the results from three biological replicates. The error bars indicate the standard deviations. Statistical significance was determined by two-tailed Student's *t*-test. *$p < 0.05$. Con, pNCIIcydA; SD1-SD4, pNCIISD1- pNCIISD4; M1-M3, pNCIIM1-pNCIIM3.

ArcB homologs in *M. smegmatis* revealed that their homologous genes are not present in the *M. smegmatis* genome. Our RNA sequencing analysis on Δaa_3 mutant strain of *M. smegmatis* grown aerobically revealed that 61% of the strongly upregulated DEGs (FC $\geq$ 4 and $p < 0.05$) in the Δaa_3 mutant relative to the WT belong to the SigF regulon, indicating that it is the SigF partner switching system that plays a predominant role in strong upregulation of gene expression in *M. smegmatis* under ETC-inhibitory conditions. Since transcription of the *cydAB* operon was shown to be independent of SigF in *M. smegmatis* (**Figure 1**), there must be (a) regulatory system(s) other than SigF that is (are) responsible for strong upregulation of *cydAB* under respiration-inhibitory conditions. Based on a previous report (Aung et al., 2014) and our results clearly showing that the *cydAB* operon is under the positive regulation of Crp in *M. smegmatis*, we assumed that Crp is involved in upregulation of *cydAB* expression in *M. smegmatis* under respiration-inhibitory conditions. This assumption is supported by our RNA sequencing result that considerable fractions of the genes, which are more than 4-fold upregulated in the Δaa_3 mutant in a SigF-independent way, were shown to be statistically downregulated in the Δ*crp1* mutant relative to the WT strain.

The Crp proteins are homodimeric global transcription factors that regulate expression of many genes involved in diverse metabolic and cellular processes in prokaryotes, including carbon utilization, respiration, virulence, cell cycle control, reactivation of non-replicating dormant cells, and stress responses, etc., (Utsumi et al., 1989; Rickman et al., 2005; Shimada et al., 2011; Aung et al., 2014; Green et al., 2014; Heroven and Dersch, 2014). The promoter of *cydA* possesses two Crp-binding sites that are

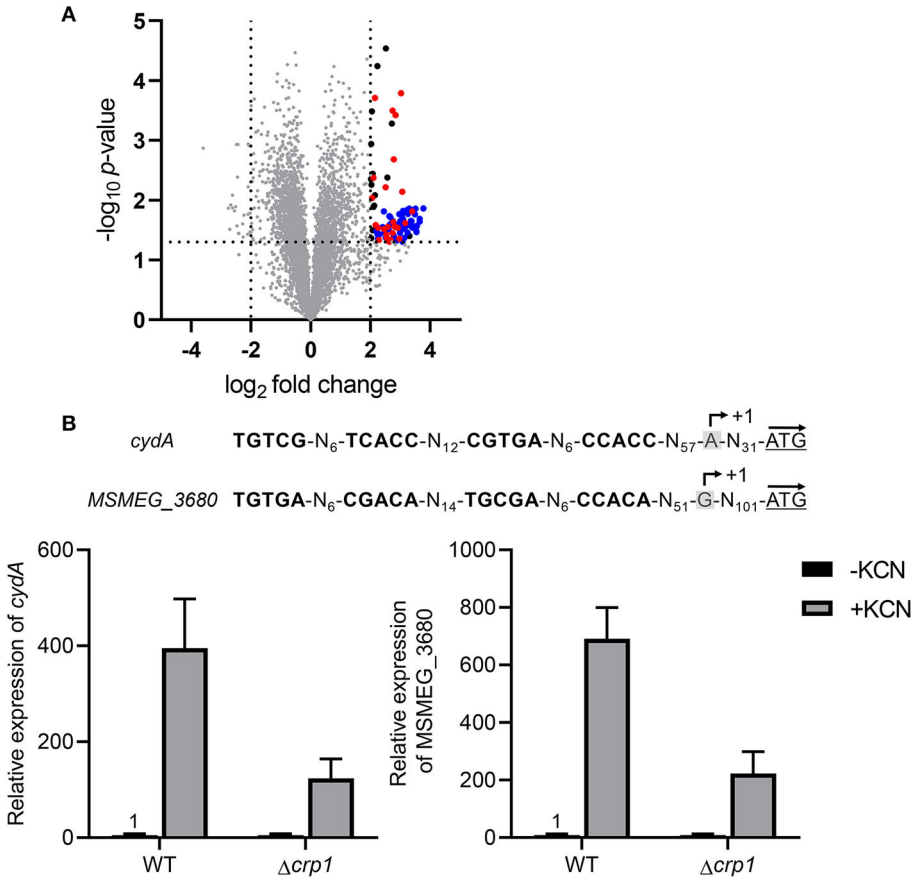


FIGURE 7 | Overlap of the SigF and Crp1 regulons with the genes induced in the Δ*aa*$_3$ mutant strain of *M. smegmatis*. **(A)** Volcano plot showing the DEGs in the Δ*aa*$_3$ mutant strain relative to the WT strain. RNA sequencing was performed using RNA prepared from three independent replicate cultures of the WT and Δ*aa*$_3$ mutant strains grown aerobically in 7H9-glucose medium to an OD_{600} of 0.45–0.5. The x-axis displays the $\log_2$ fold change of gene expression ($\log_2$FC) in the Δ*aa*$_3$ mutant relative to the WT strain, and the y-axis represents $-\log_{10}$ *p*-value. The horizontal dotted line on the graph indicates the border line indicating the *p*-value of 0.05, and the vertical dotted lines indicate the border lines indicating the $\log_2$FC values of −2 and +2. One hundred and three genes, whose expression is increased by more than $\log_2$FC ≥ 2 with $p < 0.05$, are depicted by black filled circles. Among the 103 DEGs, the genes belonging to the SigF regulon are denoted by blue filled circles (Singh et al., 2015), and the genes, whose expression is reduced by FC ≥ 1.5 with $p < 0.05$ in the Δ*crp1* mutant strain relative to the WT strain, are indicated by red filled circles. **(B)** The nucleotide sequences and locations of the Crp-binding sites in the upstream regions of *cydA* and *MSMEG_3680* and the expression levels of the genes in the WT and Δ*crp1* mutant strains of *M. smegmatis*. The Crp-binding sites are shown in bold. The previously reported TSPs of *cydA* and *MSMEG_3680* are denoted by +1 (Aung et al., 2014; Martini et al., 2019). The start codons of *cydA* and *MSMEG_3680* are underlined, and the arrows above them indicate the transcriptional direction. The expression levels of *cydA* and *MSMEG_3680* in the WT and Δ*crp1* mutant strains were quantitatively determined by qRT-PCR. The strains were aerobically grown to an OD_{600} of 0.45–0.5 and treated with 100 μM KCN for 15 min (+KCN). As controls, the WT and Δ*crp1* mutant strains grown aerobically without KCN treatment were included in the experiment (-KCN). The expression levels of *cydA* and *MSMEG_3680* determined by qRT-PCR were normalized to that of *sigA*. The expression level of each gene in the KCN-untreated WT strain is set at 1, and the relative values are expressed for the other strains. All values provide are the averages of the results from three independent experiments. The error bars indicate the standard deviations.

centered at positions −65.5 (for CBS2) and −93.5 (for CBS1). Crp was suggested to assist the binding of RNA polymerase to the promoter through its interactions with the C-terminal domain of α-subunit of RNA polymerase, when it serves as an transcriptional activator (Busby and Ebright, 1999; Lawson et al., 2004). Mutagenesis and DNA-binding analyses on CBS1 and CBS2 revealed that mutations in the distal Crp-binding CBS1 more severely affected *cydA* expression than those in CBS2 with a higher binding affinity for Crp1 (**Figures 4**, **5**, **6B**), implying that Crp bound at CBS2 might either stabilize the binding of Crp to CBS1 or help the formation of a DNA loop such that Crp bound at CBS1 can participate in recruitment or activation of RNA polymerase. Recently, it was demonstrated that expression of the *cydAB* operon is significantly reduced in a *prrA* null mutant of *M. smegmatis* compared to the WT strain (Maarsingh et al., 2019). However, our EMSA analysis using purified PrrA, which was preincubated with acetyl phosphate for phosphorylation, did not show the binding of PrrA to the upstream region of *cydA*, implying that the PrrBA TCS might indirectly participate in the regulation of the *cydAB* operon (data not shown).

Intracellular concentrations of cAMP in mycobacteria have been found to be significantly higher than those in other bacteria

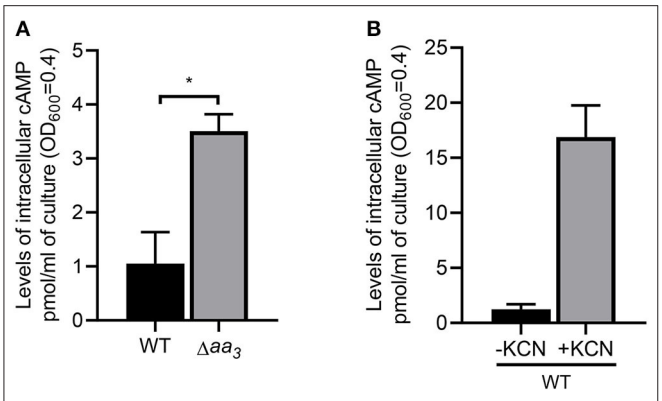


FIGURE 8 | Intracellular levels of cAMP in *M. smegmatis* under respiration-inhibitory conditions. **(A)** Intracellular levels of cAMP in the WT and Δaa_3 mutant strains of *M. smegmatis*. The WT and Δaa_3 mutant strains of *M. smegmatis* were grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium. **(B)** Effect of KCN treatment on intracellular levels of cAMP in *M. smegmatis*. The WT strain of *M. smegmatis* was grown aerobically to an OD_{600} of 0.45–0.5 in 7H9-glucose medium, and the cultures were treated with 100 μM KCN for 15 min (+KCN). As a control, the WT strain without treatment of KCN was used in the experiment (-KCN). cAMP quantification was performed using a DetectX Direct Cyclic AMP Enzyme Immunoassay kit (Arbor Assays) as described in Materials and Methods. All values provided are the average of the results from three biological replicates. The error bars indicate the standard deviations. Statistical significance was determined by two-tailed Student's *t*-test. $^*p < 0.05$.

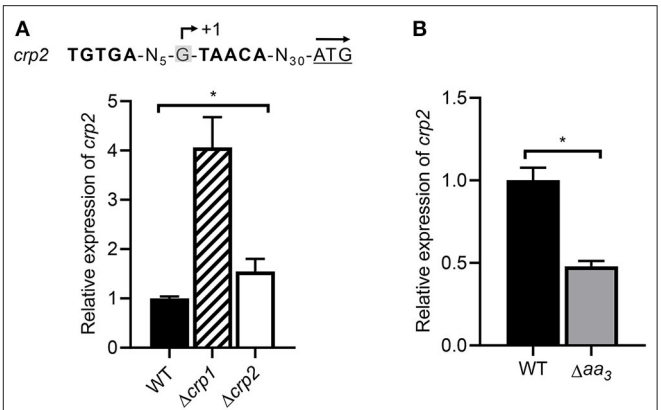


FIGURE 9 | Expression of the *crp2* gene in the WT, Δ*crp1*, Δ*crp2*, and Δaa_3 mutant strains of *M. smegmatis*. **(A)** Determination of the *crp2* transcript level in the WT, Δ*crp1*, and Δ*crp2* mutant strains by means of qRT-PCR. The positions of the identified Crp-binding site and the TSP (+1) relative to the start codon of *crp2* are shown above the graph (Martini et al., 2019). The start codon of *crp2* is underlined and the arrow above it indicates the transcriptional direction. The strains were grown aerobically to an OD_{600} of 0.45–0.5 and treated with 100 μM KCN for 15 min. The expression level of *crp2* in the KCN-treated WT strain is set at 1, and the relative values are expressed for the mutant strains. **(B)** Determination of the *crp2* transcript level in the WT and Δaa_3 mutant strains using qRT-PCR. The strains were grown aerobically to an OD_{600} of 0.45–0.5. The expression level of *crp2* in the WT strain is set at 1, and the relative values are expressed for the Δaa_3 mutant. All values provide are the average of the results from three biological replicates. The error bars indicate the standard deviations. Statistical significance was determined by two-tailed Student's *t*-test. $^*p < 0.05$.

grown under similar conditions (Padh and Venkitasubramanian, 1976; Lee, 1977; Shenoy and Visweswariah, 2006; Bai et al., 2011). Accordingly, it was suggested that mycobacteria have been evolved to possess the Crp proteins that have a low binding affinity for cAMP. Crp1 in *M. smegmatis* and Rv3676 in *M. tuberculosis* are the Crp proteins that have a low binding affinity for cAMP (Stapleton et al., 2010; Sharma et al., 2014). Although purified Rv3676 and Crp1 were shown to bind to their target DNA even in the absence of cAMP in contrast to Crp2 and *E. coli* Crp, their DNA-binding affinity was demonstrated to be enhanced by binding of cAMP (**Figures 4A**, **5A**) (Bai et al., 2005; Rickman et al., 2005; Sharma et al., 2014). Some fast-growing mycobacteria such as *Mycobacterium flavescens*, *Mycobacterium fortuitum*, and *Mycobacterium phlei* and the pathogenic slow-growing *Mycobacterium avium* contain an additional Crp protein that corresponds to Crp2 in *M. smegmatis* (Sharma et al., 2014). Compared to Crp1, Crp2 of *M. smegmatis* has been reported to have a much higher binding affinity for cAMP ($K_d = \sim 30\,\mu M$ for Crp1 and $K_d = \sim 3\,\mu M$ for Crp2). It has been also reported that Crp2 does not bind to the target DNA in the absence of cAMP (Sharma et al., 2014). Intracellular levels of cAMP were found to be 3.2-fold increased in *M. smegmatis* by the inactivation of the aa_3 cytochrome *c* oxidase and to rise even more when the aerobically grown WT strain of *M. smegmatis* was subjected to treatment of KCN (**Figure 8**), which is in line with our previous report that cAMP levels in *M. smegmatis* were increased by ∼400- and 5.7-fold under hypoxic and SNP (5 mM)-treated conditions, respectively (Jeon et al., 2014). The genome of *M. smegmatis* contains eight genes encoding adenylyl cyclase (*MSMEG_0228*, *MSMEG_3578*, *MSMEG_3780*, *MSMEG_4279*, *MSMEG_4477*, *MSMEG_4924*, *MSMEG_5018*, and *MSMEG_6154*). Although expression of *MSMEG_3780* and *MSMEG_4279* was observed to increase under both hypoxic and SNP-treated conditions, no detailed study was performed as to which adenylyl cyclase(s) is (are) implicated in an increase in cAMP levels in *M. smegmatis* under respiration-inhibitory conditions (Jeon et al., 2014). Considering both considerably different cAMP-binding affinity of two Crp paralogs and high intracellular levels of cAMP in mycobacteria, it can be assumed that Crp2 with the reportedly 10-fold higher binding affinity for cAMP than Crp1 might exist in a cAMP-bound form, while Crp1 with a low cAMP-binding affinity might be as a cAMP-unbound apo-protein in *M. smegmatis* under normal respiration conditions. If this assumption were true, Crp2 would be inappropriate to serve as a transcription factor that reflects elevating intracellular cAMP levels under respiration-inhibitory conditions to regulate gene expression. Furthermore, both significantly low cellular abundance of Crp2 relative to Crp1 and reduced expression of *crp2* under respiration-inhibitory conditions as a result of the negative regulation of *crp2* by Crp1 (**Figures 3**, **9**) renders Crp2 even more inadequate for a transcription factor that regulates gene expression in response to respiration inhibition. This assumption is consistent with our finding that Crp1, but not Crp2, plays a major role in upregulation of *cydAB* expression under respiration-inhibitory conditions.

In conclusion, we here provided the evidence that Crp1 (MSMEG_6189) of two Crp paralogs in *M. smegmatis* is the major transcription factor that is responsible for upregulation of the *cydAB* operon under ETC-inhibitory conditions. Two Crp-binding sequences were identified upstream of the *cydA* gene, both of which are necessary for induction of *cydAB* expression under ETC-inhibitory conditions. The intracellular level of cAMP in *M. smegmatis* was found to be increased under ETC-inhibitory conditions. The *crp2* gene was found to be negatively regulated by Crp1 and Crp2, which might lead to significantly low cellular abundance of Crp2 relative to Crp1 in *M. smegmatis*.

DATA AVAILABILITY STATEMENT

The RNA sequencing data have been deposited in NCBI's Gene Expression Omnibus and are accessible through the GEO Series accession number GSE158137.

AUTHOR CONTRIBUTIONS

E-MK and J-IO: conception or design of the study, analysis or interpretation of the data, and writing of the manuscript. E-MK: acquisition of the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a 2-years Research Grant of Pusan National University to J-IO.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.608624/full#supplementary-material

REFERENCES

Aung, H. L., Berney, M., and Cook, G. M. (2014). Hypoxia-activated cytochrome bd expression in *Mycobacterium smegmatis* is cyclic AMP receptor protein dependent. *J. Bacteriol.* 196, 3091–3097. doi: 10.1128/JB.01771-14

Aung, H. L., Dixon, L. L., Smith, L. J., Sweeney, N. P., Robson, J. R., Berney, M., et al. (2015). Novel regulatory roles of cAMP receptor proteins in fast-growing environmental mycobacteria. *Microbiology* 161, 648–661. doi: 10.1099/mic.0.000015

Bai, G., Knapp, G. S., and McDonough, K. A. (2011). Cyclic AMP signalling in mycobacteria: redirecting the conversation with a common currency. *Cell. Microbiol.* 13, 349–358. doi: 10.1111/j.1462-5822.2010.01562.x

Bai, G., McCue, L. A., and McDonough, K. A. (2005). Characterization of *Mycobacterium tuberculosis* Rv3676 (CRPMt), a cyclic AMP receptor protein-like DNA binding protein. *J. Bacteriol.* 187, 7795–7804. doi: 10.1128/JB.187.22.7795-7804.2005

Bekker, M., Alexeeva, S., Laan, W., Sawers, G., Teixeira de Mattos, J., and Hellingwerf, K. (2010). The ArcBA two-component system of *Escherichia coli* is regulated by the redox state of both the ubiquinone and the menaquinone pool. *J. Bacteriol.* 192, 746–754. doi: 10.1128/JB.01156-09

Belevich, I., Borisov, V. B., Bloch, D. A., Konstantinov, A. A., and Verkhovsky, M. I. (2007). Cytochrome bd from *Azotobacter vinelandii*: evidence for high-affinity oxygen binding. *Biochemistry* 46, 11177–11184. doi: 10.1021/bi700862u

Belevich, I., Borisov, V. B., Konstantinov, A. A., and Verkhovsky, M. I. (2005). Oxygenated complex of cytochrome bd from *Escherichia coli*: stability and photolability. *FEBS Lett.* 579, 4567–4570. doi: 10.1016/j.febslet.2005.07.011

Berney, M., Greening, C., Conrad, R., Jacobs, W. R. Jr., and Cook, G. M. (2014). An obligately aerobic soil bacterium activates fermentative hydrogen production to survive reductive stress during hypoxia. *Proc. Natl. Acad. Sci. U. S. A.* 111, 11479–11484. doi: 10.1073/pnas.1407034111

Boot, M., Jim, K. K., Liu, T., Commandeur, S., Lu, P., Verboom, T., et al. (2017). A fluorescence-based reporter for monitoring expression of mycobacterial cytochrome bd in response to antibacterials and during infection. *Sci. Rep.* 7:10665. doi: 10.1038/s41598-017-10944-4

Borisov, V. B., Forte, E., Davletshin, A., Mastronicola, D., Sarti, P., and Giuffre, A. (2013). Cytochrome bd oxidase from *Escherichia coli* displays high catalase activity: an additional defense against oxidative stress. *FEBS Lett.* 587, 2214–2218. doi: 10.1016/j.febslet.2013.05.047

Borisov, V. B., Forte, E., Konstantinov, A. A., Poole, R. K., Sarti, P., and Giuffre, A. (2004). Interaction of the bacterial terminal oxidase cytochrome bd with nitric oxide. *FEBS Lett.* 576, 201–204. doi: 10.1016/j.febslet.2004.09.013

Borisov, V. B., Gennis, R. B., Hemp, J., and Verkhovsky, M. I. (2011). The cytochrome bd respiratory oxygen reductases. *Biochim. Biophys. Acta* 1807, 1398–1413. doi: 10.1016/j.bbabio.2011.06.016

Brekasis, D., and Paget, M. S. (2003). A novel sensor of NADH/NAD^+ redox poise in Streptomyces coelicolor A3(2). *EMBO J.* 22, 4856–4865. doi: 10.1093/emboj/cdg453

Busby, S., and Ebright, R. H. (1999). Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* 293, 199–213. doi: 10.1006/jmbi.1999.3161

Cook, G. M., Hards, K., Vilcheze, C., Hartman, T., and Berney, M. (2014). Energetics of respiration and oxidative phosphorylation in mycobacteria. *Microbiol. Spectr.* 2:2013. doi: 10.1128/microbiolspec.MGM2-0015-2013

Cotter, P. A., Chepuri, V., Gennis, R. B., and Gunsalus, R. P. (1990). Cytochrome o (*cyoABCDE*) and d (*cydAB*) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the fnr gene product. *J. Bacteriol.* 172, 6333–6338. doi: 10.1128/jb.172.11.6333-6338.1990

Cotter, P. A., and Gunsalus, R. P. (1992). Contribution of the fnr and *arcA* gene products in coordinate regulation of cytochrome o and d oxidase (*cyoABCDE* and *cydAB*) genes in *Escherichia coli*. *FEMS Microbiol. Lett.* 70, 31–36. doi: 10.1016/0378-1097(92)90558-6

Cotter, P. A., Melville, S. B., Albrecht, J. A., and Gunsalus, R. P. (1997). Aerobic regulation of cytochrome d oxidase (*cydAB*) operon expression in *Escherichia coli*: roles of Fnr and ArcA in repression and activation. *Mol. Microbiol.* 25, 605–615. doi: 10.1046/j.1365-2958.1997.5031860.x

Cunningham, L., Pitt, M., and Williams, H. D. (1997). The *cioAB* genes from *Pseudomonas aeruginosa* code for a novel cyanide-insensitive terminal oxidase related to the cytochrome bd quinol oxidases. *Mol. Microbiol.* 24, 579–591. doi: 10.1046/j.1365-2958.1997.3561728.x

Forte, E., Borisov, V. B., Falabella, M., Colaco, H. G., Tinajero-Trejo, M., Poole, R. K., et al. (2016). The terminal oxidase cytochrome bd promotes sulfide-resistant bacterial respiration and growth. *Sci. Rep.* 6:23788. doi: 10.1038/srep23788

Fu, H. A., Iuchi, S., and Lin, E. C. (1991). The requirement of ArcA and Fnr for peak expression of the *cyd* operon in *Escherichia coli* under microaerobic conditions. *Mol. Gen. Genet.* 226, 209–213. doi: 10.1007/BF00273605

Georgellis, D., Kwon, O., and Lin, E. C. (2001). Quinones as the redox signal for the Arc two-component system of bacteria. *Science* 292, 2314–2316. doi: 10.1126/science.1059361

Giuffre, A., Borisov, V. B., Arese, M., Sarti, P., and Forte, E. (2014). Cytochrome bd oxidase and bacterial tolerance to oxidative and nitrosative stress. *Biochim. Biophys. Acta* 1837, 1178–1187. doi: 10.1016/j.bbabio.2014.01.016

Giuffre, A., Borisov, V. B., Mastronicola, D., Sarti, P., and Forte, E. (2012). Cytochrome bd oxidase and nitric oxide: from reaction mechanisms to bacterial physiology. *FEBS Lett.* 586, 622–629. doi: 10.1016/j.febslet.2011.07.035

Green, J., Stapleton, M. R., Smith, L. J., Artymiuk, P. J., Kahramanoglou, C., Hunt, D. M., et al. (2014). Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches. *Curr. Opin. Microbiol.* 18, 1–7. doi: 10.1016/j.mib.2014.01.003

Gyan, S., Shiohira, Y., Sato, I., Takeuchi, M., and Sato, T. (2006). Regulatory loop between redox sensing of the NADH/NAD^+ ratio by Rex (YdiH) and oxidation of NADH by NADH dehydrogenase Ndh in *Bacillus subtilis*. *J. Bacteriol.* 188, 7062–7071. doi: 10.1128/JB.00601-06

Heroven, A. K., and Dersch, P. (2014). Coregulation of host-adapted metabolism and virulence by pathogenic yersiniae. *Front. Cell. Infect. Microbiol.* 4:146. doi: 10.3389/fcimb.2014.00146

Jeon, H. S., Yang, K. H., Ko, I. J., and Oh, J. I. (2014). Expression of adenylyl cyclase genes in *Mycobacterium smegmatis* under hypoxic and nitric oxide conditions. *J. Life Sci.* 24, 1330–1338. doi: 10.5352/JLS.2014.24.12.1330

Jeong, J. A., Park, S. W., Yoon, D., Kim, S., Kang, H. Y., and Oh, J. I. (2018). Roles of alanine dehydrogenase and induction of its gene in *Mycobacterium smegmatis* under respiration-inhibitory conditions. *J. Bacteriol.* 200:e00152-18. doi: 10.1128/JB.00152-18

Jones-Carson, J., Husain, M., Liu, L., Orlicky, D. J., and Vazquez-Torres, A. (2016). Cytochrome bd-dependent bioenergetics and antinitrosative defenses in salmonella pathogenesis. *mBio* 7:e02052-16. doi: 10.1128/mBio.02052-16

Kalia, N. P., Shi Lee, B., Ab Rahman, N. B., Moraski, G. C., Miller, M. J., and Pethe, K. (2019). Carbon metabolism modulates the efficacy of drugs targeting the cytochrome bc1:aa3 in *Mycobacterium tuberculosis*. *Sci. Rep.* 9:8608. doi: 10.1038/s41598-019-44887-9

Kana, B. D., Weinstein, E. A., Avarbock, D., Dawes, S. S., Rubin, H., and Mizrahi, V. (2001). Characterization of the *cydAB*-encoded cytochrome bd oxidase from *Mycobacterium smegmatis*. *J. Bacteriol.* 183, 7076–7086. doi: 10.1128/JB.183.24.7076-7086.2001

Kiley, P. J., and Beinert, H. (1998). Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol. Rev.* 22, 341–352. doi: 10.1111/j.1574-6976.1998.tb00375.x

Kim, M. J., Park, K. J., Ko, I. J., Kim, Y. M., and Oh, J. I. (2010). Different roles of DosS and DosT in the hypoxic adaptation of mycobacteria. *J. Bacteriol.* 192, 4868–4875. doi: 10.1128/JB.00550-10

Koul, A., Vranckx, L., Dhar, N., Gohlmann, H. W., Ozdemir, E., Neefs, J. M., et al. (2014). Delayed bactericidal response of *Mycobacterium tuberculosis* to bedaquiline involves remodelling of bacterial metabolism. *Nat. Commun.* 5:3369. doi: 10.1038/ncomms4369

Larsson, J. T., Rogstam, A., and von Wachenfeldt, C. (2005). Coordinated patterns of cytochrome bd and lactate dehydrogenase expression in *Bacillus subtilis*. *Microbiology* 151, 3323–3335. doi: 10.1099/mic.0.28124-0

Lawson, C. L., Swigon, D., Murakami, K. S., Darst, S. A., Berman, H. M., and Ebright, R. H. (2004). Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* 14, 10–20. doi: 10.1016/j.sbi.2004.01.012

Lee, C. H. (1977). Identification of adenosine 3′,5′-monophosphate in *Mycobacterium smegmatis*. *J. Bacteriol.* 132, 1031–1033. doi: 10.1128/JB.132.3.1031-1033.1977

Lee, H. N., Ji, C. J., Lee, H. H., Park, J., Seo, Y. S., Lee, J. W., et al. (2018). Roles of three FurA paralogs in the regulation of genes pertaining to peroxide defense in *Mycobacterium smegmatis* mc^2 155. *Mol. Microbiol.* 108, 661–682. doi: 10.1111/mmi.13956

Lee, H. N., Lee, N. O., Han, S. J., Ko, I. J., and Oh, J. I. (2014). Regulation of the *ahpC* gene encoding alkyl hydroperoxide reductase in *Mycobacterium smegmatis*. *PLoS ONE* 9:e111680. doi: 10.1371/journal.pone.0111680

Lee, J. M., Cho, H. Y., Cho, H. J., Ko, I. J., Park, S. W., Baik, H. S., et al. (2008). O_2- and NO-sensing mechanism through the DevSR two-component system in *Mycobacterium smegmatis*. *J. Bacteriol.* 190, 6795–6804. doi: 10.1128/JB.00401-08

Lindqvist, A., Membrillo-Hernandez, J., Poole, R. K., and Cook, G. M. (2000). Roles of respiratory oxidases in protecting *Escherichia coli* K12 from oxidative stress. *Antonie Van Leeuwenhoek* 78, 23–31. doi: 10.1023/a:1002779201379

Maarsingh, J. D., Yang, S., Park, J. G., and Haydel, S. E. (2019). Comparative transcriptomics reveals PrrAB-mediated control of metabolic, respiration, energy-generating, and dormancy pathways in *Mycobacterium smegmatis*. *BMC Genom.* 20:942. doi: 10.1186/s12864-019-6105-3

Malpica, R., Franco, B., Rodriguez, C., Kwon, O., and Georgellis, D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13318–13323. doi: 10.1073/pnas.0403064101

Martini, M. C., Zhou, Y., Sun, H., and Shell, S. S. (2019). Defining the transcriptional and post-transcriptional landscapes of *Mycobacterium smegmatis* in aerobic growth and hypoxia. *Front. Microbiol.* 10:591. doi: 10.3389/fmicb.2019.00591

Mascolo, L., and Bald, D. (2020). Cytochrome bd in *Mycobacterium tuberculosis*: a respiratory chain protein involved in the defense against antibacterials. *Prog. Biophys. Mol. Biol.* 152, 55–63. doi: 10.1016/j.pbiomolbio.2019.11.002

Mason, M. G., Shepherd, M., Nicholls, P., Dobbin, P. S., Dodsworth, K. S., Poole, R. K., et al. (2009). Cytochrome bd confers nitric oxide resistance to *Escherichia coli*. *Nat. Chem. Biol.* 5, 94–96. doi: 10.1038/nchembio.135

Matsoso, L. G., Kana, B. D., Crellin, P. K., Lea-Smith, D. J., Pelosi, A., Powell, D., et al. (2005). Function of the cytochrome bc1-aa3 branch of the respiratory network in mycobacteria and network adaptation occurring in response to its disruption. *J. Bacteriol.* 187, 6300–6308. doi: 10.1128/JB.187.18.6300-6308.2005

Mayuri, B.agchi, G., Das, T. K., and Tyagi, J.S. (2002). Molecular analysis of the dormancy response in *Mycobacterium smegmatis*: expression analysis of genes encoding the DevR-DevS two-component system, Rv3134c and chaperone alpha-crystallin homologues. *FEMS Microbiol. Lett.* 211, 231–237. doi: 10.1111/j.1574-6968.2002.tb11230.x

Megehee, J. A., Hosler, J. P., and Lundrigan, M. D. (2006). Evidence for a cytochrome bcc-aa3 interaction in the respiratory chain of *Mycobacterium smegmatis*. *Microbiology* 152, 823–829. doi: 10.1099/mic.0.28723-0

Mouncey, N. J., and Kaplan, S. (1998). Redox-dependent gene regulation in rhodobacter sphaeroides 2.4.1(T): effects on dimethyl sulfoxide reductase (*dor*) gene expression. *J. Bacteriol.* 180, 5612–5618. doi: 10.1128/JB.180.21.5612-5618.1998

Oh, J. I., and Kaplan, S. (1999). The cbb3 terminal oxidase of *Rhodobacter sphaeroides* 2.4.1: structural and functional implications for the regulation of spectral complex formation. *Biochemistry* 38, 2688–2696. doi: 10.1021/bi9825100

Oh, Y., Song, S. Y., Kim, H. Y., Han, G., Hwang, J., Kang, H. Y., et al. (2020). The partner switching system of the SigF sigma factor in Mycobacterium smegmatis and induction of the SigF regulon under respiration-inhibitory conditions. *Front. Microbiol.* 11:588487. doi: 10.3389/fmicb.2020.588487

O'Toole, R., Smeulders, M. J., Blokpoel, M. C., Kay, E. J., Lougheed, K., and Williams, H. D. (2003). A two-component regulator of universal stress protein expression and adaptation to oxygen starvation in *Mycobacterium smegmatis*. *J. Bacteriol.* 185, 1543–1554. doi: 10.1128/jb.185.5.1543-1554.2003

Padh, H., and Venkitasubramanian, T. A. (1976). Cyclic adenosine 3′, 5′-monophosphate in mycobacteria. *Indian J. Biochem. Biophys.* 13, 413–414

Puustinen, A., Finel, M., Haltia, T., Gennis, R. B., and Wikstrom, M. (1991). Properties of the two terminal oxidases of *Escherichia coli*. *Biochemistry* 30, 3936–3942. doi: 10.1021/bi00230a019

Rahman, M. A., Cumming, B. M., Addicott, K. W., Pacl, H. T., Russell, S. L., Nargan, K., et al. (2020). Hydrogen sulfide dysregulates the immune response by suppressing central carbon metabolism to promote tuberculosis. *Proc. Natl. Acad. Sci. U. S. A.* 117, 6663–6674. doi: 10.1073/pnas.1919211117

Rickman, L., Scott, C., Hunt, D. M., Hutchinson, T., Menendez, M. C., Whalan, R., et al. (2005). A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the *rpfA* gene coding for a resuscitation promoting factor. *Mol. Microbiol.* 56, 1274–1286. doi: 10.1111/j.1365-2958.2005.04609.x

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Saini, V., Chinta, K. C., Reddy, V. P., Glasgow, J. N., Stein, A., Lamprecht, D. A., et al. (2020). Hydrogen sulfide stimulates *Mycobacterium tuberculosis* respiration, growth and pathogenesis. *Nat. Commun.* 11:557. doi: 10.1038/s41467-019-14132-y

Sambrook, J., and Green, M. R. (2012). *Molrcular Cloning: A Laboratory Manual. 4th Edn.* New York, NY: Cold Spring Harbor Laboratory Press.

Schau, M., Chen, Y., and Hulett, F. M. (2004). *Bacillus subtilis* YdiH is a direct negative regulator of the *cydABCD* operon. *J. Bacteriol.* 186, 4585–4595. doi: 10.1128/JB.186.14.4585-4595.2004

Sharma, R., Zaveri, A., Gopalakrishnapai, J., Srinath, T., Varshney, U., and Visweswariah, S. S. (2014). Paralogous cAMP receptor proteins in *Mycobacterium smegmatis* show biochemical and functional divergence. *Biochemistry* 53, 7765–7776. doi: 10.1021/bi500924v

Shenoy, A. R., and Visweswariah, S. S. (2006). New messages from old messengers: cAMP and mycobacteria. *Trends. Microbiol.* 14, 543–550. doi: 10.1016/j.tim.2006.10.005

Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V., et al. (2005). Changes in energy metabolism of *Mycobacterium tuberculosis* in mouse lung and under *in vitro* conditions affecting aerobic respiration. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15629–15634. doi: 10.1073/pnas.0507850102

Shimada, T., Fujita, N., Yamamoto, K., and Ishihama, A. (2011). Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS ONE* 6:e20081. doi: 10.1371/journal.pone.0020081

Singh, A. K., Dutta, D., Singh, V., Srivastava, V., Biswas, R. K., and Singh, B. N. (2015). Characterization of *Mycobacterium smegmatis sigF* mutant and its regulon: overexpression of SigF antagonist (MSMEG_1803) in *M. smegmatis mimics sigF mutant phenotype, loss of pigmentation, and sensitivity to oxidative stress. Microbiologyopen* 4, 896–916. doi: 10.1002/mbo3.288

Small, J. L., Park, S. W., Kana, B. D., Ioerger, T. R., Sacchettini, J. C., and Ehrt, S. (2013). Perturbation of cytochrome c maturation reveals adaptability of the respiratory chain in *Mycobacterium tuberculosis*. *mBio* 4:e00475-13. doi: 10.1128/mBio.00475-13

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., and Jacobs, W. R. Jr. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* 4, 1911–1919. doi: 10.1111/j.1365-2958.1990.tb02040.x

Stapleton, M., Haq, I., Hunt, D. M., Arnvig, K. B., Artymiuk, P. J., Buxton, R. S., et al. (2010). *Mycobacterium tuberculosis* cAMP receptor protein (Rv3676) differs from the *Escherichia coli* paradigm in its cAMP binding and DNA binding properties and transcription activation properties. *J. Biol. Chem.* 285, 7016–7027. doi: 10.1074/jbc.M109.047720

Swem, L. R., Gong, X., Yu, C. A., and Bauer, C. E. (2006). Identification of a ubiquinone-binding site that affects autophosphorylation of the sensor kinase RegB. *J. Biol. Chem.* 281, 6768–6775. doi: 10.1074/jbc.M509687200

Tseng, C. P., Albrecht, J., and Gunsalus, R. P. (1996). Effect of microaerophilic cell growth conditions on expression of the aerobic (*cyoABCDE* and *cydAB*) and anaerobic (*narGHJI, frdABCD*, and *dmsABC*) respiratory pathway genes in *Escherichia coli*. *J. Bacteriol.* 178, 1094–1098. doi: 10.1128/jb.178.4.1094-1098.1996

Utsumi, R., Noda, M., Kawamukai, M., and Komano, T. (1989). Control mechanism of the *Escherichia coli* K-12 cell cycle is triggered by the cyclic AMP-cyclic AMP receptor protein complex. *J. Bacteriol.* 171, 2909–2912. doi: 10.1128/jb.171.5.2909-2912.1989

Wu, J., and Bauer, C. E. (2010). RegB kinase activity is controlled in part by monitoring the ratio of oxidized to reduced ubiquinones in the ubiquinone pool. *mBio* 1:e00272-10. doi: 10.1128/mBio.00272-10

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: 09 December 2020
doi: 10.3389/fmicb.2020.590851

Integrated Transcriptomic and Proteomic Analyses Reveal the Role of NprR in *Bacillus anthracis* Extracellular Protease Expression Regulation and Oxidative Stress Responses

Yanchun Wang[1]†, Na Jiang[2]†, Bowen Wang[1], Haoxia Tao[1], Xin Zhang[1], Qing Guan[1] and Chunjie Liu[1]**

[1] State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Biotechnology, Beijing, China, [2] Beijing Fisheries Research Institute, Beijing, China

OPEN ACCESS

Edited by:
Jianping Xie,
Southwest University, China

Reviewed by:
Colin Harwood,
Newcastle University, United Kingdom
Shauna McGillivray,
Texas Christian University,
United States

***Correspondence:**
Chunjie Liu
liucj@nic.bmi.ac.cn
Yanchun Wang
springwyc@163.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: 03 August 2020
Accepted: 18 November 2020
Published: 09 December 2020

Citation:
Wang Y, Jiang N, Wang B, Tao H, Zhang X, Guan Q and Liu C (2020) Integrated Transcriptomic and Proteomic Analyses Reveal the Role of NprR in Bacillus anthracis Extracellular Protease Expression Regulation and Oxidative Stress Responses. Front. Microbiol. 11:590851. doi: 10.3389/fmicb.2020.590851

NprR is a protein of *Bacillus anthracis* that exhibits moonlighting functions as either a phosphatase or a neutral protease regulator that belongs to the RNPP family. We previously observed that the extracellular protease activity of an *nprR* deletion mutant significantly decreased within *in vitro* cultures. To identify the genes within the regulatory network of *nprR* that contribute to its protease activity, integrated transcriptomic and proteomic analyses were conducted here by comparing the *nprR* deletion mutant and parent strains. A total of 366 differentially expressed genes (DEGs) between the strains were observed via RNA-seq analysis. In addition, label-free LC-MS/MS analysis revealed 503 differentially expressed proteins (DEPs) within the intracellular protein fraction and 213 extracellular DEPs with significant expressional differences between the strains. The majority of DEGs and DEPs were involved in environmental information processing and metabolism. Integrated transcriptomic and proteomic analyses indicated that oxidation-reduction-related GO terms for intracellular DEPs and endopeptidase-related GO terms for extracellular DEPs were significantly enriched in the mutant strain. Notably, many genes involved in protease activity were largely downregulated in the *nprR* deletion mutant cultures. Moreover, western blot analysis revealed that the major extracellular neutral protease Npr599 was barely expressed in the *nprR* deletion mutant strain. The mutant also exhibited impaired degradation of protective antigen, which is a major *B. anthracis* toxin component, thereby resulting in higher protein yields. Concomitantly, another global transcriptional regulator, SpxA1, was also dramatically downregulated in the *nprR* deletion mutant, resulting in higher sensitivity to oxidative and disulfide stress. These data consequently indicate that NprR is a transcriptional regulator that controls genes whose products function as extracellular proteases and also is involved in oxidative stress responses. This study thus contributes to a more comprehensive understanding of the biological function of NprR, and especially in the middle growth stages of *B. anthracis*.

Keywords: neutral protease regulator, extracellular proteases, *Bacillus anthracis*, transcriptomic, proteomic, oxidative stress responses

INTRODUCTION

Extracellular proteases are secreted bacterial virulence factors that exhibit multiple roles in which they can attack host cells or tissues either directly or indirectly (Frees et al., 2013; Marshall et al., 2017). These proteases generally promote the breakdown of host tissues in order to promote bacterial dissemination into hosts and the colonization of deeper tissue layers or organs (Miyoshi and Shinoda, 2000; Vieira and Nascimento, 2016). Npr599 and InhA are the most important extracellular proteases for *Bacillus anthracis* virulence and enhance dissemination in addition to other activities (Chung et al., 2006, 2011). Npr599 and InhA first directly cleave extracellular matrix (ECM) components in order to degrade host tissues and increase barrier permeability. In addition, Npr599 and InhA activate syndecan-1 shedding by stimulating the target cell shedding mechanisms and accelerating the process through direct cleavage of the syndecan-1 ectodomain (Chung et al., 2006; Popova et al., 2006). Moreover, Npr599 and InhA can activate the fibrinolytic system via the human pro-urokinase plasminogen activator (pro-uPA) activation pathway or by cleaving the thrombin-activatable fibrinolysis inhibitor (TAFI), which has been observed in anthrax infections (Chung et al., 2011). Concomitantly, InhA1 can also directly activate prothrombin and factor X, resulting in blood coagulation (Kastrup et al., 2008). Perhaps more importantly, InhA can increase blood-brain barrier permeability and contribute to cerebral hemorrhages and is thus a critical virulence factor (Mukherjee et al., 2011). However, the expression and regulation of the *npr599* and *inhA* genes are uncharacterized.

NprR is a "moonlighting protein" of *Bacillus cereus* bacteria that functions as either a phosphatase or a neutral protease regulator and belongs to the RRNPP family that includes Rap, Rgg, NprR, PlcR, and PrgX (Rocha-Estrada et al., 2010; Rocha et al., 2012; Rice et al., 2016; Neiditch et al., 2017). NprR can modulate diverse activities by forming different polymeric structures. In general, mature NprX binds NprR and promotes the formation of a tetrameric structure of NprR that binds DNA and alters transcriptional activation of genes (Perchat et al., 2011; Zouhir et al., 2013). In the absence of NprX, NprR forms dimers and does not bind DNA, instead functioning as a phosphatase of Spo0F (Cabrera et al., 2016; Perchat et al., 2016). Furthermore, NprR is a component of NprR/NprX quorum sensing (QS) cell-cell communication that is controlled by the expression of the important neutral protease gene *nprA* (*npr599*) and a suite of other genes (Dubois et al., 2012; Perchat et al., 2016; Wu et al., 2017). When *B. anthracis* Sterne colonies were incubated at 30°C for six or more days on LB agar plates, many papillation mutants were observed, and very few of them exhibited extracellular protease activities. The *nprR* genes in most of these mutants exhibited different mutational variations including point mutations, insertions, and deletions of various lengths. These mutational NprR could not act as neutral protease regulators, which strongly indicates that the mutations lead to the loss of the extracellular protease activities in *B. anthracis* (Yang et al., 2011). Moreover, *B. anthracis* V770-NP1-R (ATCC 14185), which has been used for efficient production of PA subunit vaccines, carries one nonsense mutation within the *nprR* gene (locus GBAA0597) that affects the expression of the neutral protease Npr599 (Cohen-Gihon et al., 2014).

Previously, we constructed the *nprR* gene deletion strain *B. anthracis* A16RΔ*nprR* (Wang et al., 2018). To further characterize the NprR regulatory system, comparative strand-specific RNA-seq and label-free LC-MS/MS were used in this study to identify genes that may be associated with the NprR regulons. Transcriptomic and proteomic analyses revealed that NprR is a critical regulator for bacterial protease expression. In addition, we observed for the first time that the NprR of *B. anthracis* is involved in stress response systems. This work contributes to our comprehensive understanding of the molecular mechanism underlying the activity and function of the neutral protease regulator NprR during bacterial growth.

MATERIALS AND METHODS

Strains and Growth Conditions

The bacterial strains used in this study are shown in **Table 1**. *B. anthracis* A16RΔ*nprR* was previously described (Wang et al., 2018), while all other strains were constructed for this study and are described in further detail below. *Escherichia coli* DH5α was used for cloning experiments, and the plasmids used for electroporation of *B. anthracis* were prepared from *E. coli* SCS110. All strains were grown at 37°C with shaking at 220 rpm in Luria broth (LB). Antibiotics (Merck, Germany) were added to the media at appropriate final concentrations of erythromycin (400 μg/ml for *E. coli*, 5 μg/ml for *B. anthracis*), spectinomycin (50 μg/ml for *E. coli*, 300 μg/ml for *B. anthracis*), kanamycin (30 μg/ml for *B. anthracis*), and ampicillin (100 μg/ml, only for *E. coli*).

RNA Purification, Library Construction, and Sequencing

Three independent cultures from wild-type *B. anthracis* A16R and A16RΔ*nprR* were grown in LB over 12 h for transcriptomic experiments. Total RNA was isolated from cells using the TRIzol reagent (Invitrogen), followed by the treatment of extracts with RNase-free DNase I (Qiagen, Germany). rRNA was removed using a Ribo-Zero Magnetic kit (Epicentre, United States), and enriched mRNA was used to generate an RNA-Seq library using a TruSeq™ Stranded Total RNA Library Prep Kit (Illumina, United States). The library was sequenced on an Illumina HiSeq 4000 platform at Majorbio (Shanghai, China), and the data were analyzed on the online Majorbio Cloud Platform[1].

Clean sequence reads were mapped to the *B. anthracis* Sterne chromosome and the pXO1 plasmid using Bowtie. mRNA expression levels and transcripts per million (TPM) reads were calculated for each gene using RSEM (RNA-Seq by Expectation-Maximization). The TPM distribution density and correlation between the two groups were also analyzed to validate the reliability of the sequencing data.

Differentially expressed genes (DEGs) between A16R and A16RΔ*nprR* cultures were identified using DEseq2 with

[1]www.majorbio.com

TABLE 1 | Plasmids and strains used in this study.

Plasmids and strains	Relevant characteristics	Source
Plasmid		
pUC-PmtlA	pUC57 with PmtlA promoter	This lab
pBE2	Amp^R, Kan^R, *B. anthracis–E. coli* shuttle expression vector	This lab
pBE-nprR	pBE2 containing *nprR*	This work
pBE-spxA1	pBE2 containing *spxA1*	This work
B. anthracis		
A16R	$pXO1^+pXO2^-$, China vaccine strain	This lab
A16RΔ*nprR* (nprR)	A16R excision beginning 360 bp fragment of *nprR*	Wang et al., 2018
A16RΔ*nprR*(pBE-nprR) (c-nprR)	A16RΔnprR expressing *nprR* in plasmid pBE2	This work
A16RΔ*nprR*(pBE-spxA1) (c-spxA1)	A16RΔnprR expressing *spxA1* in plasmid pBE2	This work
A16RΔ*npr599* (npr599)	A16R deletion *npr599* gene	This lab
E. coli strains		
DH5α	cloning strain	CWBIO, China
SCS110	*dam-/dcm-* strain used to produce unmethylated plasmid	TransGen, China

a Benjamini-Hochberg false discovery rate (FDR) used to determine the threshold *p* value for assigning statistical significance after multiple tests. A fold change value ≥ 2 and FDR ≤ 0.05 were used as thresholds for identifying significant differential expression.

Functional Annotation of DEGs

Differentially expressed genes were subjected to Gene Ontology (GO) and KEGG enrichment analyses based on GO and KEGG pathway analysis using the Goatools and KOBAS software programs, respectively. The statistical significance of pathway enrichments was identified using Fisher's exact test. *P* values from the GO and KEGG enrichment analyses were also corrected using multi-testing correction methods including Bonferroni, Holm, Sidak, and false discovery rate corrections. GO and KEGG pathways were considered significantly enriched when corrected *p* values were ≤ 0.05. All data were analyzed on the Majorbio Cloud Platform[2].

Protein Extraction and Label-Free LC-MS/MS Analysis

Three independent cultures from wild-type *B. anthracis* A16R and A16RΔ*nprR* were grown in 400 mL of LB over 14 h for proteomic analyses. The cell pellets and supernatant fractions of cultures were then separated by centrifugation at 17,000 *g* for 20 min at 4°C. To investigate extracellular proteins, the supernatant fraction was filtered using a 0.44-μm syringe filter, and the filtered growth medium was concentrated using ultrafiltration tubes (3 kDa MWCO, PALL). To investigate intracellular proteins, cell pellets were washed several times with PBS buffer and then resuspended in a urea-based extraction buffer (8.0 M urea with a protease inhibitor cocktail) and sonicated for 10 min following centrifugation at 12,000 *g* for 30 min at 4°C. Protein concentrations were estimated using a BCA Protein Assay Kit (Thermo Fisher Scientific, United States), and protein quality was assessed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). A 100 μg/vial protein sample was reduced and alkylated with 100 mM triethylammonium bicarbonate (TEAB) buffer along with 100 mM tris (2-carboxyethyl) phosphine (TCEP) and 40 mM iodoacetamide (IAM) by incubating for 40 min at room temperature in the dark. Six volumes of cold acetone were then added (cold acetone: sample v/v = 6:1), and proteins were precipitated for 4 h at −20°C. After centrifugation at 4°C for 3 min, the resultant precipitates were dissolved in 100 μL of 100 mM TEAB buffer.

Dissolved proteins were digested with sequencing grade trypsin (Promega) at an enzyme: substrate ratio of 1:50 (w/w) overnight at 37°C. After digestion, the peptides were vacuum-dried and dissolved in 0.1% trifluoroacetic acid (TFA), followed by desalting with an HLB SPE column (AISIMO) and drying by vacuum centrifugation. The concentrations of desalted peptides were then measured using a Pierce Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific) and dissolved in MS loading buffer (2% acetonitrile and 0.1% formic acid) to a final concentration of 0.25 μg/μL.

Peptides were separated with an acetonitrile gradient (2–80% over 120 min) in 0.1% formic acid using a reversed-phase analytical column (C18 column; Thermo Fisher Scientific), with a constant flow rate of 300 nl/min on an EASY-nLC 1200 UPLC system (Thermo Fisher Scientific, United States). Eluted peptides were analyzed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific, United States) operated in data-dependent acquisition (DDA) mode. Full-scan MS spectra were acquired between 350 and 1,300 m/z, with the 20 most intense signals per cycle fragmented. The resolutions of the MS and MS/MS scans were 60,000 and 15,000, respectively, and the dynamic exclusion was set to 18 s. Each sample was analyzed in triplicate.

The peptides were identified from the raw data using the Proteome Discoverer v.2.2 software program (Thermo Fisher Scientific). Software parameters were specified as follows: iodoacetamide for Cys alkylation, oxidation (M), and acetyl (protein N-terminus) for dynamic modification, carbamidomethyl for static modification, trypsin as the enzyme

[2]https://cloud.majorbio.com/

name, a maximum of two missed cleavage sites, 20 ppm for mass tolerance, and 1% false discovery rate. All data were analyzed on the Majorbio Cloud Platform (see text footnote 2). The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium[3] via the iProX partner repository with the dataset identifier PXD020800 (Ma et al., 2019).

Construction of *nprR* and *spxA1* Complemented Strains

The genomic DNA of *B. anthracis* A16R was used as a template in order to amplify an *nprR* fragment with the promoter region using the nprRF and nprRR primers. The PmtlA promoter was amplified by mtlAF and mtlAR using pUC-PmtlA as template, while the spxA1 gene was amplified by spxA1F and spxA1R from A16R genomic DNA. All the sequences of the primers are listed in **Table 2**. The fragments were then cloned into the shuttle plasmid pBE2 to construct the expression plasmids for NprR and SpxA1. The expression plasmids were transformed into *B. anthracis* A16RΔ*nprR* as described previously (Wang et al., 2011).

Phenotypic and Protein Expression Analysis

Extracellular protease activity phenotypes of the strains were investigated on milk plates, as previously described (Yang et al., 2011; Pineda-Castellanos et al., 2015). Briefly, fresh colonies grown on LB agar were used to inoculate an overnight LB culture (37°C, 220 rpm). The cultures (5 μl) were spotted onto milk plates and then incubated at 37°C. A distinct halo around the colonies was then identified at different time points as a phenotype.

To analyze the expression of secreted proteins like Npr599, InhA1, and PA, the *B. anthracis* strains were grown at 37°C in LB for 14 h, and secreted proteins were analyzed by Western blot. The supernatant fraction was collected from the bacterial cultures by centrifugation at 8,000 *g* for 10 min at 4°C, followed by filtering with a 0.45 μm syringe filter. The conditioned growth medium was concentrated 20-fold using Amicon Ultra-15 membranes (Millipore), and equivalent volumes (10 μL) of each sample were subjected to 12% SDS-PAGE. Separated proteins were transferred to an NC membrane (GE Healthcare, United States) using the Bio-Rad Trans-Blot system. The membranes were then incubated in blocking solutions containing 5% fat-free milk for 1 h and incubated with respective primary antibodies for 1.5 h at 37°C. After washing, the membranes were incubated with measurable IgG conjugated to HRP at a dilution of 1:5,000 (Jackson, United States). The membranes were treated with ECL reagents (Engreen Biosystem, Beijing, China) and visualized using a chemiluminescent imaging system (Tanon, China), according to the manufacturer's instructions.

To analyze the expression of cytoplasmic proteins, cell pellets were resuspended using a 1/10 volume of PBS buffer and briefly sonicated followed by centrifugation at 16,000 *g* for 10 min at 4°C. The supernatant fractions were collected and analyzed by SDS-PAGE and western blot analysis, as described above. The expression of the ribosomal protein L6 was used as a loading control and was detected in samples from all strains (Wang et al., 2011).

Quantitative Real-Time PCR

Quantitative real-time PCR analysis was performed to validate the RNA-seq data using RNA preparations that were generated as described above. The RNAs were subjected to cDNA synthesis using the HiScript Q RT SuperMix for qPCR (Vazyme, China). qPCR reactions were performed using a 2X ChamQ SYBR COLOR qPCR master mix (Vazyme, China) on a 7500 Fast Real-Time PCR System (Applied Biosystems, United States). All the sequences of the primers are listed in **Table 2**. Experimental duplicates were conducted for reactions using three independent cDNA preparations. Levels of expression of *spxA1* gene were normalized to a constitutively expressed gene (*fusA*). The efficiency of qPCR primers used was validated using thermal gradient and gel electrophoresis. The relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Stress Challenges

Fresh colonies grown on LB agar were used to inoculate overnight LB cultures at 37°C and 220 rpm to investigate the role of NprR in *B. anthracis* stress responses. Ten-fold serial dilutions of the cultures were spotted (5 μl) onto LB agar either supplemented or not supplemented with 200 μmol/L diamide, 0.44 mmol/L H_2O_2, or 0.0025% deoxycholate. Sensitivity on the LB agar plates to the supplemented chemicals was evaluated after 16 h of incubation at 37°C.

RESULTS

Transcriptomic Analysis Following *nprR* Disruption

To evaluate the effect of *nprR* on transcriptomic profiles, RNA-Seq analysis was conducted using triplicate biological samples of the *nprR* mutant strain and A16R reference strain from early stationary phase cultures (∼ 12 h of growth). An average of 23.5 million clean reads was obtained per sample. All total mapping ratios and uniquely mapped ratios of clean reads to the *B. anthracis* Sterne strain genome were nearly 99%, indicating that sequencing was sufficient to cover all transcripts within the cells. Genome mapping revealed that 5,662 and 5,670 genes were expressed in the A16R and *nprR* mutant strain cells, respectively. A total of 5752 genes were detected and, among these, 5,580 were commonly expressed within the two groups. The distribution density of TPM values did not obviously differ between the two groups, and correlation analysis also illustrated a high degree of correlation of expression between the two samples (**Supplementary Figure 1**). Differentially expressed genes (DEGs) between the *nprR* mutant and A16R strains were identified that had at least 2-fold changes in expression ($|\log_2 FC(NprR/A16R)| \geq 1$) and a corrected *p*-value (FDR) < 0.05 (**Figure 1** and **Supplementary Table 1**). Gene expression comparisons indicated that 279 genes were significantly downregulated (green in **Figure 1**) and 50

[3]http://proteomecentral.proteomexchange.org

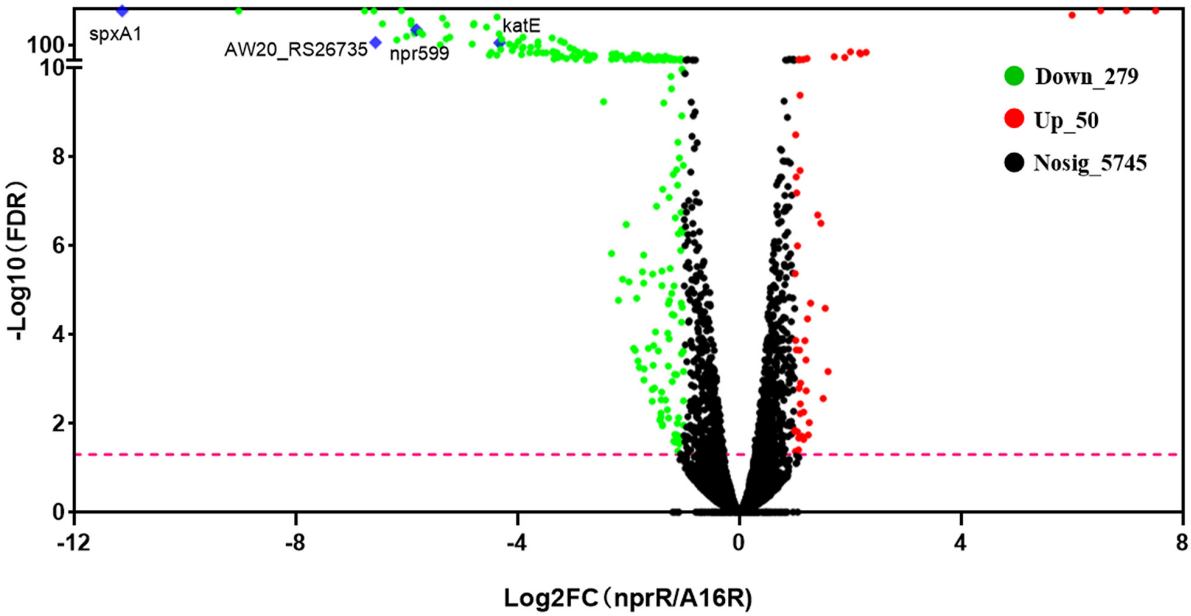


FIGURE 1 | Volcano plots of differentially expressed genes (DEGs). Plots showing log2-fold change (FC) and $-\log_{10}$ FDR (corrected *p*-value) values of the DEGs when comparing the RNA-Seq profile of the *nprR* deletion mutant strain against that of the A16R strain. Green, red, and black points indicate downregulated, upregulated, and non-differentially expressed genes, respectively. The pink line represents FDR = 0.05. Some focused genes also were marked by blue diamond.

were significantly upregulated (red in **Figure 1**) in the *nprR* deletion mutant relative to A16R cells. Only one gene on the plasmid pXO1 (AW20_RS00735) was downregulated in the mutant. The *npr599* gene, the direct target of NprR activation, was one of the most downregulated genes in the mutant ($\log_2$FC(NprR/A16R) = −5.81).

GO and KEGG Enrichment Analysis of DEGs

To explore the functions of the DEGs and the pathways that they participate in, the DEGs were classified into GO categories, yielding three predominant classifications: biological process (BP), cellular component (CC), and molecular function (MF). Specifically, the DEGs were assigned to 24 functional groups by GO annotation (**Figure 2A**). GO functional enrichment analysis of DEGs was also conducted based on the annotation data. GO enrichment results of all DEGs were provided. A total of 24 GO terms were significantly enriched (corrected *p*-value, FDR < 0.05). In the top 20 GO terms, most of the terms were classified within the "biological process" and "molecular function" categories (**Figure 2B**). Among these, "secondary metabolite biosynthetic process," "siderophore metabolic process," "non-ribosomal peptide biosynthetic process," "oxidoreductase activity," "siderophore biosynthetic process," and "secondary metabolic process" were the primary significantly enriched downregulated GO terms (FDR < 0.01), while the aldonic acid metabolic process and D-gluconate metabolic process were the primary significantly enriched upregulated GO terms (FDR < 0.01). KEGG pathway annotation and enrichment analysis were also used to evaluate significantly enriched DEGs (**Figures 2C,D**). Among the 20 most represented pathways for these DEGs, "biosynthesis of siderophore group non-ribosomal peptides" (pathway ID: map01053) was the only one that was significantly downregulated in *nprR* mutant strains (FDR = 0.0001).

Proteomic Analysis and Identification of Differentially Expressed Proteins (DEPs)

To further understand the effects of nprR gene deletion on *B. anthracis* physiology, label-free LC-MS/MS analysis was used to evaluate the expression of intra- and extra-cellular proteins

TABLE 2 | Primers used in this study.

Name	Sequence (5′→3′)	Purpose
nprRF	ACATGCATGCACAACTTCACCTTCTTGCATAC	PCR of *nprR* cassette
nprRR	CGGAATTCTTATTCCTCCTTATCATTC	
mtlAF	ACATGCATGCTTTATTTTTAAAAAATTGTC	PCR of PmtlA promoter
mtlAR	TTTGGTCTCTTTATATATTTCC	
spxA1F	TTTGGTCTCTATAATATGGTAACATTATATAGTTC	PCR of *spxA1* gene
spxA1R	CGGAATTCGAAGGTTTTCATATTATTATAA	
q-spxA1F	TGTACGTCTTGTAGAAAGGC	qPCR of spxA1 gene
q-spxA1R	AATAATCTCATCCGTTCCGC	
fusAF	TTGGTATCATGGCTCACATC	qPCR of fusA gene
fusAR	GCAGCAGAAGTAATTGTGAT	

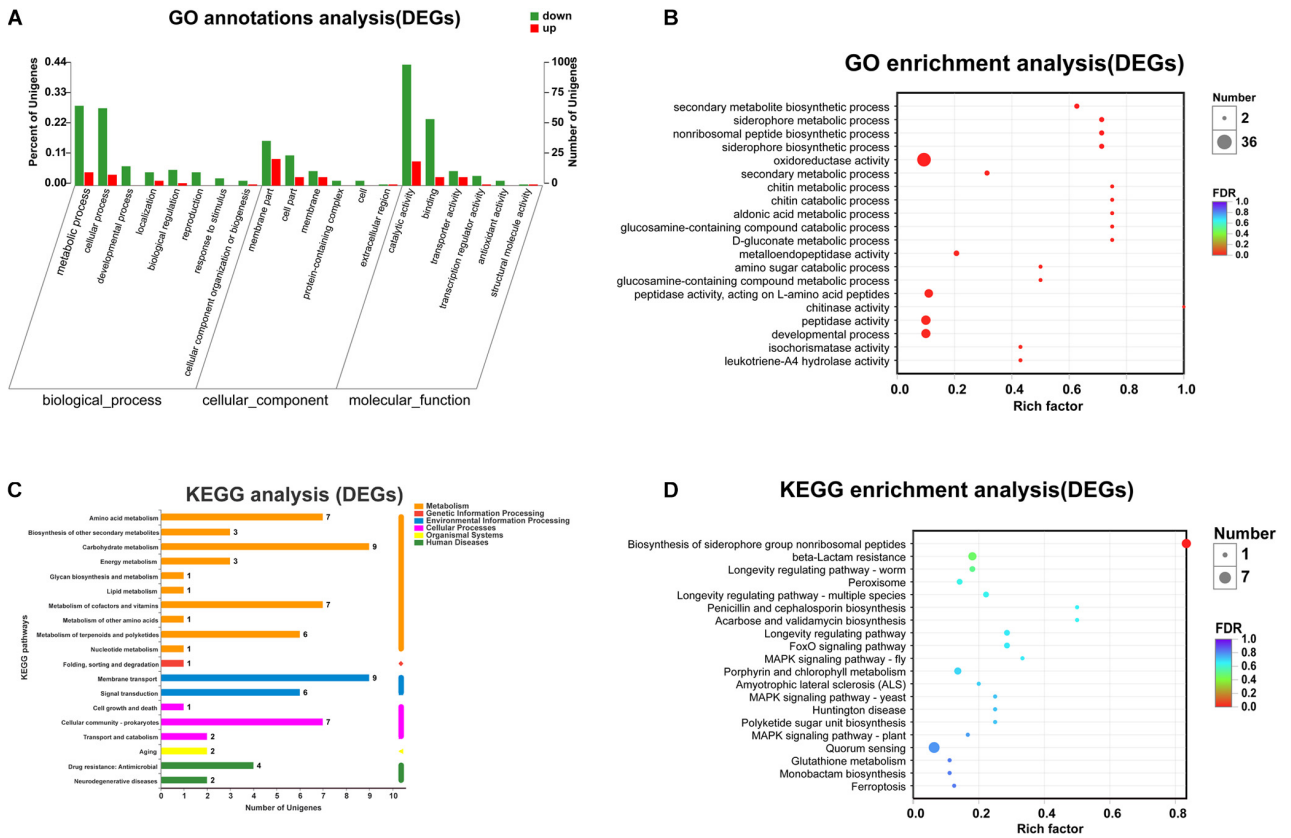


FIGURE 2 | Functional annotation of the DEGs in the mutant and reference strains. **(A)** Top 20 terms of GO annotations. Red and green columns represent the number of upregulated and downregulated unigenes in GO terms, respectively. **(B)** GO functional enrichment analysis (The 20 most enriched GO terms). **(C)** KEGG pathway annotations. **(D)** KEGG pathway enrichment analysis. "Rich factor" means that the ratio of the DEGs number and the number of genes has been annotated in this term or pathway. The greater the Rich factor, the greater the degree of enrichment. The size of each point indicates the number of differentially expressed genes in that term (pathway).

in the *B. anthracis* A16R and the *nprR* deletion mutant strains. DEPs were those exhibiting $P < 0.05$ and fold change (FC) >1.20 or <0.87 between groups. A total of 2,135 intracellular proteins were quantified, of which 503 were identified as DEPs, including 269 downregulated and 234 upregulated proteins in *nprR* deletion mutant strains (**Figure 3A** and **Supplementary Table 2**). Likewise, 379 extracellular proteins were identified, with 213 being DEPs and 30 downregulated, while 183 were upregulated (**Figure 3B** and **Supplementary Table 3**).

GO annotation revealed that the 503 intracellular DEPs were assigned to 35 upregulated GO terms and 30 downregulated. The 20 most abundant GO terms (based on protein abundances) are shown in **Figure 4A**. GO functional enrichment analysis of the DEGs was also conducted using the annotation data (**Figure 4B** and **Supplementary Table 4**). The 20 most significantly enriched GO terms were primarily associated with the BP and MF categories and included "oxidoreductase activity," "oxidation-reduction process," "reactive oxygen species metabolic process," "FMN binding," "oligopeptide transport," "oligopeptide transporter activity," and "response to oxidative stress" ($p < 0.001$). In addition, GO term analysis of the 279 extracellular DEPs resulted in an annotation of 39 upregulated and 21 downregulated GO terms, and the 20 most abundant GO terms are shown in **Figure 4C**. However, only two of these were significantly differentially enriched and were associated with "protein binding" and "endopeptidase activity" (**Supplementary Figure 2**). In addition to the GO term analysis, KEGG pathway analysis was also conducted to assign the DEPs of the *nprR* deletion mutant and A16R strains to metabolic pathways. A total of 503 intracellular DEPs were assigned into 85 KEGG pathways, the majority of which were associated with metabolism (M) (**Figure 4D**). Four specific KEGG pathways comprising "beta-lactam resistance," "ABC transporters," "quorum sensing," and "biofilm formation" were significantly enriched in the DEP profiles (FDR < 0.05) (**Figure 4E**). Extracellular DEPs were also assigned to 61 KEGG pathways, but none were significantly enriched between groups (**Figure 4F**).

Integrated Transcriptomic and Proteomic Analysis

To better understand the correlation between mRNA and protein expression profiles, an integrated analysis of the proteomic and transcriptomic data was conducted and the results are

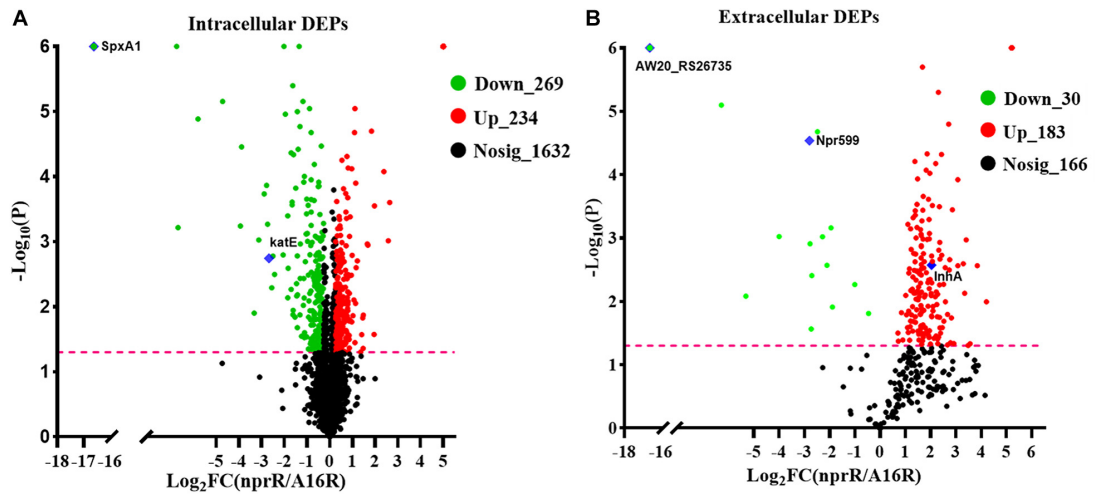


FIGURE 3 | Volcano plots of differentially expressed proteins (DEPs) in the *nprR* deletion mutant and A16R strain. The plots show log2-fold change (FC) and $-\log_{10}$ (*p*-value) comparisons of the DEPs from the *nprR* deletion mutant strain against the A16R strain. A FC > 1.2 was used to identify upregulated DEPs and a FC < 0.83 for downregulated DEPs. **(A)** intracellular protein fraction. **(B)** Extracellular protein fraction. Green, red, and black points represent downregulated, upregulated, and non-differentially regulated proteins, respectively. The pink line represents FDR = 0.05. Some focused proteins are marked with blue diamonds.

listed in **Supplementary Tables 5**, **6**. Pearson correlation coefficient *rho* was used to assess the correlation between the proteomic and transcriptomic data. All of the quantified proteins, including 2,135 intracellular and 379 extracellular proteins corresponded to those inferred from the RNA-seq data. The distribution of the corresponding mRNA: protein ratios is shown in a scatterplot of the log2-transformed ratios. For intracellular proteins, the integration analyses data revealed that 57 genes (green plot) and their corresponding proteins were downregulated, while 10 (red plot) were upregulated (**Figure 5A**). Likewise, among extracellular proteins, there were 14 upregulated and two downregulated (**Figure 5B**) genes/proteins that had corresponding changes of expression. The details of the genes and proteins that had corresponding changes in expression are listed in **Supplementary Tables 7**, **8**. As in many other studies (Haider and Pal, 2013), a poor positive correlation was observed between transcript and protein abundances in the mutant and A16R strains, with rho = 0.3542 for intracellular proteins and rho = 0.2763 for extracellular proteins (**Figures 5A,B**). Integrated analyses were subsequently conducted based on the DEPs and DEGs functional enrichment analysis in order to explore the correlations between the proteomes and transcriptomes from the perspective of protein function. Significantly enriched GO terms for intracellular proteins (FDR < 0.05) for both DEGs and DEPs are shown in **Figure 5C**. The GO terms "oxidoreductase activity," "oxidation-reduction process," "FMN binding," and "monooxygenase activity" were significantly enriched in both the DEPs and DEGs profiles, while the KEGG pathway "beta-lactam resistance" was the only pathway significantly enriched in both the DEPs and DEGs profiles (**Figure 5D**). Regarding extracellular proteins, "endopeptidase activity" was the only GO term significantly enriched in both the extracellular DEGs and DEPs profiles (**Supplementary Figure 3**, FDR < 0.05), and no KEGG pathways were significantly enriched in the extracellular DEGs and DEPs profiles. This result was fully consistent with the priorities of neutral protease regulators.

The Deletion of *nprR* in *B. anthracis* Decreases Proteolytic Enzyme Activity

Some DEGs and DEPs were clearly associated with protease activity based on annotation information. To better understand the regulation of protease-associated protein expression by NprR, several downregulated genes involved in protease activity identified in the RNA-seq and MS analyses were further evaluated (**Figure 6A**). Numerous proteins (including Npr599 and a peptidase M4 family protein) that were associated with protease activity were downregulated in the *nprR* deletion mutant strain. Furthermore, the MS analysis indicated that many proteases (peptidases) were also upregulated in the *nprR* deletion mutant strain, but not in RNA-seq result (**Figure 6A**). In particular, the InhA protein, another important protease for *B. anthracis* infection, was upregulated in the *nprR* deletion mutant strain relative to the A16R strain. Consequently, the extracellular proteolytic activity of strains was evaluated using phenotypic tests on skim milk agar media. Extracellular enzymatic activity is expressed as a halo (i.e., a zone of hydrolysis) around colonies in this test. Based on halo size, the *nprR* deletion mutant strain exhibited a large decrease in extracellular proteolytic enzymatic activity compared to the original A16R strain (**Figure 6B**). In contrast, the *nprR* complemented strain exhibited similar extracellular proteolytic activity as the A16R strain, indicating that plasmid-expressed NprR could recover the extracellular proteolytic enzyme activity of *nprR* deletion mutant strains. *npr599* encodes the neutral protease Npr599, which is the most important gene regulated by NprR. However, the *npr599* deletion mutant strain had weak extracellular proteolytic activity compared to

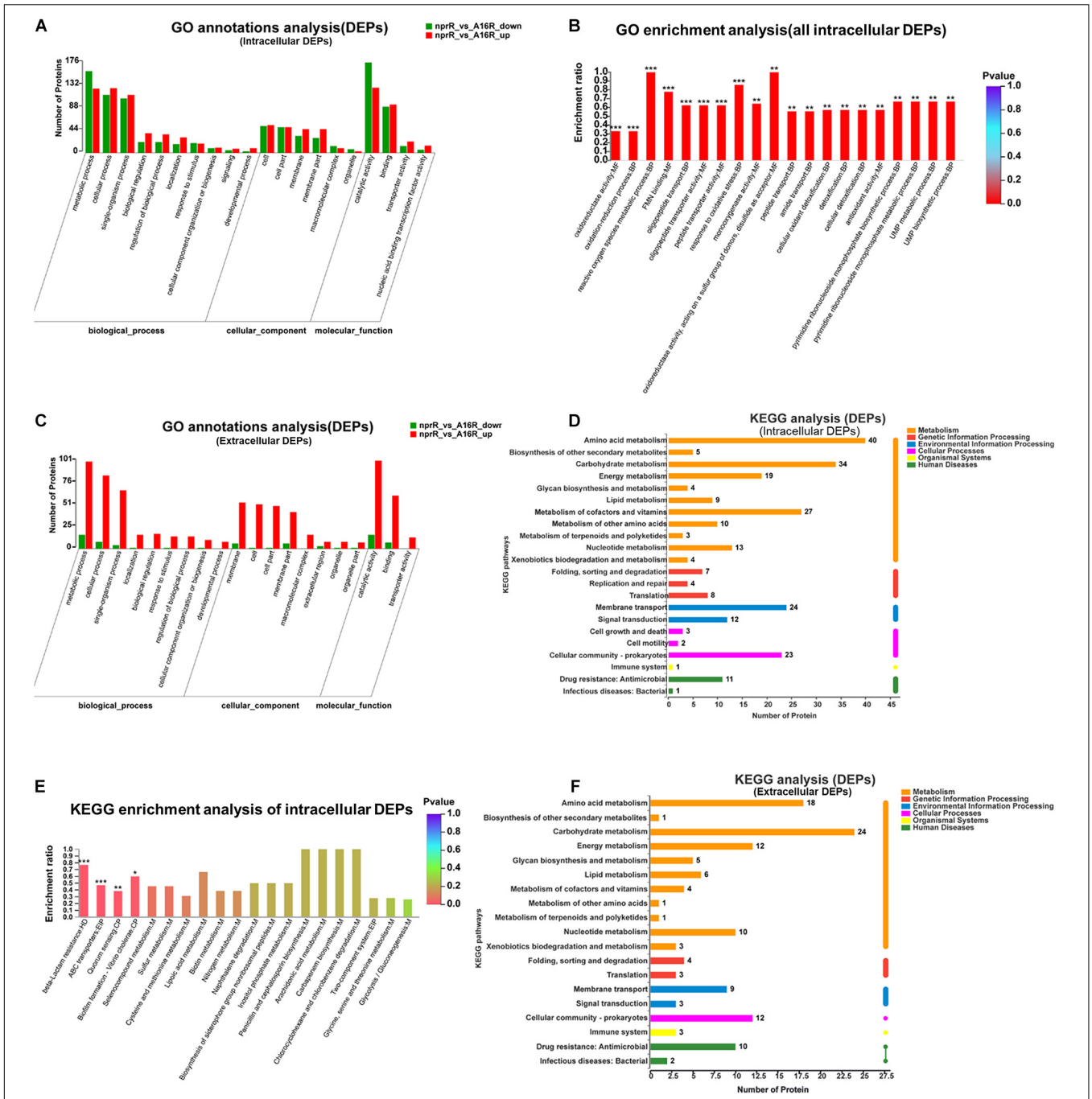


FIGURE 4 | Functional classification of DEPs. **(A)** GO annotation of intracellular DEPs. **(B)** GO functional enrichment analysis of intracellular DEPs. **(C)** GO annotation of extracellular DEPs. **(D)** KEGG annotation of intracellular DEPs. **(E)** KEGG pathway enrichment analysis of intracellular DEPs. **(F)** KEGG pathway annotations of extracellular DEPs. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

the *nprR* deletion mutant strain, but the latter had almost no extracellular enzymatic activity. Thus, Npr599 is not the only extracellular proteolytic enzyme regulated by NprR, and others are also downregulated by this important neutral protease regulator.

To further investigate the effects of extracellular proteolytic enzymes on the expression of other important proteins, the culture media of the *nprR* deletion mutant was used for further western blot analysis by recovering equal concentrations of protein samples from both the mutant and A16R strain cultures. The expression of the neutral protease Npr599 was almost undetectable in the *nprR* deletion mutant strain compared to the A16R strain. In addition, *nprR* complementation resulted in normal Npr599 expression relative to the A16R strain

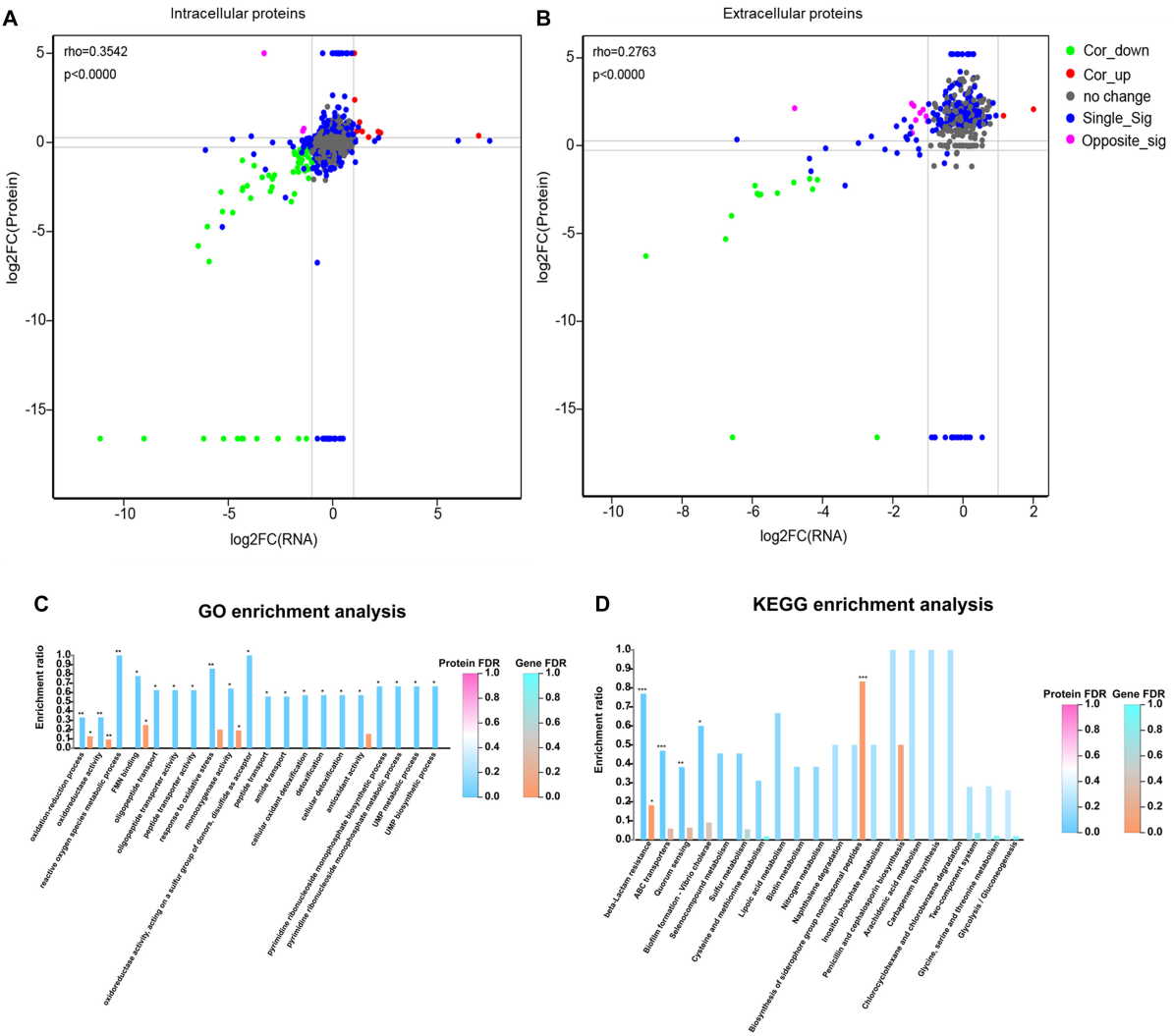


FIGURE 5 | Integrated analysis of the RNA-seq and label-free MS data for comparisons of the mutant and reference strains. **(A,B)** Correlation analysis between the RNA-seq and MS data for intracellular and extracellular protein fractions, respectively. In the nine-quadrant diagram, green (Cor_down) represents downregulated genes and proteins, red (Cor_up) represents upregulated genes and proteins, gray (no change) represents genes and proteins with no significant difference, blue (Single_Sig) represents one of the genes and proteins with no significant difference, purple (Opposite_Sig) indicates that DEGs and DEPs show opposite upregulation and downregulation. **(C)** Integrated GO term enrichment analysis of DEGs and DEPs for the intracellular protein fraction. **(D)** Integrated KEGG pathway enrichment analysis of DEGs and DEPs for the intracellular protein fraction. *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001.

(**Figure 6C**). In contrast, the expression of another important extracellular proteolytic enzyme, InhA, was only slightly upregulated (**Figure 6D**). These results were mostly consistent with the RNA-seq [*npr599*, $\log_2$FC(NprR/A16R) = −5.82; *inhA1*, $\log_2$FC(NprR/A16R) = 0.06] and MS [Npr599, $\log_2$FC(NprR/A16R) = −2.8; InhA, $\log_2$FC(NprR/A16R) = 2.03] analyses. In addition, the degradation of protective antigen (PA) proteins was significantly reduced in the mutant strains, while complemented NprR expression in the *nprR* deletion mutant led to similar degradation as the A16R strain (**Figure 6E**). Concomitantly, the lethal factor (LF) expression level was also slightly upregulated (**Figure 6F**). These results generally coincided with those of the MS analysis [PA, $\log_2$FC(NprR/A16R) = 1.536; LF, $\log_2$FC(NprR/A16R = 1.458)].

Deletion of NprR in *B. anthracis* Increases Oxidative Stress Susceptibility

The *spxA1* gene was significantly downregulated according to the RNA-seq and MS analyses [$\log_2$FC(NprR/A16R) = −11.13 and −16.61, respectively]. qPCR and western blot analyses were thus used to validate these results. The expression of SpxA1 was almost undetectable in the *nprR* deletion mutant strain compared to the A16R and *nprR* complemented strains (**Figure 7A**). Because cells lacking *spxA1* are sensitive to

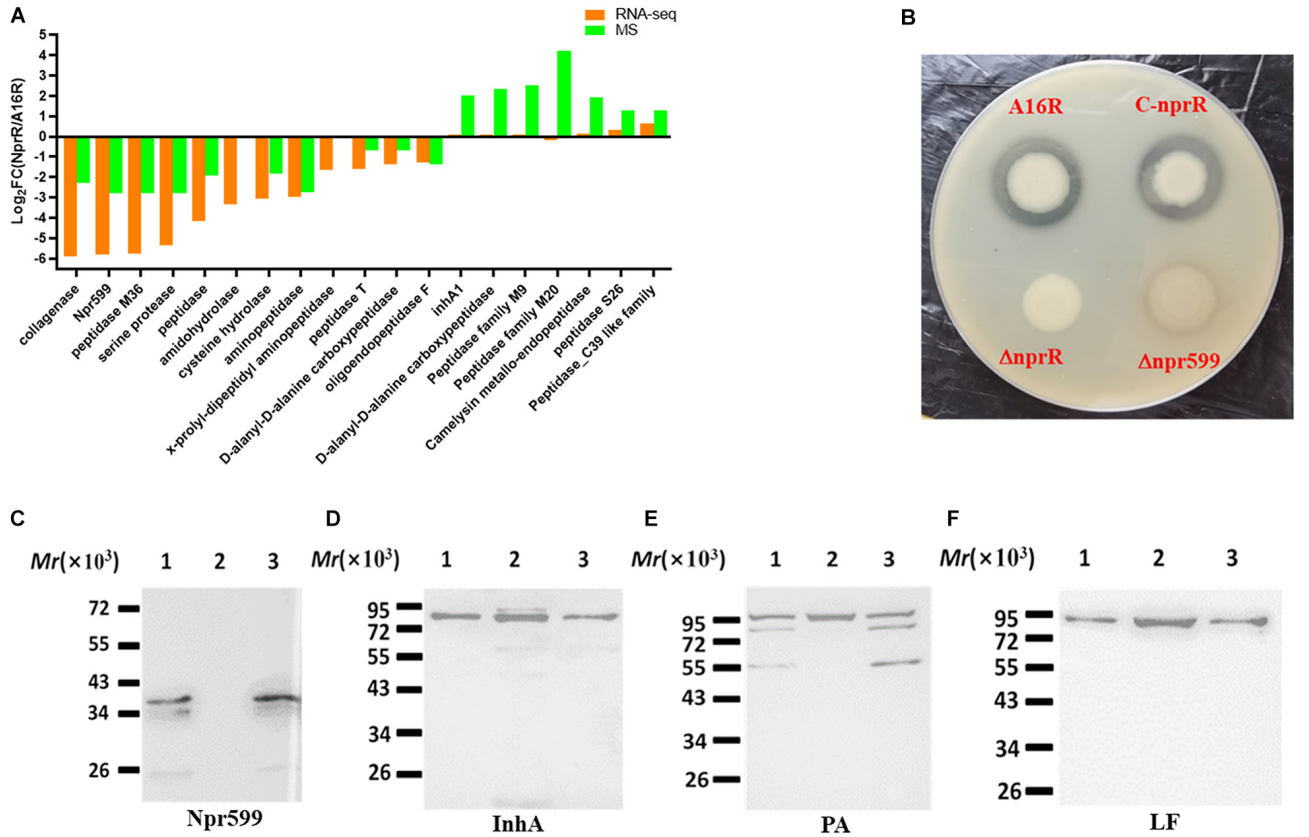


FIGURE 6 | Analysis of protease-related phenotypic characteristics of the *nprR* deletion mutant. **(A)** Comparison of protease-related DEGs and DEPs in the *nprR* deletion mutant strain against the A16R strain. **(B)** Extracellular proteolytic enzyme activity assays for different strains. Strain colonies with considerable extracellular proteolytic enzyme activity are surrounded by halos. **(C–F)** Western blot analysis to evaluate the expression of important *B. anthracis* proteins including Npr599 **(C)**, InhA **(D)**, PA **(E)**, and LF **(F)**. Lanes are indicated as follows: 1, A16R; 2, the *nprR* deletion mutant of A16R, 3, the *nprR* recovered strain. All Western blot experiments were performed at least three times.

oxidative stress, an oxidative stress susceptibility test was applied to the *nprR* mutant strain using several oxidants in culture agar plates. The mutant strain was more sensitive to peroxide, diamide, and salt deoxycholate relative to the A16R strain (**Figure 7B**). Thus, deletion of *nprR* made the strain more sensitive to oxidative stress and disulfide formation. To explore whether the loss of stress resistance in the *nprR* mutant is due to the downregulation of *spxA1* gene directly, the *nprR* mutant complemented with an SpxA1 expression plasmid was constructed. Oxidative stress susceptibility testing showed that defects in peroxide, diamide, and salt deoxycholate resistance could be corrected by introducing *spxA1* in plasmid, suggesting that downregulation of SpxA1 protein might be the main reason that the *nprR* deletion mutant strain was more sensitive to oxidative stress (**Figure 7C**). RNA-seq and MS analyses also indicated that the expression levels of AW20_RS26735 encoding peroxidase [$\log_2$FC(NprR/A16R) = −6.563 or −16.61, respectively] and AW20_RS27305 encoding catalase [*katE*, $\log_2$FC(NprR/A16R) = −4.323 or −2.679, respectively] were significantly downregulated in the *nprR* deletion mutant strain compared to the A16R strain (**Figures 1**, **3**).

DISCUSSION

NprR is an important regulatory protein of *B. anthracis*, with a primary function to regulate the expression of extracellular proteases like Npr599 in the early stages of bacterial growth (Rice et al., 2016). Here, we evaluated the role of this regulatory factor in the post-exponential growth phase. To explore the mechanism of NprR regulation of extracellular protease expression, an *nprR* deletion mutant was used to systematically investigate downstream genes regulated by NprR using RNA-seq and label-free MS approaches, with targeted validation by qPCR and western blot experiments. The analyses yielded clarification of NprR function during the middle stages of bacterial growth. Extracellular protease activity of the *nprR* deletion mutant was less than that of the *npr599* mutant. Thus, in addition to the downstream gene *npr599* that is regulated by NprR, the expression of other genes involved in extracellular protease activity was also likely regulated, which was supported by the RNA-seq and MS analyses. For example, the only GO term significantly enriched in both the DEGs and DEPs ($P < 0.05$) was "endopeptidase activity." Nevertheless, exceptions to this activity were also observed. Specifically, InhA is another important

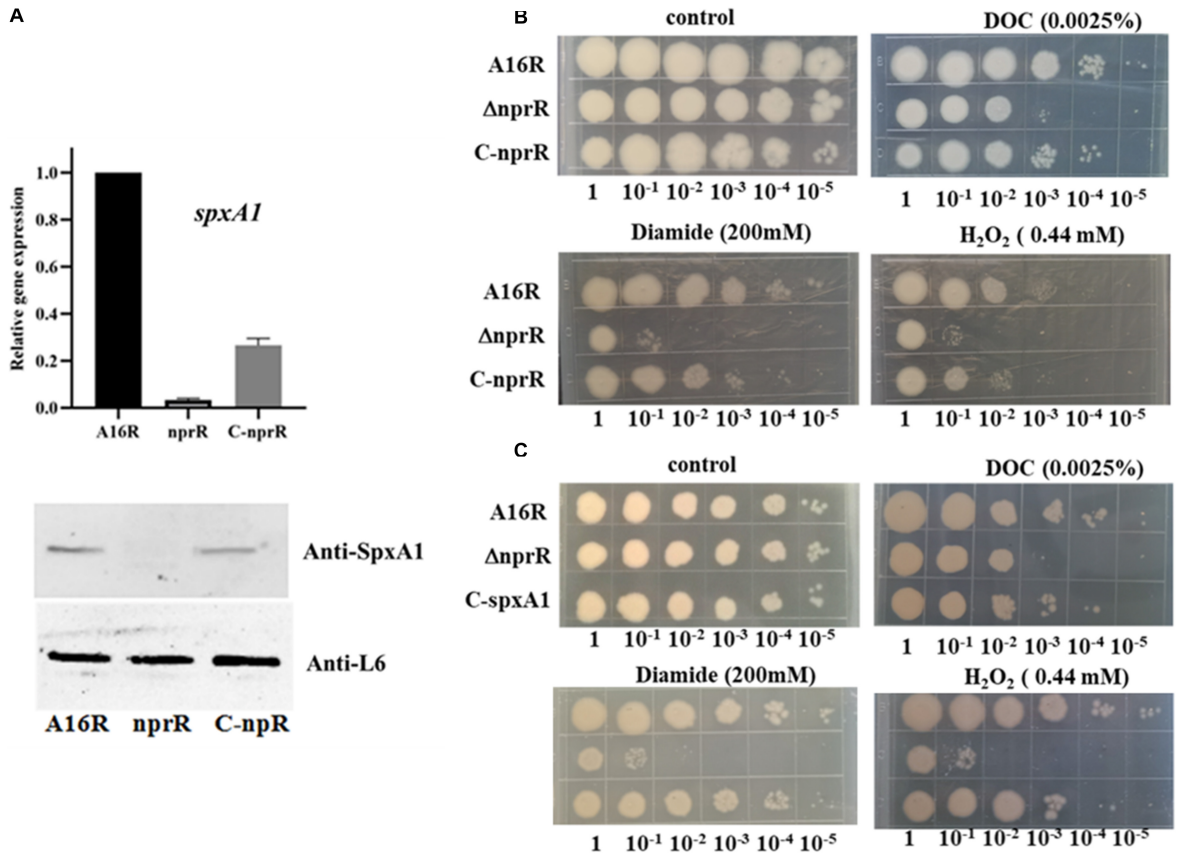


FIGURE 7 | Oxidative stress susceptibility analysis of the *nprR* deletion mutant. **(A)** qPCR and western blot analysis of *spxA1* expression. **(B,C)** Strains were grown on LB media with or without 0.0025% DOC, 200 mmol/L diamide, or 0.44 mmol/L H_2O_2. Five microliters at the indicated dilutions were spotted onto LB agar (Materials and Methods). All experiments were performed at least three times.

extracellular protease, and its expression was significantly upregulated based on the MS analysis, although mRNA levels did not change in the RNA-seq analyses. We speculate that this difference may primarily arise because InhA is one of the targets of Npr599, so that when the expression of Npr599 was significantly downregulated, InhA expression was upregulated due to reduced enzyme digestion effects.

In our opinion, the protein degradation caused by Npr599 was the main reason for the decrease in the amounts of some extracellular protein, such as PA and LF, and the degradation made it possible for this mutant strain to develop an effective host for production of recombinant anthrax antigen proteins (Pomerantsev et al., 2011). In addition, owing to the downregulation of Npr599 expression in the *nprR* mutant strain, the degradation of many extracellular proteins was significantly reduced, but there were no readily visible differences in the transcriptional level of these genes (coded extracellular proteins) between mutant strain and parental strain A16R. This might explain why the positive correlation between transcript and protein abundances was poor, especially regarding extracellular proteins (**Figures 5A,B**).

We also observed that NprR is associated with the oxidative stress response of *B. anthracis*. RNA-seq analysis indicated that the expression level of the *spxA1* (AW20_RS08630) and *spxA2* (AW20_RS25945) genes significantly downregulated in the *nprR* deletion mutant strain. Both SpxA1 and SpxA2 are closely related to the oxidative stress responses of *B. anthracis*, and *spxA1* is essential for peroxide resistance and participates in disulfide stress tolerance (Barendt et al., 2013, 2016). SpxA1 is also involved in oxidative stress response of many gram-positive bacteria (Chen et al., 2012; Kajfasz et al., 2017; Nilsson et al., 2019; Cesinger et al., 2020). For example, Chen et al. (2012) reported that SpxA1 is a global regulator required to activate genes encoding catalase in *Listeria monocytogenes*, and Nilsson et al. (2019) also reported that SpxA1 is involved in the oxidative stress response in *Streptococcus mutans*. As far as our results are concerned, SpxA1 exhibited the greatest change in expression levels in the mutant strain compared to the control strain, according to the RNA-seq and MS analyses. So that we speculated that *spxA1* is a newfound target gene of the moonlight protein NprR. At the same time, our result also indicated that the defects of *nprR* mutant in peroxide, diamide and salt deoxycholate resistance could be corrected by introducing spxA1 in plasmids and the downregulation of SpxA1 protein might be the direct reason that the *nprR* deletion mutant strain was more sensitive to oxidative stress. However, it remains unclear whether

NprR protein downregulated the expression of *spxA1* directly or indirectly. It is necessary to determine whether the signal polypeptide NprX plays an important role in this process. The biological mechanism underlying the omics data merits in-depth study in our future work.

Integrated transcriptomic and proteomic analyses also showed that most significantly enriched GO terms for intracellular proteins were associated with oxidative stress responses. We thus speculated that the mutant strain was more sensitive to oxidative and disulfide stress, and subsequent phenotypic experiments confirm this hypothesis. Specifically, the mutant strain was more sensitive to oxidants like hydrogen peroxide and deoxycholate. This sensitivity was reflected in the significantly inhibited growth of the mutant relative to the A16R and complemented strains when exposed to the same concentrations of hydrogen peroxide. The RNA-seq and MS analyses also suggested that the expression levels of peroxidase (AW20_RS26735) and catalase (katE, AW20_RS27305) were significantly downregulated in the mutant (**Figure 1**), thereby suggesting that degradation of hydrogen peroxide was impaired in the mutant strain. These results may explain why the mutant strain is more sensitive to hydrogen peroxide. Previous work showed that the analogous genes were also downregulated in *spxA1* deletion mutant and these mutant strains were more sensitive to hydrogen peroxide in *Streptococcus sanguinis* and *Listeria monocytogenes* than wild-type strains (Chen et al., 2012; Cesinger et al., 2020). The mechanism by which BprR protein regulates these genes remains unclear, and further work is still needed to provide an adequate explanation. The RNA-seq and MS analyses from this study will provide useful reference materials for future related oxidative stress responses research in *B. anthracis*.

NprR protein is a typical moonlight protein. It might also be involved in many other biological processes, not including regulation the expression of extracellular proteases. For instance, in the GO term and KEGG pathway analyses, several DEGs and DEPs involved secondary metabolic process, such as siderophore biosynthetic and metabolic process, were enriched (**Figures 2**, **4**). Siderophore is a secondary metabolite secreted by microorganisms. It serves primarily to transport iron across cell membranes. According to published work, *B. anthracis* requires siderophore biosynthesis for growth in macrophages and for mouse virulence (Cendrowski et al., 2004). Normally, siderophore biosynthesis occurs via two pathways: the non-ribosomal peptide synthetase (NRPS) pathway and the NRPS-independent siderophore (NIS) synthetase pathway (Carroll and Moore, 2018). Coincidentally, several DEGs involving non-ribosomal peptide biosynthetic process were also enriched in GO term enrichment analyses. These results indicated that the NprR protein might also play an important role in the NRPS-dependent siderophore biosynthesis of *B. anthracis*. Siderophore synthesis can also be regulated by oxidative stress in some microorganisms, including *B. anthracis* (Tindale et al., 2000; Lee et al., 2011; Adler et al., 2014). We speculate that there may be a connection between the downregulation of genes related to NRPS-dependent siderophore biosynthesis and the genes related to oxidative stress responses in the *nprR* mutant strain of *B. anthracis* and the corresponding mechanism and explanation of these results will be explored in greater depth in our next work.

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: ProteomeXchange [accession: PXD020800].

AUTHOR CONTRIBUTIONS

YW and CL designed the research. YW, NJ, HT, BW, QG, and XZ performed all the experiments. YW, NJ, and CL analyzed the data. YW wrote the manuscript. All authors reviewed the final manuscript.

FUNDING

This work was supported in part by the Beijing Academy of Agriculture and Forestry Sciences Youth Foundation (No. QNJJ202026).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.590851/full#supplementary-material

Supplementary Figure 1 | Summary of the RNA-seq analyses of the nprR deletion mutant and A16R strains. **(A)** Venn diagram, **(B)** distribution density analysis, and **(C)** correlation analysis showing the overlap in gene expression between the two strains.

Supplementary Figure 2 | GO term enrichment analysis of the extracellular DEPs. MF, molecular function. *$P < 0.05$.

Supplementary Figure 3 | Integrated GO term enrichment analysis of DEGs and DEPs for the extracellular protein fraction. *FDR < 0.05.

Supplementary Table 1 | The detail of differentially expressed genes (DEGs) between the *nprR* mutant and A16R strains.

Supplementary Table 2 | The detail of intracellular differentially expressed proteins (DEPs) between the *nprR* mutant and A16R strains.

Supplementary Table 3 | The detail of extracellular differentially expressed proteins (DEPs) between the *nprR* mutant and A16R strains.

Supplementary Table 4 | GO functional enrichment analysis data of the DEGs.

Supplementary Table 5 | Integrated analysis data of DEGs and DEPs (intracellular proteins).

Supplementary Table 6 | Integrated analysis data of DEGs and DEPs (extracellular proteins).

Supplementary Table 7 | The details of the DEGs and intracellular DEPs that had corresponding changes.

Supplementary Table 8 | The details of the DEGs and extracellular DEPs that had corresponding changes.

REFERENCES

Adler, C., Corbalan, N. S., Peralta, D. R., Pomares, M. F., de Cristobal, R. E., and Vincent, P. A. (2014). The alternative role of enterobactin as an oxidative stress protector allows *Escherichia coli* colony development. *PLoS One* 9:e84734. doi: 10.1371/journal.pone.0084734

Barendt, S., Birch, C., Mbengi, L., and Zuber, P. (2016). Evidence that oxidative stress induces spxA2 transcription in *Bacillus anthracis* sterne through a mechanism requiring SpxA1 and positive autoregulation. *J. Bacteriol.* 198, 2902–2913. doi: 10.1128/jb.00512-16

Barendt, S., Lee, H., Birch, C., Nakano, M. M., Jones, M., and Zuber, P. (2013). Transcriptomic and phenotypic analysis of paralogous spx gene function in *Bacillus anthracis* sterne. *Microbiologyopen* 2, 695–714. doi: 10.1002/mbo3.109

Cabrera, R., Rodriguez-Romero, A., Guarneros, G., and de la Torre, M. (2016). New insights into the interaction between the quorum-sensing receptor NprR and its DNA target, or the response regulator Spo0F. *FEBS Lett.* 590, 3243–3253. doi: 10.1002/1873-3468.12371

Carroll, C. S., and Moore, M. M. (2018). Ironing out siderophore biosynthesis: a review of non-ribosomal peptide synthetase (NRPS)-independent siderophore synthetases. *Crit. Rev. Biochem. Mol. Biol.* 53, 356–381. doi: 10.1080/10409238.2018.1476449

Cendrowski, S., MacArthur, W., and Hanna, P. (2004). *Bacillus anthracis* requires siderophore biosynthesis for growth in macrophages and mouse virulence. *Mol. Microbiol.* 51, 407–417. doi: 10.1046/j.1365-2958.2003.03861.x

Cesinger, M. R., Thomason, M. K., Edrozo, M. B., Halsey, C. R., and Reniere, M. L. (2020). *Listeria monocytogenes* SpxA1 is a global regulator required to activate genes encoding catalase and heme biosynthesis enzymes for aerobic growth. *Mol. Microbiol.* 114, 230–243. doi: 10.1111/mmi.14508

Chen, L., Ge, X., Wang, X., Patel, J. R., and Xu, P. (2012). SpxA1 involved in hydrogen peroxide production, stress tolerance and endocarditis virulence in *Streptococcus sanguinis*. *PLoS One* 7:e40034. doi: 10.1371/journal.pone.0040034

Chung, M. C., Jorgensen, S. C., Tonry, J. H., Kashanchi, F., Bailey, C., and Popov, S. (2011). Secreted *Bacillus anthracis* proteases target the host fibrinolytic system. *FEMS Immunol. Med. Microbiol.* 62, 173–181. doi: 10.1111/j.1574-695X.2011.00798.x

Chung, M. C., Popova, T. G., Millis, B. A., Mukherjee, D. V., Zhou, W., Liotta, L. A., et al. (2006). Secreted neutral metalloproteases of *Bacillus anthracis* as candidate pathogenic factors. *J. Biol. Chem.* 281, 31408–31418. doi: 10.1074/jbc.M605526200

Cohen-Gihon, I., Israeli, O., Beth-Din, A., Levy, H., Cohen, O., Shafferman, A., et al. (2014). Whole-genome sequencing of the nonproteolytic *Bacillus anthracis* V770-NP1-R strain reveals multiple mutations in peptidase loci. *Genome Announc.* 2:e00075-14.

Dubois, T., Faegri, K., Perchat, S., Lemy, C., Buisson, C., Nielsen-LeRoux, C., et al. (2012). Necrotrophism is a quorum-sensing-regulated lifestyle in *Bacillus thuringiensis*. *PLoS Pathog.* 8:e1002629. doi: 10.1371/journal.ppat.1002629

Frees, D., Brondsted, L., and Ingmer, H. (2013). Bacterial proteases and virulence. *Subcell. Biochem.* 66, 161–192. doi: 10.1007/978-94-007-5940-4_7

Haider, S., and Pal, R. (2013). Integrated analysis of transcriptomic and proteomic data. *Curr. Genomics* 14, 91–110. doi: 10.2174/1389202911314020003

Kajfasz, J. K., Ganguly, T., Hardin, E. L., Abranches, J., and Lemos, J. A. (2017). Transcriptome responses of *Streptococcus mutans* to peroxide stress: identification of novel antioxidant pathways regulated by Spx. *Sci. Rep.* 7:16018.

Kastrup, C. J., Boedicker, J. Q., Pomerantsev, A. P., Moayeri, M., Bian, Y., Pompano, R. R., et al. (2008). Spatial localization of bacteria controls coagulation of human blood by 'quorum acting'. *Nat. Chem. Biol.* 4, 742–750. doi: 10.1038/nchembio.124

Lee, J. Y., Passalacqua, K. D., Hanna, P. C., and Sherman, D. H. (2011). Regulation of petrobactin and bacillibactin biosynthesis in *Bacillus anthracis* under iron and oxygen variation. *PLoS One* 6:e20777. doi: 10.1371/journal.pone.0020777

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Ma, J., Chen, T., Wu, S., Yang, C., Bai, M., Shu, K., et al. (2019). iProX: an integrated proteome resource. *Nucleic Acids Res.* 47, D1211–D1217. doi: 10.1093/nar/gky869

Marshall, N. C., Finlay, B. B., and Overall, C. M. (2017). Sharpening host defenses during Infection: proteases cut to the chase. *Mol. Cell. Proteomics* 16, S161–S171. doi: 10.1074/mcp.O116.066456

Miyoshi, S., and Shinoda, S. (2000). Microbial metalloproteases and pathogenesis. *Microbes Infect.* 2, 91–98. doi: 10.1016/S1286-4579(00)00280-X

Mukherjee, D. V., Tonry, J. H., Kim, K. S., Ramarao, N., Popova, T. G., Bailey, C., et al. (2011). *Bacillus anthracis* protease InhA increases blood-brain barrier permeability and contributes to cerebral hemorrhages. *PLoS One* 6:e17921. doi: 10.1371/journal.pone.0017921

Neiditch, M. B., Capodagli, G. C., Prehna, G., and Federle, M. J. (2017). Genetic and structural analyses of RRNPP intercellular peptide signaling of Gram-positive bacteria. *Annu. Rev. Genet.* 51, 311–333. doi: 10.1146/annurev-genet-120116-023507

Nilsson, M., Jakobsen, T. H., Givskov, M., Twetman, S., and Tolker-Nielsen, T. (2019). Oxidative stress response plays a role in antibiotic tolerance of *Streptococcus mutans* biofilms. *Microbiology* 165, 334–342. doi: 10.1099/mic.0.000773

Perchat, S., Dubois, T., Zouhir, S., Gominet, M., Poncet, S., Lemy, C., et al. (2011). A cell-cell communication system regulates protease production during sporulation in bacteria of the *Bacillus cereus* group. *Mol. Microbiol.* 82, 619–633. doi: 10.1111/j.1365-2958.2011.07839.x

Perchat, S., Talagas, A., Poncet, S., Lazar, N., de la Sierra-Gallay, I. L., Gohar, M., et al. (2016). How quorum sensing connects sporulation to necrotrophism in *Bacillus thuringiensis*. *PLoS Pathog.* 12:e1005779. doi: 10.1371/journal.ppat.1005779

Pineda-Castellanos, M. L., Rodríguez-Segura, Z., Villalobos, F. J., Hernández, L., Lina, L., and Nuñez-Valdez, M. E. (2015). Pathogenicity of isolates of *Serratia Marcescens* towards Larvae of the Scarab *Phyllophaga Blanchardi* (Coleoptera). *Pathog.* 4, 210–228. doi: 10.3390/pathogens4020210

Pomerantsev, A. P., Pomerantseva, O. M., Moayeri, M., Fattah, R., Tallant, C., and Leppla, S. H. (2011). A *Bacillus anthracis* strain deleted for six proteases serves as an effective host for production of recombinant proteins. *Protein Expr. Purif.* 80, 80–90. doi: 10.1016/j.pep.2011.05.016

Popova, T. G., Millis, B., Bradburne, C., Nazarenko, S., Bailey, C., Chandhoke, V., et al. (2006). Acceleration of epithelial cell syndecan-1 shedding by anthrax hemolytic virulence factors. *BMC Microbiol.* 6:8. doi: 10.1186/1471-2180-6-8

Rice, A. J., Woo, J. K., Khan, A., Szypulinski, M. Z., Johnson, M. E., Lee, H., et al. (2016). Over-expression, purification, and confirmation of *Bacillus anthracis* transcriptional regulator NprR. *Protein Expr. Purif.* 125, 83–89. doi: 10.1016/j.pep.2015.08.030

Rocha, J., Flores, V., Cabrera, R., Soto-Guzman, A., Granados, G., Juaristi, E., et al. (2012). Evolution and some functions of the NprR-NprRB quorum-sensing system in the *Bacillus cereus* group. *Appl. Microbiol. Biotechnol.* 94, 1069–1078. doi: 10.1007/s00253-011-3775-4

Rocha-Estrada, J., Aceves-Diez, A. E., Guarneros, G., and de la Torre, M. (2010). The RNPP family of quorum-sensing proteins in Gram-positive bacteria. *Appl. Microbiol. Biotechnol.* 87, 913–923. doi: 10.1007/s00253-010-2651-y

Tindale, A. E., Mehrotra, M., Ottem, D., and Page, W. J. (2000). Dual regulation of catecholate siderophore biosynthesis in *Azotobacter vinelandii* by iron and oxidative stress. *Microbiology* 146(Pt 7), 1617–1626. doi: 10.1099/00221287-146-7-1617

Vieira, M. L., and Nascimento, A. L. T. O. (2016). Interaction of spirochetes with the host fibrinolytic system and potential roles in pathogenesis. *Crit. Rev. Microbiol.* 42, 573–587. doi: 10.3109/1040841X.2014.972336

Wang, Y. C., Jiang, N., Zhan, D. W., Tao, H. X., Yuan, S. L., Wang, P., et al. (2011). Surface display of the 20-kDa N-terminal fragment of anthrax protective antigen based on attenuated recombinant *Bacillus anthracis*. *World J. Microbiol. Biotechnol.* 27, 2575–2581. doi: 10.1007/s11274-011-0729-z

Wang, Y. C., Yuan, L. S., Tao, H. X., Jiang, W., and Liu, C. J. (2018). pheS* as a counter-selectable marker for marker-free genetic manipulations in *Bacillus anthracis*. *J. Microbiol. Biotechnol. Meth.* 151, 35–38. doi: 10.1016/j.mimet.2018.05.024

Wu, L. S., Guo, X. L., Liu, X. L., and Yang, H. (2017). NprR-NprX quorum-sensing system regulates the algicidal activity of *Bacillus* sp strain S51107 against bloom-forming *Cyanobacterium Microcystis aeruginosa*. *Front. Microbiol.* 8:1968. doi: 10.3389/Fmicb.2017.01968

Yang, H., Sikavi, C., Tran, K., McGillivray, S. M., Nizet, V., Yung, M., et al. (2011). Papillation in *Bacillus anthracis* colonies: a tool for finding new mutators. *Mol. Microbiol.* 79, 1276–1293. doi: 10.1111/j.1365-2958.2011.07519.x

Zouhir, S., Perchat, S., Nicaise, M., Perez, J., Guimaraes, B., Lereclus, D., et al. (2013). Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor NprR. *Nucleic Acids Res.* 41, 7920–7933. doi: 10.1093/nar/gkt546

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 January 2021
doi: 10.3389/fmicb.2021.625705

Mfd Affects Global Transcription and the Physiology of Stressed *Bacillus subtilis* Cells

*Holly Anne Martin[1], Anitha Sundararajan[2], Tatiana S. Ermi[1], Robert Heron[1], Jason Gonzales[3], Kaiden Lee[4], Diana Anguiano-Mendez[1], Faye Schilkey[2], Mario Pedraza-Reyes[5] and Eduardo A. Robleto[1]**

[1] School of Life Sciences, University of Nevada, Las Vegas, Las Vegas, NV, United States, [2] National Center for Genome Resources, Santa Fe, NM, United States, [3] West Career and Technical Academy, Las Vegas, NV, United States, [4] The College of Idaho, Caldwell, ID, United States, [5] Division of Natural and Exact Sciences, Department of Biology, University of Guanajuato, Guanajuato, Mexico

For several decades, Mfd has been studied as the bacterial transcription-coupled repair factor. However, recent observations indicate that this factor influences cell functions beyond DNA repair. Our lab recently described a role for Mfd in disulfide stress that was independent of its function in nucleotide excision repair and base excision repair. Because reports showed that Mfd influenced transcription of single genes, we investigated the global differences in transcription in wild-type and *mfd* mutant growth-limited cells in the presence and absence of diamide. Surprisingly, we found 1,997 genes differentially expressed in Mfd$^-$ cells in the absence of diamide. Using gene knockouts, we investigated the effect of genetic interactions between Mfd and the genes in its regulon on the response to disulfide stress. Interestingly, we found that Mfd interactions were complex and identified additive, epistatic, and suppressor effects in the response to disulfide stress. Pathway enrichment analysis of our RNASeq assay indicated that major biological functions, including translation, endospore formation, pyrimidine metabolism, and motility, were affected by the loss of Mfd. Further, our RNASeq findings correlated with phenotypic changes in growth in minimal media, motility, and sensitivity to antibiotics that target the cell envelope, transcription, and DNA replication. Our results suggest that Mfd has profound effects on the modulation of the transcriptome and on bacterial physiology, particularly in cells experiencing nutritional and oxidative stress.

Keywords: oxidative stress, protein oxidation, stationary-phase, transcription-coupled repair, *Bacillus*

OPEN ACCESS

Edited by:
Monika Glinkowska, University of Gdańsk, Poland

Reviewed by:
Mitsuo Ogura, Tokai University, Japan
Seav-ly Tran, INRAE Centre Jouy-en-Josas, France

***Correspondence:**
Eduardo A. Robleto
eduardo.robleto@unlv.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *03 November 2020*
Accepted: *07 January 2021*
Published: *28 January 2021*

Citation:
Martin HA, Sundararajan A, Ermi TS, Heron R, Gonzales J, Lee K, Anguiano-Mendez D, Schilkey F, Pedraza-Reyes M and Robleto EA (2021) Mfd Affects Global Transcription and the Physiology of Stressed Bacillus subtilis Cells. Front. Microbiol. 12:625705. doi: 10.3389/fmicb.2021.625705

INTRODUCTION

Short for Mutation frequency decline, Mfd is the transcription-coupling repair factor, which coalesces a stalled RNAP with the Nucleotide Excision Repair (NER) pathway to preferentially repair lesions in the template strand of actively transcribed genes before lesions in the coding strand or in non-actively transcribed genes (Hanawalt and Spivak, 2008). However, as early as the 1990s, evidence suggested that Mfd influenced phenotypes unrelated to transcription-coupled repair. Reports showed that Mfd affected carbon catabolite repression of operons (Zalieckas et al., 1998b). In addition, *in vitro* studies showed

that Mfd can facilitate repression of transcription by roadblock clearance at genes regulated by the global transcription regulator CodY (Belitsky and Sonenshein, 2011). Our lab demonstrated a role for Mfd in the expression of amino acid biosynthesis genes and protection from oxidative stress (Pybus et al., 2010; Martin et al., 2011, 2019). Of note, a recent report proposed that Mfd functions at hard-to-transcribe genes and affected gene expression and survival associated with toxin-antitoxin gene modules in *B. subtilis* (Ragheb et al., 2021).

In bacterial species other than *B. subtilis*, observations suggest that Mfd influences traits other than DNA repair. While in some organisms Mfd has the potential to increase the frequency of mutations conferring resistance to antibiotics (Han et al., 2008; Ragheb et al., 2019; Merrikh and Kohli, 2020), in other bacteria such as *H. pylori*, it can increase antibiotic sensitivity (Lee et al., 2009). In *Staphylococcus aureus*, inactivation of *mfd* resulted in decreased biofilm formation (Tu Quoc et al., 2007). Interestingly, *E. coli* cells can perform transcription-coupled repair through an Mfd-independent mechanism and the activity of UvrD (Epshtein et al., 2014). This Mfd-independent transcription-coupled repair has been postulated to occur in *B. subtilis* (Moreno-Del Alamo et al., 2020). Also, DNA lesions of oxidative nature modulate growth and stationary-phase associated Mfd-dependent mutagenic events in this Gram-positive microorganism (Leyva-Sánchez et al., 2020). Moreover, recent single-molecule resolution experiments showed that Mfd can translocate on undamaged DNA independently of its interactions with RNAP (Ho et al., 2018). Given the previous biochemical observations, and the different phenotypes in different bacterial species associated with Mfd, we hypothesized that Mfd, in addition to mediating transcription-coupled repair, modulates the cell transcriptome. More specifically, we tested whether Mfd affects the global transcription profile in stationary-phase *B. subtilis* cells and in conditions of diamide exposure. Diamide is an oxidizing agent that transforms protein thiol groups into disulfide bonds and subject cells to protein oxidation stress (Pother et al., 2009). Cells experiencing disulfide stress activate the oxidative and electrophile stress stimulon to repair and process cellular damage (Antelmann and Helmann, 2011). The activation of this stimulon is complemented and overlaps with regulons controlled by SigB (general stress response), PerR (peroxide stress), OhrR (organic peroxide stress), AdhR (formaldehyde stress), and YodB (disulfide stress) to produce factors that detoxify metabolic intermediates and reactive oxygen species, prevent protein damage, and protect DNA from accumulating lesions (Antelmann and Helmann, 2011).

Using RNA extracted from stationary-phase *B. subtilis* cultures exposed to either 0 or 1 mM of the protein oxidant diamide, we found that nearly half of the genes in the transcriptome and several biological functions were expressed differentially in the absence of Mfd. Pathway enrichment analysis indicated that the absence of Mfd dysregulates expression of genes affecting biological processes that include translation, endospore formation, and flagellar motility. This dysregulation associated with phenotypic changes in growth in a defined medium, motility, and sensitivity to antibiotics. In addition to identifying pathways affected by Mfd expression, we were also interested in investigating potential gene targets that work in concert with Mfd to confer protection against disulfide stress. Genetic interaction experiments showed that the effect of Mfd on the cellular response to diamide was complex. For example, disruption of *sodA*, which encodes superoxide dismutase, abrogated the ability of cells to survive exposure to diamide, but overexpression of *mfd* in the SodA$^-$ background restored it. Our genetic interaction assays uncovered, additive, epistatic, and suppressor effects on the response to disulfide stress.

In conclusion, the results presented here expand the roles of Mfd in the cell beyond transcription-coupled repair and suggest that this factor is a global modulator of transcription with profound effects on bacterial physiology and adaptation to stress.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

The parental strain, YB955, is a prophage- "cured" *B. subtilis* strain 168 derivative that contains the point mutations *metB5*, *hisC952*, and *leuC427*. *B. subtilis* strains employed in this study (**Table 1**) were routinely isolated on tryptic blood agar base (TBAB) (Acumedia Manufacturers, Inc., Lansing, MI, United States), and liquid cultures were grown in Penassay broth (PAB) (antibiotic medium 3, Difco Laboratories, Sparks, MD, United States) supplemented with 1X Ho-Le trace elements (*Methods for General and Molecular Bacteriology*. Washington, D.C.: American Society for Microbiology) (1994). When required, tetracycline (Tet; 10 $\mu g{\cdot}mL^{-1}$), spectinomycin (Sp; 100 $\mu g{\cdot}mL^{-1}$), ampicillin (Amp; 100 $\mu g{\cdot}mL^{-1}$), chloramphenicol (Cm; 5 $\mu g{\cdot}mL^{-1}$), erythromycin (Em; 1 $\mu g{\cdot}mL^{-1}$) or isopropyl-β-D-thiogalactopyranoside (IPTG; 1 mM) were added to media.

Construction of Mutant Strains

To construct single mutant strains, genomic DNA was isolated from the corresponding BKE (Bacillus Knockout Erythromycin collection; Koo et al., 2017) strains using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI). Of note, the BKE gene deletion constructs are designed to minimize functional interference on the flanking open reading frames. Isolated genomic DNA was then transformed into YB955 using the competence procedures for *Bacillus* described previously (Yasbin et al., 1975). Briefly, YB955 was grown to T_{90}, 90 min after the cessation of growth (stationary phase), in GM1 broth (0.5% dextrose, 0.1% yeast extract, 0.2% casein hydrolysate, essential amino acids 50 μg/mL, 1X Spizizen salt solution and then diluted 10-fold into GM2 broth (GM1 broth plus 50 μM $CaCl_2$, 250 μM $MgCl_2$). After 1 h of incubation at 37°C with aeration, genomic DNA (100 ng) was added. Cells were plated on TBAB containing 5 μg/mL erythromycin to select for the BKE allele. Transformants were confirmed by PCR.

To construct double mutant strains, genomic DNA from YB9801 (Mfd$^-$) was isolated and transformed into *B. subtilis* strains with single mutations as described above (**Table 1**). Cells were plated on TBAB containing 10 μg/mL tetracycline to select for the *mfd*$^-$ allele and 5 μg/mL erythromycin for maintenance

TABLE 1 | Strains and plasmids used in this study.

Strain name	Genotype	Reference or source
YB955	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS}	Sung and Yasbin, 2002
YB9801	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *mfd:tet*	Ross et al., 2006
PERM1134	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *mfd:tet amyE:pHS-mfd*	Ramirez-Guadiana et al., 2013
HAM800	*hisC952 metB5 leuC427* xin-1 SpβSENS *ohrR:erm*	This study
HAM801	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *sigB:erm*	This study
HAM802	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *perR:erm*	This study
HAM803	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *yodB:erm*	This study
HAM806	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *cysK:erm*	This study
HAM807	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ssuC:erm*	This study
HAM810	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *cysK:erm mfd:tet*	This study
HAM812	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ssuC:erm mfd:tet*	This study
HAM826	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *bshA:erm*	This study
HAM827	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *bshA:erm mfd:tet*	This study
HAM828	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *bshB1:erm*	This study
HAM829	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *bshB1:erm mfd:tet*	This study
HAM830	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *sodA:erm*	This study
HAM831	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *sodA:erm mfd:tet*	This study
HAM832	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *yodB:erm mfd:tet*	This study
HAM833	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ykuV:erm mfd:tet amyE:pHS-mfd*	This study
HAM834	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ohrR:erm mfd:tet amyE:pHS-mfd*	This study
HAM835	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *perR:erm mfd:tet amyE:pHS-mfd*	This study
HAM836	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *sodA:erm mfd:tet amyE:pHS-mfd*	This study
HAM837	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *yodB:erm mfd:tet amyE:pHS-mfd*	This study
JG001	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *polYB:erm*	This study
JG002	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ohrB:erm*	This study
JG003	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *bstA:erm*	This study
JG008	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ohrR:erm mfd:tet*	This study
JG009	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *perR:erm mfd:tet*	This study
JG010	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *sigB:erm mfd:tet*	This study
JG011	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *bstA:erm mfd:tet*	This study
JG013	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ohrB:erm mfd:tet*	This study
JG014	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *polYB:erm mfd:tet*	This study
KL101	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *cypC:erm*	This study
KL105	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *aldY:erm*	This study
KL201	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *aldY:erm mfd:tet*	This study
KL205	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *cypC:erm mfd:tet*	This study
BH001	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ykuV:erm*	This study
BH002	*hisC952 metB5 leuC427* xin-1 Spβ^{SENS} *ykuV:erm mfd:tet*	This study

of the BKE allele. Transformants were confirmed by PCR with specific oligonucleotide primers.

To construct the *mfd*-restored strains, BKE or PERM1134 DNA was isolated and transformed as described above. Of note, in this construct, the *mfd* gene is expressed from an IPTG-dependent promoter, and previous experiments showed that IPTG amendment results in restoration of Mfd functions to levels above those observed in the parent strain (Martin et al., 2019). Cells were plated on TBAB containing 100 μg/mL spectinomycin, 10 μg/mL tetracycline, and 5 μg/mL of erythromycin. Transformants were confirmed by PCR.

RNA Sequencing and Differential Gene Expression Analysis

Briefly, a single colony was used to start a 2-mL PAB overnight culture. The next morning 0.5 mL was used to start a 15 mL PAB culture. Cultures were grown in flasks containing PAB and Ho-Le trace elements with aeration (250 rpm) at 37°C until 90 min after the cessation of exponential growth [designated T_{90} (90 min after the time point in the culture when the slopes of the logarithmic and stationary phases of growth intercepted)]. Growth was monitored with a spectrophotometer measuring the optical density at 600 nm (OD_{600}). At T90, cultures were divided, and half were treated with 1 mM diamide, and incubated for another 2 h.

Total RNA from three biological replicates was harvested from cells differing in Mfd proficiency and treated or untreated with diamide, using the MP FastRNA Pro Blue Kit, and treated with DNase to remove residual DNA (Waltham, MA). Ribosomal RNA was removed by Ribo-Zero Magnetic Kit for Gram-Positive Bacteria, and the remaining RNA was then fragmented. The RNA samples were reverse transcribed into cDNA and sequenced. High quality sequence reads were generated using a HiSeq platform (2 × 150 bp read length) and aligned using HISAT2

(v 2.1.0) short read aligner to the latest version of reference in the Pubmed database (GCA_000009045.1). All sequence data were deposited to the NCBI SRA database under the bio project ID PRJNA673980. Read counts were generated using featureCounts (v1.6.2). Gene expression was quantified as the total number of reads uniquely aligning to the reference, binned by annotated gene coordinate. Differential gene expression and related quality control analyses was determined using the Bioconductor package DESEQ2. Normalization of raw read counts was performed by a scaling method implemented within DESEQ2 package, which accounts for differences in sequencing depth and composition. Differential expression of pairwise comparisons (of the different conditions) was assessed using the negative binomial test with a Benjamini–Hochberg false discovery rate (FDR) adjustment applied for multiple testing corrections.

Growth Assays in Complex and Defined Media

A single colony was used to start a 2-mL PAB (complex) overnight culture. To start cultures for the growth curve, the OD_{600} for each overnight culture was measured. Cells were diluted to an OD_{600} of 0.4 for each strain and replicate. Then in a 96-well flat-bottom plate, 200 μL of PAB, 10 μL of the diluted overnight cultures, and either 0 mM or 0.5 mM diamide were mixed. Complemented strains were supplemented with 1 mM IPTG to induce expression of *mfd*. The growth curve was incubated at 37°C with shaking on the Synergy HTX plate reader. Readings were taken every 5 min for 16 h. Each strain was replicated at least nine times. For the assay in defined medium, we used Spizizen medium with histidine, methionine, and leucine supplements (the test strains carry these three auxotrophic markers) (Spizizen, 1958; Sung and Yasbin, 2002).

Motility Assay

To test for flagellum-based motility, we conducted experiments in media containing 0.7 and 0.3% agar concentrations. We examined the spread of colonies of Mfd derivative strains on 0.7% TBAB containing IPTG and measured the colony diameter at 0, 3, and 6 days after inoculation. A single colony was used to start a 2-mL PAB overnight culture. The next morning, the OD_{600} for each overnight culture was measured. Cells were diluted to an OD_{600} of 0.1 for each strain and replicate. 10 μL of the cell dilution was spotted on the TBAB plates containing (0.7% agar) and incubated at 37°C lid-side up, as previously described (Patrick and Kearns, 2009). We spotted three strains per plate. At least three biological replicates were completed. Assays using 0.3% agar were conducted similarly. However, we measured swimming diameter 10 h after inoculation. Measurements were analyzed by ANOVA, the differences between means were tested using the LSD test at $P \leq 0.05$ and $P \leq 0.01$. See statistical analysis for more details.

Minimal Inhibitory Concentration Assays

BioMerieux antibiotic strips containing a gradient of concentrations were used to test the effects of Mfd on sensitivity to linezolid, ampicillin, rifampicin, trimethoprim, and daptomycin. Protocols for preparation of cultures and media were followed according to manufacturer's instructions.

Statistical Analysis

ANOVA was used to test for differences between means. When ANOVA indicated statistical significance between treatments, we used the Least Significant Difference (LSD) method at $P \leq 0.05$ or $P \leq 0.01$. ANOVA (complete randomized design) and LSD analyses were conducted using the IBM SPSS 27 software and the statistical package in the GraphPad Prism 9 graphing software. Mean comparisons were conducted in pairwise combinations, and statistically significant differences between any two means were denoted by assigning different letters. We assigned "a" to the means that were not significantly different from the mean with the highest value, "b" to means that were different from the "a" group, and so on. A similar analysis was used in the study that first documented the effect of Mfd on tolerance to diamide (Martin et al., 2019).

Pathway Analysis by Gene Enrichment

Gene enrichment or pathway analyses was performed using the ClueGO plug-in module of the Cytoscape software program, which annotates a list of genes to biological functions (gene ontologies) in a hierarchical way against an annotated genome (Bindea et al., 2009). Lists of genes down or up regulated in the absence of Mfd and in conditions of diamide exposure were used as input into Cytoscape and analyzed for overrepresented gene ontologies. Also, Kappa statistics, within the ClueGO plug-in, were calculated to link gene networks. For example, the list of genes that were downregulated mapped to 13 gene ontology terms and were grouped into 3 groups based on Kappa scores (translation, cell differentiation, and protein folding (**Supplementary Material**). These grouping results were obtained using the ClueGO plug-in of Cytoscape with the following settings: biological functions for gene ontologies, medium network specificity, GO tree interval with level 3 as minimum and level 8 as maximum, enrichment (right -sided hypergeometric test) with a pV value of 0.05 or less with the Bonferroni correction, and a GO term/pathway network connectivity (Kappa Score threshold) of 0.4.

RESULTS

Mfd Modulates Global Transcription in Stationary-Phase and During Disulfide Stress in *B. subtilis* Cells

We conducted transcriptomic analysis assays in stationary-phase cells untreated or treated with 1 mM diamide of the parental (YB955) and Mfd^- (YB9801) strains. We used 3 independent cultures for each condition, which totaled 12 independent observations. The overall results showed that sufficient depth of coverage was established with reads mapping to genes uniquely (mapping to a single location), that expression patterns between independent cultures clustered according to experimental conditions. The results also showed that almost half

of the genome was differentially regulated by Mfd and as affected by diamide exposure. This result was striking (**Figures 1A–C**). Interestingly, the pairwise comparison between the parent and the Mfd mutant in the absence of diamide exposure showed that mRNA levels of a significant number of genes were dysregulated in the absence of Mfd (**Figure 1C**). These results suggested that Mfd has a major impact on the transcription profile of the cell in stationary-phase conditions. Pairwise comparisons for changes in gene expression between the parent strain and the Mfd$^-$ mutant in untreated and diamide-treated cells are presented in **Supplementary Table S1**.

In the untreated condition, 1,066 genes were downregulated in the absence of Mfd. The values in fold change expression [$Log_{(2)}$ in these genes were from −0.39 to −3.55]. Pathway-enrichment analysis showed that genes in thirteen biological functions were downregulated in the absence of Mfd. Three major biological pathways were highlighted by Kappa scores: protein folding (Gene Ontology ID 006457), cell differentiation (GO ID 0030154-including spore synthesis and germination), and translation (GO ID 0006412) (**Supplementary Table S2** and **Supplementary Figure S4**). On the other hand, expression of 931 genes was upregulated, and fold expression (Log_2) values ranged from 0.42 to 7.78. Thirty-six biological functions were disproportionately upregulated in the absence of Mfd. These 36 biological functions were coalesced into 10 major groups by Kappa scores and include pentose metabolism (GO ID 0019321), cellular nitrogen and organonitrogen compound biosynthetic processes (GO ID 19001566, GO ID 004271), phosphorus metabolism (GO ID 0006793), toxin metabolism (GO ID 0009404), and transport of carbohydrates and organics (GO ID 0008643 and 0071702), ribonucleoside monophosphate and pyrimidine-containing compound biosynthesis (GO ID 0009156 and GO ID 0072528), and flagellum dependent motility (GO ID 0001539) (**Supplementary Table S3** and **Supplementary Figure S5**).

In the diamide treated condition, Mfd deficiency resulted in the down-regulation of 1,365 genes. Enrichment pathway analysis showed that genes in 23 gene ontologies were down regulated in the absence of Mfd. These biological functions were grouped by Kappa scores into the ones observed in the untreated condition, proteolysis (GO ID 0006508) and glutamine biosynthesis (GO ID 0009084), isoprenoid metabolism (GO ID 0006720), antibiotic metabolism (GO ID 0016999), carboxylic acid metabolism (GO ID 0019752), and protein metabolism (GO ID 0019538). The fold change in gene expression (Log_2) ranged from −0.4 to −6.1 (**Supplementary Table S4** and **Supplementary Figure S6**). In addition, the Mfd$^-$ background displayed up-regulation of 1,040 genes, the enrichment analysis showed 24 gene ontologies that included inosine monophosphate (IMP) biosynthesis (GO ID 0006188), cell projection organization (GO ID 0030030), flagellum motility and chemotaxis (GO ID 0001539, GO ID0006935), transmembrane and sodium transport (GO ID 0006814 and 0055085), and cellular nitrogen biosynthesis (GO ID 0044271) (**Supplementary Table S5** and **Supplementary Figure S7**). In summary, these results showed that Mfd affects many biological processes in stationary-phase *B. subtilis*.

Mfd Influences Sensitivity to Antibiotics, Growth in Defined Medium, and Motility

Transcriptomic results indicated that loss of Mfd caused dysregulation of global gene expression and prompted us to test for growth and other phenotypes associated with the biological functions that were highlighted by the pathway enrichment analysis (ribonucleoside phosphate biosynthesis, and motility). Also, 108 essential genes were dysregulated in Mfd$^-$ cells (**Supplementary Table S1**). The genes *liaF*, *liaR*, and *liaS*, which are activated during oxidative and cell-envelop stress (Radeck et al., 2017), were differentially expressed by the loss of Mfd. Mfd affected expression of *rpoB* and *rpoE*. These genes encode the β and Δ subunits of the RNA polymerase (Boor et al., 1995). Given these observations, and previous reports of Mfd effects on gene expression of amino acid biosynthesis, we were motivated to investigate the ability to grow on a defined medium (Spizizen medium) and measure minimal inhibitory concentrations (MIC) for ampicillin, linezolid (controls), trimethoprim (thymidine synthesis), daptomycin (cell membrane), and rifampicin (transcription) and in cultures of the parent, Mfd$^-$, and Mfd$^-$ carrying an *mfd*-overexpressing construct in the *amyE* chromosomal locus.

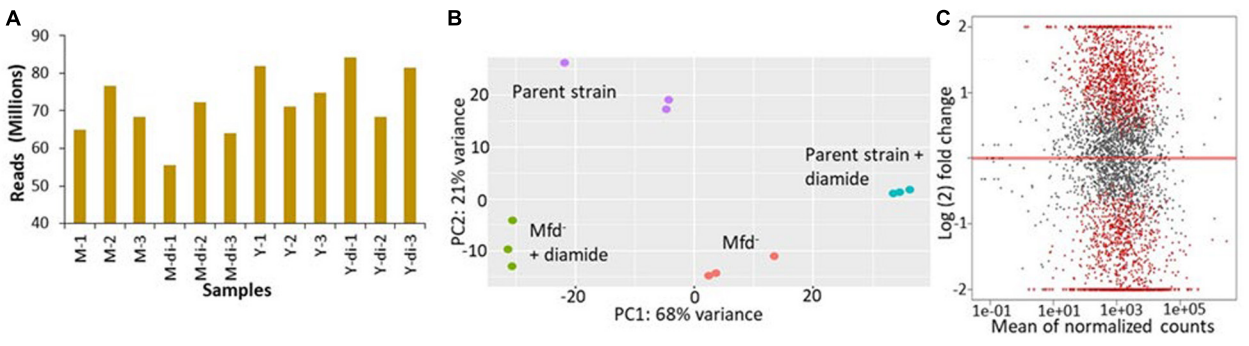


FIGURE 1 | Mfd alters the transcriptome in stationary-phase cells in the presence and absence of diamide. **(A)** Mapped sequencing reads generated by independent RNA samples from M: *mfd* mutant, Y9801; Y: YB955, the parental strain; and di: exposure to diamide. **(B)** Principal component analysis of gene expression results from independent samples (*n* = 3/strain and treatment combination). **(C)** Gene expression pairwise comparison between the parent, YB955 (reference) and the *mfd* mutant, YB9801, in untreated cells. Each dot represents fold change value in expression of a single gene. Red dots indicate significant change in expression.

We measured growth in defined medium for 16 h and showed similar values for doubling time, cell density, and growth lag between the parent and Mfd$^-$ cells (**Figure 2A**). However, the *mfd*-overexpressing strain showed marked differences in growth lag and cell density compared to the parent or Mfd$^-$ cells (**Figure 2A**). To better quantify overall growth dynamics, which integrate doubling time and cell density, we used the area under the curve, as calculated by the R-based program GrowthCurver (Sprouffske and Wagner, 2016), to measure the effect of Mfd on growth. The area under the curve (AUC) mean values indicated that the *mfd*-overexpressing cells displayed a significant increase in growth compared to the parent, and the parent's AUC was significantly higher than the Mfd$^-$ strain (**Figure 2B**).

All the strains were equally sensitive to ampicillin (lowest concentration—0.016 μg/ml) and resistant to linezolid (highest concentration—256 μg/ml). Interestingly, the loss of Mfd resulted in lower, but not significant, MIC values than the parent for trimethoprim and daptomycin (**Figures 2C,D**). However, overexpression of *mfd* produced significant increases in MIC values when compared to the Mfd$^-$ strain (∼2- and ∼3-fold for trimethoprim and daptomycin, respectively). On the other hand, Mfd$^-$ cells did not show significant differences to rifampicin when compared to the parent, but the cultures that overexpressed *mfd* were more sensitive than the parent (**Figure 2E**).

Cells lacking Mfd showed upregulation of 15 genes (*flgC, flgK, flgL, flhO, flhP, fliE, fliI, fliJ, fliK, fliL, fliM, fliY, hag, motA, and motB*—**Supplementary Table S1**) controlling flagellum-based motility, which prompted us to test if there was a difference in motility-associated phenotypes of *B. subtilis* differing in Mfd. Interestingly, Mfd$^-$ cells upregulated the expression of *srfAB*, *srfAC*, *srfAD* (**Supplementary Table S1**), which contain genes for surfactin biosynthesis and factors required for genetic competence (Dubnau, 1991; Nakano et al., 1991; Cosmina et al., 1993; Fuma et al., 1993). We tested Mfd derivatives on TBAB plates containing 0.7% and 0.3%, and incubated at 37°C. Colonies lacking Mfd displayed increased spread after 3 days, and the Mfd-defective cells were significantly increased in colony spread 6 days after inoculation on 0.7% agar compared to the parent cells and those that overexpressed *mfd* (**Figure 3**). The swimming motility assay showed significant differences between strains. The *mfd* mutant showed a diminished ability to swim compared to the parent strain, albeit not significantly. However, *mfd*

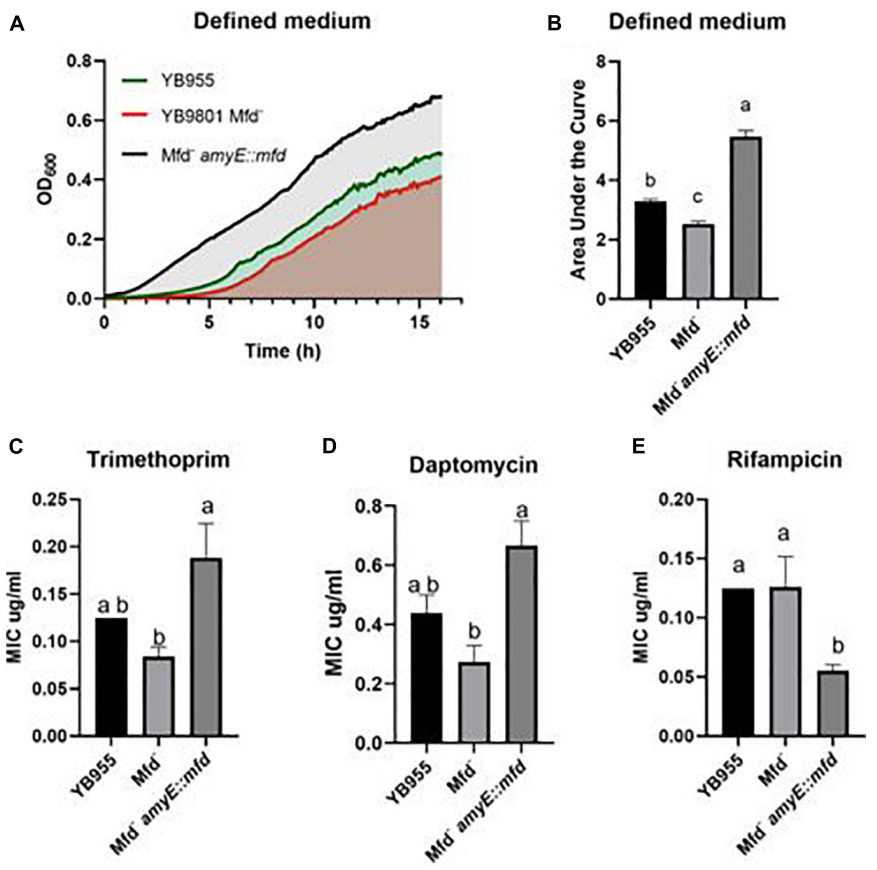


FIGURE 2 | Effects of Mfd on growth in defined medium and minimal inhibitory concentrations MIC to antibiotics. **(A)** Average OD_{600} readings of strains differing in Mfd measured by a plate reader (n = 9). **(B)** Average area under the curve as calculated by the R package, Growthcurver, of growth curves from **(A)**. **(C–E)** Average MIC of strains differing in Mfd determined by ETEST strips (n = 3). AUC was analyzed by ANOVA, and means were separated using the Least Significant Difference (LSD) test. Lower case letters distinguish significant differences between means. "a," "b," "c" and onward are significantly different mean groups at $P \leq 0.05$. Bars on top the columns represent SEM.

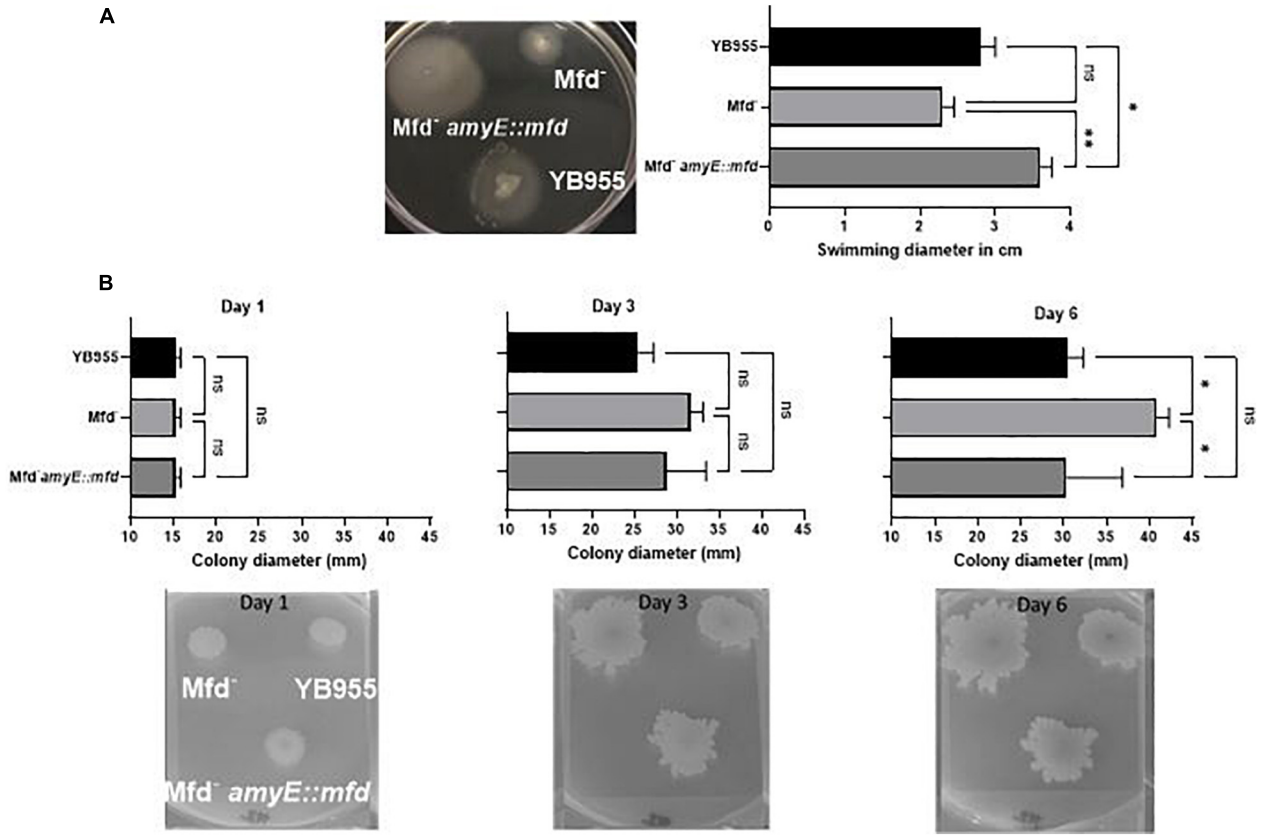


FIGURE 3 | Deficiencies in Mfd reduce swimming motility and increase colony spread. **(A)** Effect of Mfd on swimming motility. Mfd derivatives were spot inoculated on 0.3% TBAB agar (n = 3). Swimming diameter was measured 10 h after inoculation analyzed by ANOVA; means were separated using the LSD test. **(B)** Effect of Mfd on colony spread on 0.7% TBAB agar. ns, *, ** denote no significant differences, different at $P \leq 0.05$, and $P \leq 0.01$, respectively.

overexpression resulted in a significant increase in swimming motility compared to the parent and the *mfd* mutant. Altogether, these results suggest that Mfd affects fundamental aspects of bacterial physiology.

Mfd Modulates the Response to Disulfide Stress via Complex Genetic Interactions in *B. subtilis*

In addition to the experiments measuring growth in defined medium, sensitivity to antibiotics, and motility, we examined the effect of Mfd in stressed cells and in the context of protein oxidation. Our previous report showed that Mfd protected cells against disulfide stress, but not much is known about how this factor operates to provide cells such protection (Martin et al., 2019). Growth in complex broth in the absence and presence of diamide was followed by OD_{600} measurements every 5 min for 16 h. The growth dynamics of the parent and Mfd$^-$ strains were indistinguishable in the untreated cells, and both genetic backgrounds showed a similar growth lag in the presence of diamide (**Figure 4A**). However, the Mfd$^-$ cells displayed an increase in doubling time during mid exponential growth and a decrease in cell density compared to the parent strain in the presence of diamide. Cells that overexpressed *mfd* displayed the same growth rate but an increase in cell density compared to the parent strain in untreated conditions (**Figure 4A**). Strikingly, the *mfd*-overexpressing cells showed a shorter growth lag and a higher cell density than the parent strain, but both strains showed a similar doubling time in the treated cells (**Figure 4A**). The AUC values showed that growth in untreated cells was significantly different among the tested strains and that the *mfd*-overexpressing cells responded better than the parent and Mfd$^-$ counterparts ($P \leq 0.05$) (**Figure 4B**). Overexpressing *mfd* increased tolerance to diamide levels significantly better than in the parent and Mfd$^-$ counterparts. In fact, the response in the *mfd*-overexpressing strain was like the one observed in the untreated parent strain. The Mfd$^-$ cells displayed the lowest tolerance to diamide amongst the tested strains (**Figure 4B**). The results observed in untreated and diamide-treated cells suggest that Mfd controls the physiology of cells experiencing nutrient-limiting conditions and the response to disulfide stress.

To better understand the role of Mfd in the response to disulfide stress, we examined the genetic interactions between the Mfd factor and those genes that are either affected in expression by Mfd, are known transcription factors that control gene expression during oxidative damage, or are components of cysteine and sulfur metabolism, which produce

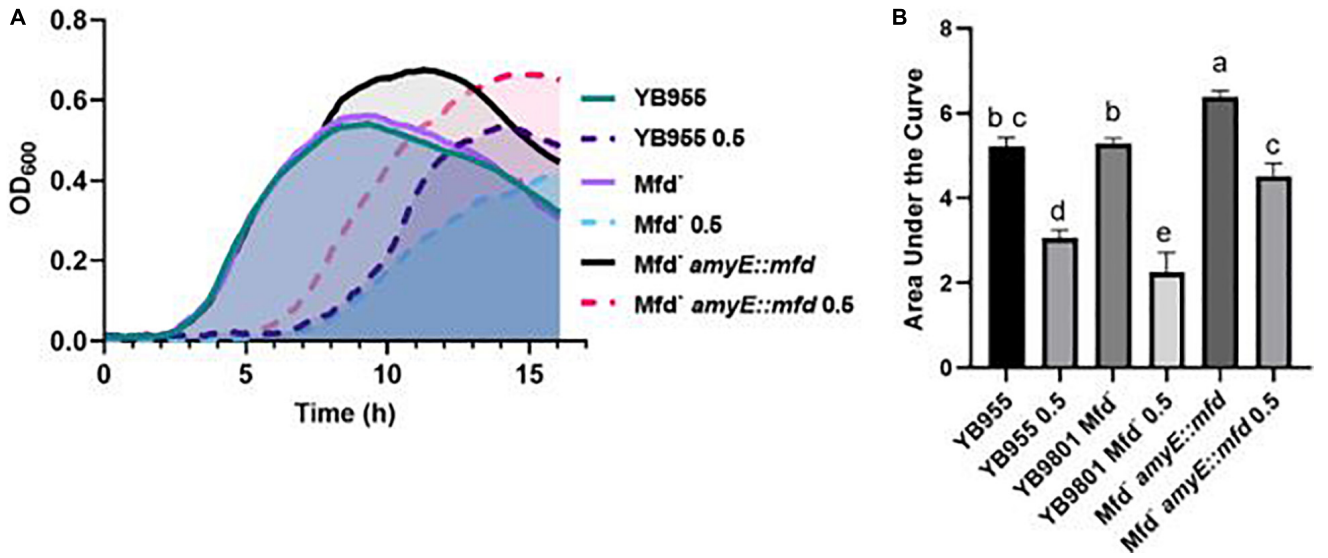


FIGURE 4 | Effects of Mfd on growth in a complex medium and in the presence of diamide. **(A)** Average OD_{600} readings of strains differing in Mfd measured by a plate reader (n = 9). **(B)** Average area under the curve (AUC) as calculated by the R package, Growthcurver, of growth curves from **(A)**. AUC was analyzed by ANOVA, and means were separated using the Least Significant Difference (LSD) test. Lower case letters distinguish significant differences between means. "a," "b," "c" and onward are significantly different mean groups at $P \leq 0.05$. Bars on top the columns represent SEM.

compounds important for disulfide and electrophile stress tolerance (Hochgräfe et al., 2007; Gaballa et al., 2010). We selected 15 genes to generate single knockouts in our parental strain, YB955 (**Table 1** and **Supplementary Table S6**). Then, we produced double knockouts of *mfd* and each of the 15 selected genes by transforming genomic DNA from our *mfd* mutant strain, YB9801, into the single knockout. Also, for 5 of the 15 genes under study, we generated strains in which the double knockout was transformed with genomic DNA from our *mfd*-restored strain, PERM1134. This process generated three types of strains: (i) single gene knockouts, (ii) double knockout combinations of *mfd* and each of the 15 genes, and (iii) double knockouts with an *mfd* overexpressing construct ectopically placed in the chromosome.

We conducted growth assays in the presence and absence of 0.5 mM of diamide and used the area under the curve to measure the growth response and tolerance to diamide in all the tested strains. Strains with single mutations in *bshA, bshB1*, and *ohrB* showed a similar growth response and tolerance to diamide to the parent (**Figures 5A–C** and **Supplementary Figure S1**). The *bshA*, and *bshB1* gene products are components of the bacillithiol biosynthesis, and *ohrB* encodes a peroxiredoxin (Völker et al., 1998; Fuangthong et al., 2001; Gaballa et al., 2010). These results suggest that the single contributions of these factors to disulfide stress tolerance are negligible (**Figures 5A–C**). However, the mean response in tolerance to diamide in cells with mutations in *mfd* and each of these three factors were statistically the same as the one observed in the parent. These results suggest that mutations in *bshB1*, *bshA*, and *ohrB* suppress the effect of *mfd* on disulfide stress.

Mutations in *cypC*, *bstA*, and *aldY* did not affect growth, but exerted increased tolerance to diamide when compared to the parent or Mfd mutant (**Figures 5D–F** and **Supplementary Figure S1**). The factors coded by these genes are involved in synthesis of lipopeptides and bacillithiol, and tolerance to alcohol exposure (Matsunaga et al., 1999; Petersohn et al., 1999; Perera et al., 2018). Interestingly, the levels of increased tolerance to diamide in the CypC$^-$ and BstA$^-$ backgrounds were unaffected by mutations in *mfd*. These results suggest that the loss of these factors alters the cell physiology to better withstand disulfide stress and exert an epistatic effect on *mfd*. Mutations in *aldY* did not affect growth but increased tolerance to diamide compared to the parent. In fact, there was no difference between treated and untreated AldY$^-$ cells. The combined mutations in *mfd* and *aldY* resulted in diminished growth compared to the parent or *aldY* mutant, suggesting that both functions are required for optimal growth. Interestingly, the absence of both gene products showed similar AUC values in treated (Mfd$^-$ AldY$^-$ 0.5) and untreated cells (Mfd$^-$ AldY$^-$), indicating that deficiencies in AldY desensitize cells to diamide (**Figure 5F**).

We conducted experiments that included overexpression of the Mfd factor in strains with mutations in transcription factors that affect the response to either oxidative or disulfide stress, as well as in strains with mutations in genes coding for a thiol-oxidoreductase and a superoxide dismutase (Inaoka et al., 1998; Fuangthong et al., 2001; Zhang et al., 2006). Mutating *ohrR* did not affect growth, but decreased tolerance to diamide treatment to the levels observed in the Mfd$^-$ background (**Figure 6A** and **Supplementary Figure S2**). The results shown by the double inactivation of *ohrR* and *mfd* suggest that these genes contribute to diamide additively and, therefore, are part of different cellular pathways (**Figure 6A**). Restoring *mfd* with an overexpressing construct in the *ohrR mfd* mutant resulted in increased tolerance to diamide levels higher than the ones observed in the parent

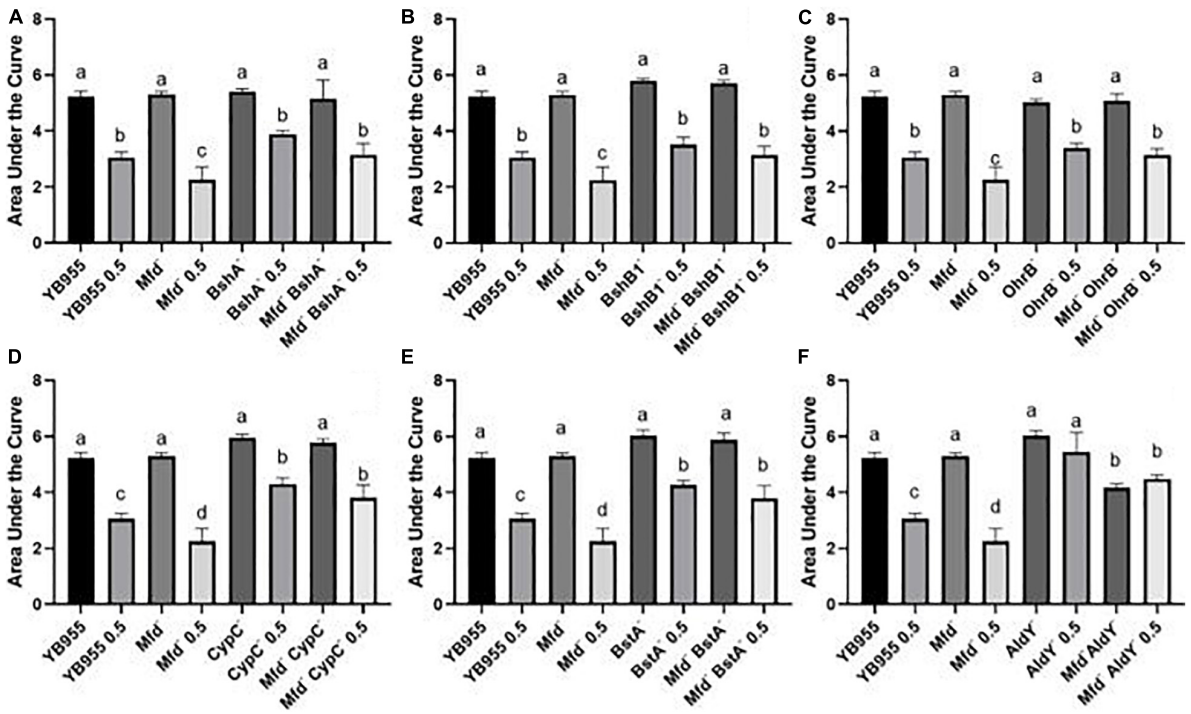


FIGURE 5 | Growth dynamics of single- and double-gene-deletion strains in the presence and absence of diamide show gene suppressors and recessive interactions. Average area under the curve as calculated by the R package, Growthcurver, of growth curves from **Supplementary Figure S1**. Each graph represents an independent trial of single and combined mutations of **(A)** *bshA*, **(B)** *bshB1*, **(C)** *ohrB*, **(D)** *cypC*, **(E)** *bstA*, and **(F)** *aldY*, and *mfd*, and each AUC mean is based on $n = 9$. AUC was analyzed by ANOVA, and means were separated using the Least Significant Difference (LSD) test. Lower case letters distinguish significant differences between means. "a," "b," "c" and onward are significantly different mean groups. Bars on top the columns represent SEM.

(**Figure 6A**). Deficiencies in YodB, a repressor that controls expression of genes that respond to disulfide stress (Chi et al., 2010), did not affect growth, but displayed decreased tolerance to diamide levels comparable to the Mfd$^-$ strain. However, the combined effects of deficiencies in Mfd and YodB resulted in increased growth and tolerance to diamide, and overexpressing *mfd* in this background sensitized cells to it (**Figure 6B** and **Supplementary Figure S2**). Inactivation of *perR*, the gene that encodes the repressor that controls the response to peroxide (Bsat et al., 1998), increased growth but showed no differences in tolerance to protein oxidation compared to the parent strain. The response in the *mfd perR* strain suggests that a component of the increase in growth seen in the Per$^-$ cells is dependent on Mfd; restoring this factor restored growth to the levels seen in the Per$^-$ cells but did not change the response to diamide compared to the double mutant (**Figure 6C** and **Supplementary Figure S2**).

Genetically inactivating SodA increased growth but abrogated tolerance to protein oxidation, and a similar response was observed when this mutation combined with Mfd deficiencies in treated and untreated cells (**Figure 6D** and **Supplementary Figure S2**). So, the phenotype caused by the loss of function in SodA persisted in Mfd$^+$ cells. This result suggests that the *sodA* gene has a recessive epistatic effect onto *mfd*. However, tolerance to diamide was restored to levels higher than the parent in the *sodA mfd* strain that overexpressed *mfd*. This indicates that excess Mfd in the cell can remodel physiology to bypass the absence of SodA. Interestingly, cells that lack YkuV, which encodes a thiol oxidoreductase, showed an increase in growth and tolerance to diamide compared to the parent, and the values in diamide tolerance seen in the double mutant (*mfd ykuV*) suggest a suppressor effect onto the *mfd* mutation. Overexpression of *mfd* restored growth to the parent levels and increased tolerance to diamide significantly compared to the double mutant (**Figure 6E** and **Supplementary Figure S2**). This result suggests that excess Mfd can function to increase tolerance to diamide independently of the suppressor effect exerted by the loss of YkuV (compare the means of the double mutant and the YkuV$^-$ mutant overexpressing *mfd* in the presence of diamide—**Figure 6E**).

The experiments examining the pairwise interactions between *mfd* and *sigB* (general stress response; Petersohn et al., 1999), *ssuC* (aliphatic sulfonate transporter; Van Der Ploeg et al., 1998), *polYB* (translesion synthesis polymerase; Duigou et al., 2004), and *cysK* (cysteine synthase; Van Der Ploeg et al., 2001); upregulated in cells exposed to diamide (Leichert et al., 2003) showed that single mutations in the latter four genes resulted in an increase in growth and tolerance to protein oxidation (**Supplementary Figure S3**). The double mutants *mfd sigB*, *mfd polY*, and *mfd ssuC* showed a similar response to the ones observed in the Mfd$^+$ background, which suggests that strains with deficiencies in either SigB, PolYB, or SsuC respond to disulfide stress independently of Mfd. In contrast, the double mutant *mfd*

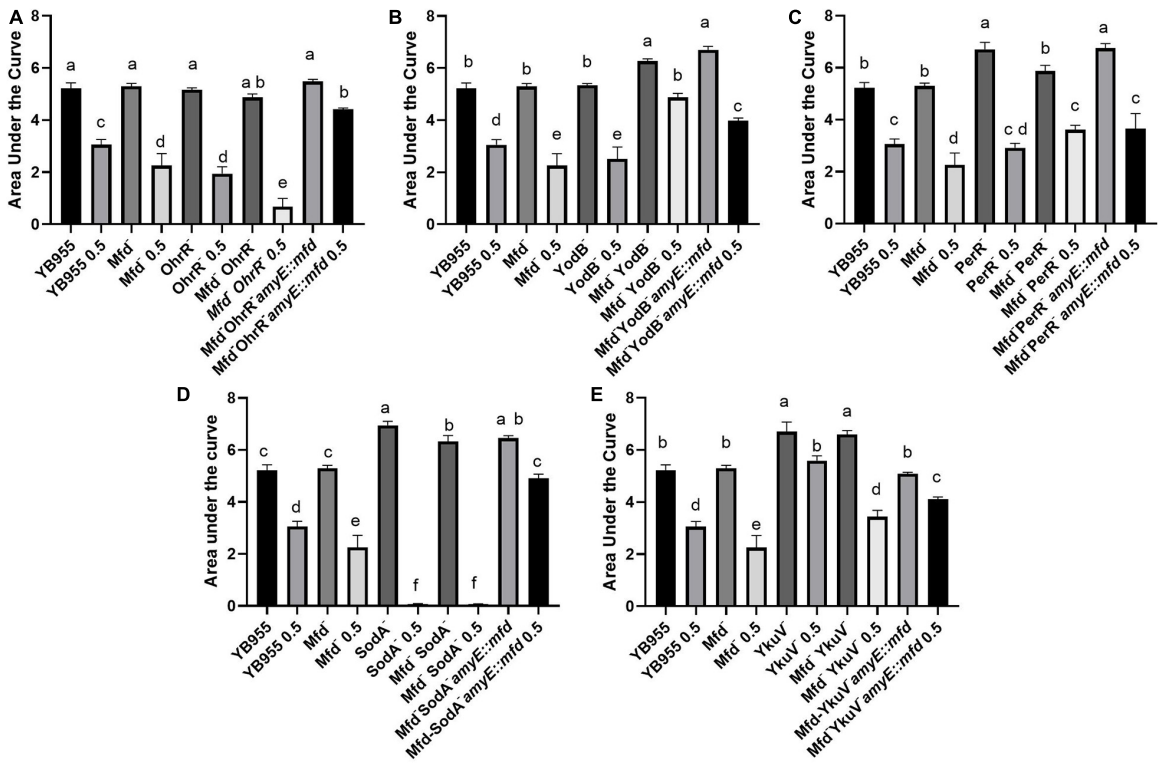


FIGURE 6 | Overexpression of *mfd* can rescue or sensitize cells to diamide. Growth dynamics of single- and double-gene-deletion and overexpressing-*mfd* strains in the presence and absence of diamide. Average area under the curve as calculated by the R package, Growthcurver, of growth curves from **Supplementary Figure S2**. Each graph represents an independent trial of single and combined mutations of **(A)** *ohrR*, **(B)** *yodB*, **(C)** *perR*, **(D)** *sodA*, and **(E)** *ykuV*, and *mfd*. Each AUC mean is based on $n = 9$. AUC was analyzed by ANOVA, and means were separated using the Least Significant Difference (LSD) test. Lower case letters distinguish significant differences between means. "a," "b," "c" and onward are significantly different mean groups. Bars on top the columns represent SEM.

cysK showed a decreased growth and tolerance to diamide compared to the CysK$^-$ cells but expressed similar values in treated and untreated conditions (**Supplementary Figure S3**). This result suggests that the combined deficiencies in CysK and Mfd decrease growth but do not influence overall tolerance to diamide. Altogether, these results indicate that the effects of Mfd on the growth and the response to protein oxidation is complex and dependent on gene interactions, and that this factor can remodel cell physiology to sensitize or increase tolerance to disulfide stress.

DISCUSSION

Mfd affected expression of almost 2,000 genes in stationary-phase cells and several biological functions, even in the absence of disulfide stress. There are no reports on the effects of Mfd on overall cell physiology on stationary-phase bacterial cells. Mfd did not affect the growth dynamics in a complex medium compared to the parental strain (**Figure 4**). Nevertheless, Mfd$^-$ cells showed differences in expression of 108 essential genes (83 were down regulated in the absence of Mfd—see **Supplementary Table S1**) that affect DNA replication, cell division, transcription, protein synthesis, and other biological functions. Given the changes in expression on essential genes, whose products are targeted by antibiotics, we tested Mfd derivatives for sensitivity to antibiotics that target cell envelope, transcription, and DNA replication. Also, previous reports on the effect of Mfd on gene transcription controlled by amino acid starvation (Gerhardt et al., 1994; Belitsky and Sonenshein, 2011) prompted us to test the different Mfd backgrounds for the ability to grow in a defined medium, in the absence of amino acid supplements. Mfd defects did not result in significant changes in MIC for trimethoprim, daptomycin, and rifampicin but its overexpression did (**Figure 2**). The results in growth in defined medium and antibiotic sensitivity support the idea that the Mfd-dependent changes in gene transcription alter cell physiology.

In the absence of Mfd, 65% of the genes for inosine monophosphate metabolism were upregulated. Other genes involved in nucleoside metabolism were affected by Mfd. These changes in gene expression are likely to induce a more sensitive state to trimethoprim; such sensitivity was reduced when *mfd* was overexpressed. Interestingly, activity of the ribonucleotide reductase, encoded by *nrdEF*, influences Mfd-dependent mutagenesis (Castro-Cerritos et al., 2017, 2018). The response in daptomycin correlated with low gene expression of the *liaRS* operon, which codes for factors that sense cell-envelope stress. It is worth noting that previous studies have implicated Mfd in the formation of mutations that confer resistance to antibiotics (Han et al., 2008; Ragheb et al.,

2019); however, no reports associate Mfd-dependent changes in gene expression to antibiotic sensitivity. The effect of Mfd on rifampicin MIC is likely due to the changes in expression in *rpoB* and *rpoE* in the Mfd$^-$ cells (**Supplementary Table S1**). Mfd is active during transcription elongation and at regions distant from the promoter (Haines et al., 2014) and processes backtracked RNAP into active transcription (Park et al., 2002). Perhaps, excess Mfd alters the equilibrium between the RNAPs engaged in initiation and elongation, and such alteration results in the increased sensitivity observed here.

Our assays showed that the absence of Mfd influenced motility (**Figure 3**). Our results on the colony spread on 0.7% were interesting. Our expectation was that Mfd derivatives would not differ in 0.7% agar because we worked with a domesticated strain; domesticated strains do not display swarming motility (Kearns et al., 2004; Patrick and Kearns, 2009), which is what we observed. However, we did note differences in colony spread after 6 days, suggesting that colony properties are affected by Mfd. The results from the RNASeq indicated that 15, but not all, structural genes for flagellum biosynthesis were upregulated in the absence of Mfd. Strikingly, Mfd$^-$ cells were diminished in swimming. Flagellum biosynthesis is regulated in a complex manner in *B. subtilis* (Kearns et al., 2004; Mukherjee and Kearns, 2014), and the disparate upregulation of structural genes could have compromised the ability to swim in Mfd$^-$ cells. To our knowledge, this is the first report that shows Mfd to be a factor that controls motility. Forty percent of genes involved in endospore formation were downregulated and suggest that the loss of Mfd affects sporulation. This assertion agrees with our previous report (Ramirez-Guadiana et al., 2013) and that of Koo et al. (2017) which showed that Mfd$^-$ deficient cells display a 65% sporulation efficiency compared to Mfd$^+$ cells. The decrease on sporulation efficiency could be the result of the combined deficiencies in gene expression and repair of oxidative damage during endospore formation in Mfd$^-$ cells. Mfd$^-$ cells with defects in DNA repair systems that target 8-oxo-G showed an exacerbated decrease in sporulation efficiency during oxidant exposure compared to the Mfd$^+$ cells (personal communication Pedraza-Reyes).

The loss of Mfd significantly decreased growth in defined medium, and overexpressing of this factor produced an increase in the growth response in defined medium compared to the parent (**Figure 2**). On the other hand, growth in PAB (complex medium) showed no differences between the parent and the Mfd$^-$ cells, but overexpression of Mfd conferred PAB cultures the ability to attain a higher cell density than the parent (**Figure 4**). These results suggest that Mfd facilitates gene expression of amino acid biosynthetic genes and a more efficient growth physiology in defined medium; in complex medium, Mfd-dependent changes in physiology allowed cells to continue doubling during the transition to nutrient scarcity. Strains containing gene deletions in *mfd* showed a slight growth defect in LB at 37°C, as measured by relative fitness (Koo et al., 2017). The results from our growth assays suggest that the biological consequences associated with deficiencies in Mfd express in cells nutritionally stressed or conditions in which amino acid biosynthesis is active.

The effects of Mfd on the overall response to diamide exposure were complex. In untreated conditions, several biological functions were affected by Mfd, and there were no consequences in the ability to grow on PAB, but it could be argued that the changes in gene expression observed in the untreated condition sensitize cells to protein oxidation. Contrastingly, the gene expression associated with the cellular functions that were affected in the absence of Mfd in treated cells only (protein degradation, glutamine, isoprenoid and carboxylic acid metabolism, antibiotic metabolism, inosine monophosphate biosynthesis, cell projection and organization, as well as transmembrane and sodium transport) are likely to compromise the response to disulfide stress. Future research will examine the specific contributions of those biological processes to sensitivity and adaptation to diamide exposure.

We studied the effects of Mfd on tolerance to protein oxidation further by measuring how genetic interactions influenced the cellular response to disulfide stress (**Figures 5**, **6** and **Supplementary Figure S3**). The results indicated different kinds of interactions, and the strains with single mutations in *bshA*, *bshB1*, *ohrB*, *cypC*, *bstA*, *sigB*, and *polYB* showed no significant differences compared to the strains with double mutations in *mfd bshA*, *mfd bshB1*, *mfd ohrB*, *mfd cypC*, *mfd bstA*, *mfd sigB*, and *mfd polYB*. Then, the effects on growth in a complex medium and tolerance to diamide caused by the single mutations in these genes (*bshA*, *bshB1*, *ohrB*, *cypC*, *bstA*, *sigB*, and *polYB*) are Mfd-independent. Contrastingly, the effects of mutations in *aldY*, *ohrR*, *yodB*, *perR*, *sodA*, *ykuV*, *ssuC*, and *cysK* on growth in PAB (*aldY*, *yodB*, *perR*, *sodA*, and *cysK*) or tolerance to diamide (*aldY*, *ohrR*, *yodB*, *ykuV*, and *cysK*) were partially dependent on Mfd. The effect of Mfd overexpression on growth and tolerance to diamide was dependent on the genetic background tested (**Figure 6**). Interestingly, overexpression of *mfd* combined with mutations in *ohrR* and *sodA* resulted in increased tolerance to disulfide stress when compared to Mfd$^-$ OhrR$^-$ and Mfd$^-$ SodA$^-$ cells, respectively. We interpreted the *mfd*-overexpression results to mean that Mfd can function to modulate the response to diamide by different pathways.

Mfd is a multidomain enzyme that can translocate on DNA independently of its direct interactions with RNAP, and the Mfd number of molecules is estimated to range between 30 and 300 (Deaconescu et al., 2006; Ho et al., 2018; Le et al., 2018). Then, it is likely that a large fraction of the ∼2,000 genes affected by Mfd reported here are caused by indirect effects. However, the experiments that tested the effects of gene interactions and *mfd* overexpression on tolerance to diamide provide candidates for genes and cellular pathways whose expression are directly affected by Mfd. For example, tolerance to diamide in the *sodA* and *mfd sodA* mutants was abrogated but overexpression of *mfd* rescued it, which suggests an Mfd-dependent effect. Direct effects on genes can be identified by combining transcriptome experiments and assays that test Mfd interactions with the transcription machinery (NETSeq or ChIPSeq) (Ragheb et al., 2021). Similar assays were used recently to characterize transcriptional pauses as affected by NusG (Yakhnin et al., 2020). Mechanistically, *in vitro* studies have shown that Mfd can function as a transcription modulator

(Selby and Sancar, 1995; Belitsky and Sonenshein, 2011; Le et al., 2018). Also, mutagenesis reports suggest that Mfd acts at highly transcribed regions that accumulate lesions and mediates the formation of mutations (Gomez-Marroquin et al., 2016). In the presence of Mfd, active transcription at catabolite-repressed genes is terminated at DNA sites distant from the promoter region and occupied by the *cre*-CcpA complex (DNA-protein repression block) (Zalieckas et al., 1998a). This mode of transcription termination operates at genes controlled by the CodY repressor (Belitsky and Sonenshein, 2011). On the other hand, if Mfd is required to process RNAP paused at intrinsic sites or pauses caused by non-B DNA structures back into active transcription, then loss of Mfd would result in a decrease of complete gene transcripts. *In vitro* studies demonstrated that non-B DNA structures can stall transcription (Tornaletti et al., 2008; Pandey et al., 2015).

A recent genome association report showed that expression of almost 380 genes were affected by Mfd during exponential growth; genes that were directly affected by this factor encode toxin-antitoxin modules (Ragheb et al., 2021). Further, those studies showed that overexpression of the *txpA*, *bsrH*, genes in Mfd^- cells decreased cell survival compared to the parent strain and led to propose a model in which Mfd directly interacts with RNAP to modulate transcription of regions with structured RNAs (Ragheb et al., 2021). Interestingly, our transcriptome assays were conducted in stationary-phase cells and showed no changes in expression of *txpA* and *bsrH* genes in untreated cells. However, Mfd^- cells treated with diamide were downregulated in transcription of those genes [–1.14 $Log_{(2)}$ for *bsrH* and –1.79 for *txpA*], which suggests that Mfd-dependent modulation of transcription changes according to growth conditions.

CONCLUSION

The results reported here and those in exponentially growing cells provide strong evidence for the concept that Mfd acts as global modulator of transcription and that the biological consequences associated with the loss of this factor manifest in stressed cells. From a mechanism standpoint, Mfd resolves pauses or blocks to transcription by dissociating a halted RNAP from the DNA or reactivating it into active transcription (Park et al., 2002; Deaconescu et al., 2006; Savery, 2007; Ho et al., 2018). Therefore, the loss of Mfd can increase or decrease the ability of an RNAP to complete a full transcript directly. Structural biology studies indicate that key features for Mfd to process a halted RNAP are its translocase activity and RNAP interactions (Smith and Savery, 2005; Brugger et al., 2020; Shi et al., 2020). While the conditions that lead Mfd to displace a halted RNAP (DNA-protein repression blocks and bulky-distortive DNA lesions) have been characterized, we know little about the type of events that is resolved by Mfd through realignment and reactivation of transcription complexes, particularly in stressed cells and at the systems' level (Selby and Sancar, 1995; Saxowsky and Doetsch, 2006; Smith and Savery, 2008; Belitsky and Sonenshein, 2011).

In summary, this work demonstrates that Mfd has profound effects on the transcriptome and phenotypes of stationary-phase *B. subtilis*. The loss of Mfd results in deficiencies in transcription-coupled DNA repair, the cellular responses to nutrient deprivation, cell envelope stress, antibiotic exposure, and protein oxidation. Also, Mfd influences cell differentiation behaviors that include endospore formation (Ramirez-Guadiana et al., 2013; Koo et al., 2017) (personal communication/manuscript in review Pedraza-Reyes) and motility. Because Mfd is well-conserved in bacteria, it would not be surprising if the pleiotropic effects observed in *B. subtilis* extend to other bacterial species, including those that are pathogenic. Considering the mutagenesis functions of Mfd that confer fitness to stationary-phase cells and resistance to antibiotics (Pybus et al., 2010; Ragheb et al., 2019), our results suggest that this factor operates at the intersection between gene expression and mutagenesis that mediates adaptation to stress and bacterial evolution.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, ID PRJNA673980.

AUTHOR CONTRIBUTIONS

HM, AS, FS, MP-R, and ER planned the experiments. HM, TE, RH, JG, KL, DA-M, and ER performed the laboratory work. HM, AS, MP-R, and ER analyzed the data. HM and ER wrote the manuscript. AS, TE, RH, JG, KL, DA-M, FS, and MP-R edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the NIH (GM131410 and P20), the NSF (DBI069267), and the CONACYT (A-1S-27116) grants. Also, research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103451.

ACKNOWLEDGMENTS

This research was performed as a collaboration with the New Mexico IDeA Networks for Biomedical Research Excellence (INBRE) supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.625705/full#supplementary-material

REFERENCES

Antelmann, H., and Helmann, J. D. (2011). Thiol-based redox switches and gene regulation. *Antioxid Redox Signal* 14, 1049–1063. doi: 10.1089/ars.2010.3400

Belitsky, B. R., and Sonenshein, A. L. (2011). Roadblock repression of transcription by *Bacillus subtilis* CodY. *J. Mol. Biol.* 411, 729–743.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093.

Boor, K. J., Duncan, M. L., and Price, C. W. (1995). Genetic and transcriptional organization of the region encoding the beta subunit of *Bacillus subtilis* RNA polymerase. *J. Biol. Chem.* 270, 20329–20336. doi: 10.1074/jbc.270.35.20329

Brugger, C., Zhang, C., Suhanovsky, M. M., Kim, D. D., Sinclair, A. N., Lyumkis, D., et al. (2020). Molecular determinants for dsDNA translocation by the transcription-repair coupling and evolvability factor Mfd. *Nat. Comm.* 11, 1–12.

Bsat, N., Herbig, A., Casillas-Martinez, L., Setlow, P., and Helmann, J. D. (1998). *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. *Mol. Microbiol.* 29, 189–198. doi: 10.1046/j.1365-2958.1998.00921.x

Castro-Cerritos, K. V., Lopez-Torres, A., Obregón-Herrera, A., Wrobel, K., Wrobel, K., and Pedraza-Reyes, M. (2018). LC–MS/MS proteomic analysis of starved *Bacillus subtilis* cells overexpressing ribonucleotide reductase (*nrdEF*): implications in stress-associated mutagenesis. *Curr. Genet.* 64, 215–222. doi: 10.1007/s00294-017-0722-7

Castro-Cerritos, K. V., Yasbin, R. E., Robleto, E. A., and Pedraza-Reyes, M. (2017). Role of ribonucleotide reductase in *Bacillus subtilis* stress-associated mutagenesis. *J. Bacteriol.* 199, 15–16.

Chi, B. K., Albrecht, D., Gronau, K., Becher, D., Hecker, M., and Antelmann, H. (2010). The redox-sensing regulator YodB senses quinones and diamide via a thiol-disulfide switch in *Bacillus subtilis*. *Proteomics* 10, 3155–3164.

Cosmina, P., Rodriguez, F., De Ferra, F., Grandi, G., Perego, M., Venema, G., et al. (1993). Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mole. Microbiol.* 8, 821–831.

Deaconescu, A. M., Chambers, A. L., Smith, A. J., Nickels, B. E., Hochschild, A., Savery, N. J., et al. (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell* 124, 507–520. doi: 10.1016/j.cell.2005.11.045

Dubnau, D. (1991). Genetic competence in *Bacillus subtilis*. *Microbiol. Mole. Biol. Rev.* 55, 395–424. doi: 10.1128/mmbr.55.3.395-424.1991

Duigou, S., Ehrlich, S. D., Noirot, P., and Noirot-Gros, M. F. (2004). Distinctive genetic features exhibited by the Y-family DNA polymerases in *Bacillus subtilis*. *Mol. Microbiol.* 54, 439–451. doi: 10.1111/j.1365-2958.2004.04259.x

Epshtein, V., Kamarthapu, V., Mcgary, K., Svetlov, V., Ueberheide, B., Proshkin, S., et al. (2014). UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* 505, 372–377. doi: 10.1038/nature12928

Fuangthong, M., Atichartpongkul, S., Mongkolsuk, S., and Helmann, J. D. (2001). OhrR is a repressor of *ohrA*, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. *J. Bacteriol.* 183, 4134–4141. doi: 10.1128/jb.183.14.4134-4141.2001

Fuma, S., Fujishima, Y., Corbell, N., D'souza, C., Nakano, M. M., Zuber, P., et al. (1993). Nucleotide sequence of 5'portion of srfA that contains the region required for competence establishment in *Bacillus subtilus*. *Nucleic Acids Res.* 21, 93–97. doi: 10.1093/nar/21.1.93

Gaballa, A., Newton, G. L., Antelmann, H., Parsonage, D., Upton, H., Rawat, M., et al. (2010). Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in *Bacilli*. *Proc. Natl. Acad. Sci.* 107, 6482–6486. doi: 10.1073/pnas.1000928107

Gerhardt, P., Murray, R. G. E., Wood, W. A., and Krieg, N. R. (1994). *Methods for General and Molecular Bacteriology*. Washington, D.C: American Society for Microbiology.

Gomez-Marroquin, M., Martin, H. A., Pepper, A., Girard, M. E., Kidman, A. A., Vallin, C., et al. (2016). Stationary-Phase Mutagenesis in Stressed *Bacillus subtilis* Cells Operates by Mfd-Dependent Mutagenic Pathways. *Genes* 7:33. doi: 10.3390/genes7070033

Haines, N. M., Kim, Y. I., Smith, A. J., and Savery, N. J. (2014). Stalled transcription complexes promote DNA repair at a distance. *Proc. Natl. Acad. Sci. U S A* 111, 4037–4042.

Han, J., Sahin, O., Barton, Y. W., and Zhang, Q. (2008). Key role of Mfd in the development of fluoroquinolone resistance in *Campylobacter jejuni*. *PLoS Pathog* 4:e1000083. doi: 10.1371/journal.ppat.1000083

Hanawalt, P. C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell. Biol.* 9, 958–970. doi: 10.1038/nrm2549

Ho, H. N., Van Oijen, A. M., and Ghodke, H. (2018). The transcription-repair coupling factor Mfd associates with RNA polymerase in the absence of exogenous damage. *Nat. Commun.* 9:1570.

Hochgräfe, F., Mostertz, J., Pöther, D.-C., Becher, D., Helmann, J. D., and Hecker, M. (2007). S-cysteinylation is a general mechanism for thiol protection of *Bacillus subtilis* proteins after oxidative stress. *J. Biol. Chem.* 282, 25981–25985. doi: 10.1074/jbc.c700105200

Inaoka, T., Matsumura, Y., and Tsuchido, T. (1998). Molecular cloning and nucleotide sequence of the superoxide dismutase gene and characterization of its product from *Bacillus subtilis*. *J. Bacteriol.* 180, 3697–3703. doi: 10.1128/jb.180.14.3697-3703.1998

Kearns, D. B., Chu, F., Rudner, R., and Losick, R. (2004). Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. *Mole. Microbiol.* 52, 357–369. doi: 10.1111/j.1365-2958.2004.03996.x

Koo, B. M., Kritikos, G., Farelli, J. D., Todor, H., Tong, K., Kimsey, H., et al. (2017). Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. *Cell Syst.* 4:e297.

Le, T. T., Yang, Y., Tan, C., Suhanovsky, M. M., Fulbright, R. M. Jr., Inman, J. T., et al. (2018). Mfd Dynamically Regulates Transcription via a Release and Catch-Up Mechanism. *Cell* 173:1823. doi: 10.1016/j.cell.2018.06.002

Lee, G. H., Jeong, J. Y., Chung, J. W., Nam, W. H., Lee, S. M., Pak, J. H., et al. (2009). The *Helicobacter pylori* Mfd protein is important for antibiotic resistance and DNA repair. *Diagn. Microbiol. Infect. Dis.* 65, 454–456.

Leichert, L. I., Scharf, C., and Hecker, M. (2003). Global characterization of disulfide stress in *Bacillus subtilis*. *J. Bacteriol.* 185, 1967–1975. doi: 10.1128/jb.185.6.1967-1975.2003

Leyva-Sánchez, H. C., Villegas-Negrete, N., Abundiz-Yanez, K., Yasbin, R. E., Robleto, E. A., and Pedraza-Reyes, M. (2020). Role of Mfd and GreA in *Bacillus subtilis* Base Excision Repair-Dependent Stationary-Phase Mutagenesis. *J. Bacteriol.* 202. doi: 10.1128/JB.00807-19

Martin, H. A., Pedraza-Reyes, M., Yasbin, R. E., and Robleto, E. A. (2011). Transcriptional de-repression and Mfd are mutagenic in stressed *Bacillus subtilis* cells. *J. Mol. Microbiol. Biotechnol.* 21, 45–58. doi: 10.1159/000332751

Martin, H. A., Porter, K. E., Vallin, C., Ermi, T., Contreras, N., Pedraza-Reyes, M., et al. (2019). Mfd protects against oxidative stress in *Bacillus subtilis* independently of its canonical function in DNA repair. *BMC Microbiol.* 19:26. doi: 10.1186/s12866-019-1394-x

Matsunaga, I., Ueda, A., Fujiwara, N., Sumimoto, T., and Ichihara, K. (1999). Characterization of the *ybdT* gene product of *Bacillus subtilis*: novel fatty acid beta-hydroxylating cytochrome P450. *Lipids* 34, 841–846. doi: 10.1007/s11745-999-0431-3

Merrikh, H., and Kohli, R. M. (2020). Targeting evolution to inhibit antibiotic resistance. *FEBS J.* 287, 4341–4353. doi: 10.1111/febs.15370

Moreno-Del Alamo, M., Torres, R., Manfredi, C., Ruiz-Maso, J. A., Del Solar, G., and Alonso, J. C. (2020). *Bacillus subtilis* PcrA Couples DNA Replication. *Trans. Recomb. Segreg. Front. Mol. Biosci.* 7:140. doi: 10.3389/fmolb.2020.00140

Mukherjee, S., and Kearns, D. B. (2014). The structure and regulation of flagella in *Bacillus subtilis*. *Ann. Rev. Genet.* 48, 319–340.

Nakano, M. M., Xia, L., and Zuber, P. (1991). Transcription initiation region of the srfA operon, which is controlled by the comP-comA signal transduction system in *Bacillus subtilis*. *J. Bacteriol.* 173, 5487–5493. doi: 10.1128/jb.173.17.5487-5493.1991

Pandey, S., Ogloblina, A. M., Belotserkovskii, B. P., Dolinnaya, N. G., Yakubovskaya, M. G., Mirkin, S. M., et al. (2015). Transcription blockage by stable H-DNA analogs in vitro. *Nucleic Acids Res.* 43, 6994–7004. doi: 10.1093/nar/gkv622

Park, J. S., Marr, M. T., and Roberts, J. W. (2002). *E. coli* Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* 109, 757–767. doi: 10.1016/s0092-8674(02)00769-9

Patrick, J. E., and Kearns, D. B. (2009). Laboratory strains of *Bacillus subtilis* do not exhibit swarming motility. *J. Bacteriol.* 191, 7129–7133. doi: 10.1128/jb.00905-09

Perera, V. R., Lapek, J. D. Jr., Newton, G. L., Gonzalez, D. J., and Pogliano, K. (2018). Identification of the S-transferase like superfamily bacillithiol transferases encoded by *Bacillus subtilis*. *PLoS One* 13:e0192977. doi: 10.1371/journal.pone.0192977

Petersohn, A., Antelmann, H., Gerth, U., and Hecker, M. (1999). Identification and transcriptional analysis of new members of the σB regulon in *Bacillus subtilis*. *Microbiology* 145, 869–880. doi: 10.1099/13500872-145-4-869

Pother, D. C., Liebeke, M., Hochgrafe, F., Antelmann, H., Becher, D., Lalk, M., et al. (2009). Diamide triggers mainly S Thiolations in the cytoplasmic proteomes of *Bacillus subtilis* and *Staphylococcus aureus*. *J. Bacteriol.* 191, 7520–7530. doi: 10.1128/jb.00937-09

Pybus, C., Pedraza-Reyes, M., Ross, C. A., Martin, H., Ona, K., Yasbin, R. E., et al. (2010). Transcription-associated mutation in *Bacillus subtilis* cells under stress. *J. Bacteriol.* 192, 3321–3328. doi: 10.1128/jb.00354-10

Radeck, J., Fritz, G., and Mascher, T. (2017). The cell envelope stress response of *Bacillus subtilis:* from static signaling devices to dynamic regulatory network. *Curr. Genet.* 63, 79–90. doi: 10.1007/s00294-016-0624-0

Ragheb, M. N., Merrikh, C., Browning, K., and Merrikh, H. (2021). Mfd regulates RNA polymerase association with hard-to-transcribe regions in vivo, especially those with structured RNAs. *Proc. Natl. Acad. Sci.* 118:e2008498118. doi: 10.1073/pnas.2008498118

Ragheb, M. N., Thomason, M. K., Hsu, C., Nugent, P., Gage, J., Samadpour, A. N., et al. (2019). Inhibiting the evolution of antibiotic resistance. *Mole. Cell* 73, 157.–165.

Ramirez-Guadiana, F. H., Del Carmen, Barajas-Ornelas, R., Ayala-Garcia, V. M., Yasbin, R. E., Robleto, E., et al. (2013). Transcriptional coupling of DNA repair in sporulating *Bacillus subtilis* cells. *Mol. Microbiol.* 90, 1088–1099.

Ross, C., Pybus, C., Pedraza-Reyes, M., Sung, H. M., Yasbin, R. E., and Robleto, E. (2006). Novel role of mfd: effects on stationary-phase mutagenesis in *Bacillus subtilis*. *J. Bacteriol.* 188, 7512–7520. doi: 10.1128/JB.00980-06

Savery, N. J. (2007). The molecular mechanism of transcription-coupled DNA repair. *Trends Microbiol.* 15, 326–333. doi: 10.1016/j.tim.2007.05.005

Saxowsky, T. T., and Doetsch, P. W. (2006). RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? *Chem. Rev.* 106, 474–488. doi: 10.1021/cr040466q

Selby, C. P., and Sancar, A. (1995). Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J. Biol. Chem.* 270, 4890–4895. doi: 10.1074/jbc.270.9.4890

Shi, J., Wen, A., Zhao, M., Jin, S., You, L., Shi, Y., et al. (2020). Structural basis of Mfd-dependent transcription termination. *Nucleic Acids Res.* 48, 11762–11772. doi: 10.1093/nar/gkaa904

Smith, A. J., and Savery, N. J. (2005). RNA polymerase mutants defective in the initiation of transcription-coupled DNA repair. *Nucleic Acids Res.* 33, 755–764. doi: 10.1093/nar/gki225

Smith, A. J., and Savery, N. J. (2008). Effects of the bacterial transcription-repair coupling factor during transcription of DNA containing non-bulky lesions. *DNA Repair* 7, 1670–1679. doi: 10.1016/j.dnarep.2008.06.020

Spizizen, J. (1958). Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxiribonucleate. *Proc. Natl. Acad. Sci.* 44, 1072–1078. doi: 10.1073/pnas.44.10.1072

Sprouffske, K., and Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinform.* 17:172. doi: 10.1186/s12859-016-1016-7

Sung, H.-M., and Yasbin, R. E. (2002). Adaptive, or Stationary-Phase. Mutagenesis, a Component of Bacterial Differentiation in Bacillus subtilis. *J. Bacteriol.* 184, 5641–5653. doi: 10.1128/jb.184.20.5641-5653.2002

Tornaletti, S., Park-Snyder, S., and Hanawalt, P. C. (2008). G4-forming sequences in the non-transcribed DNA strand pose blocks to T7 RNA polymerase and mammalian RNA polymerase II. *J. Chem.* 283, 12756–12762. doi: 10.1074/jbc.m705003200

Tu Quoc, P. H., Genevaux, P., Pajunen, M., Savilahti, H., Georgopoulos, C., Schrenzel, J., et al. (2007). Isolation and characterization of biofilm formation-defective mutants of *Staphylococcus aureus*. *Infect. Immun.* 75, 1079–1088. doi: 10.1128/iai.01143-06

Van Der Ploeg, J. R., Barone, M., and Leisinger, T. (2001). Functional analysis of the *Bacillus subtilis cysK* and *cysJI* genes. *FEMS Microbiol. Lett.* 201, 29–35. doi: 10.1016/s0378-1097(01)00225-7

Van Der Ploeg, J. R., Cummings, N. J., Leisinger, T., and Connerton, I. F. (1998). *Bacillus subtilis* genes for the utilization of sulfur from aliphatic sulfonates. *Microbiology* 144(Pt 9), 2555–2561. doi: 10.1099/00221287-144-9-2555

Völker, U., Andersen, K. K., Antelmann, H., Devine, K. M., and Hecker, M. (1998). One of two OsmC homologs in *Bacillus subtilis* is part of the ςB-dependent general stress regulon. *J. Bacteriol.* 180, 4212–4218.

Yakhnin, A. V., Fitzgerald, P. C., Mcintosh, C., Yakhnin, H., Kireeva, M., Turek-Herman, J., et al. (2020). NusG controls transcription pausing and RNA polymerase translocation throughout the *Bacillus subtilis* genome. *Proc. Natl. Acad. Sci. U S A* 117, 21628–21636.

Yasbin, R. E., Wilson, G. A., and Young, F. E. (1975). Transformation and transfection in lysogenic strains of *Bacillus subtilis*: evidence for selective induction of prophage in competent cells. *J. Bacteriol.* 121, 296–304. doi: 10.1128/jb.121.1.296-304.1975

Zalieckas, J. M., Wray, L. V. Jr., and Fisher, S. H. (1998b). Expression of the *Bacillus subtilis acsA* gene: position and sequence context affect cre-mediated carbon catabolite repression. *J. Bacteriol.* 180, 6649–6654. doi: 10.1128/jb.180.24.6649-6654.1998

Zalieckas, J. M., Wray, L. V. Jr., Ferson, A. E., and Fisher, S. H. (1998a). Transcription-repair coupling factor is involved in carbon catabolite repression of the *Bacillus subtilis hut* and *gnt* operons. *Mol. Microbiol.* 27, 1031–1038. doi: 10.1046/j.1365-2958.1998.00751.x

Zhang, X., Hu, Y., Guo, X., Lescop, E., Li, Y., Xia, B., et al. (2006). The *Bacillus subtilis* YkuV is a thiol:disulfide oxidoreductase revealed by its redox structures and activity. *J. Biol. Chem.* 281, 8296–8304.

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

PERSPECTIVE
published: 29 April 2021
doi: 10.3389/fmicb.2021.662907

Connection Between Chromosomal Location and Function of CtrA Phosphorelay Genes in Alphaproteobacteria

Jürgen Tomasch[1]†, Sonja Koppenhöfer[2†] and Andrew S. Lang[2]**

[1] Department of Molecular Bacteriology, Helmholtz-Center for Infection Research, Braunschweig, Germany, [2] Department of Biology, Memorial University of Newfoundland, St. John's, NL, Canada

OPEN ACCESS

Edited by:
Morigen Morigen,
Inner Mongolia University, China

Reviewed by:
Gregory Marczynski,
McGill University, Canada
Emanuele G. Biondi,
UMR 7283 Laboratoire de Chimie Bactérienne (LCB), France
Justine Collier,
University of Lausanne, Switzerland

***Correspondence:**
Jürgen Tomasch
juergen.tomasch@helmholtz-hzi.de
Andrew S. Lang
aslang@mun.ca

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: 01 February 2021
Accepted: 09 April 2021
Published: 29 April 2021

Citation:
Tomasch J, Koppenhöfer S and Lang AS (2021) Connection Between Chromosomal Location and Function of CtrA Phosphorelay Genes in Alphaproteobacteria. Front. Microbiol. 12:662907. doi: 10.3389/fmicb.2021.662907

Most bacterial chromosomes are circular, with replication starting at one origin (*ori*) and proceeding on both replichores toward the terminus (*ter*). Several studies have shown that the location of genes relative to *ori* and *ter* can have profound effects on regulatory networks and physiological processes. The CtrA phosphorelay is a gene regulatory system conserved in most alphaproteobacteria. It was first discovered in *Caulobacter crescentus* where it controls replication and division into a stalked and a motile cell in coordination with other factors. The locations of the *ctrA* gene and targets of this response regulator on the chromosome affect their expression through replication-induced DNA hemi-methylation and specific positioning along a CtrA activity gradient in the dividing cell, respectively. Here we asked to what extent the location of CtrA regulatory network genes might be conserved in the alphaproteobacteria. We determined the locations of the CtrA phosphorelay and associated genes in closed genomes with unambiguously identifiable *ori* from members of five alphaproteobacterial orders. The location of the phosphorelay genes was the least conserved in the Rhodospirillales followed by the Sphingomonadales. In the Rhizobiales a trend toward certain chromosomal positions could be observed. Compared to the other orders, the CtrA phosphorelay genes were conserved closer to *ori* in the Caulobacterales. In contrast, the genes were highly conserved closer to *ter* in the Rhodobacterales. Our data suggest selection pressure results in differential positioning of CtrA phosphorelay and associated genes in alphaproteobacteria, particularly in the orders Rhodobacterales, Caulobacterales and Rhizobiales that is worth deeper investigation.

Keywords: CtrA phosphorelay, replication, genome evolution, genome organization, gene expression

INTRODUCTION

Most bacteria possess one circular chromosome. Replication is initiated through unwinding the two DNA strands at the origin of replication (*ori*) and proceeds on both replichores toward the terminus (*ter*). Here, the dimer of newly synthesized chromosomes is resolved, and cell division can be completed (reviewed by Reyes-Lamothe et al., 2012). Close links between replication and

organization of genes on the chromosome became evident with the first complete bacterial genomes (Rocha, 2004; Touchon and Rocha, 2016). In many bacteria, genes are preferentially oriented co-directional to replication progression. This pattern probably evolved to avoid collisions between DNA and RNA polymerase complexes (Lang and Merrikh, 2018). In recent years it also became apparent that the specific locations of genes can have a major influence on transcription levels and thereby control physiological processes (Slager and Veening, 2016). For instance, chromosomal location results in differences in the copy-number of genes during replication. Therefore, genes that are more highly expressed, such as those encoding transcription and translation proteins, tend to be conserved near *ori* (Couturier and Rocha, 2006). The importance of gene location has also been validated experimentally: Relocating *Vibrio cholerae* genes encoding ribosomal proteins to the *ter* region resulted in severe growth defects (Soler-Bistué et al., 2015).

Positioning of genes on the chromosome might also be dictated by regulatory needs. In *Escherichia coli* and other gammaproteobacteria, genes coding for nucleoid-associated proteins and regulators are ordered according to their activities during the growth cycle (Sobetzko et al., 2012). For example, *rpoN*, expressed during exponential growth, is located closer to *ori* while *rpoS*, expressed in the stationary phase, is located closer to *ter*. The same trend was found for the targets of these sigma factors. One of the most fascinating examples is how replication-oriented location of regulatory genes is employed to control the timing of *Bacillus subtilis* spore formation (Narula et al., 2015). Here, imbalance between the expression of *ori* and *ter* located members of a phosphorylation chain during replication inhibits activation of the sporulation-inducing transcription factor Spo0A. This ensures that spore formation is only induced in cells with two complete chromosomes.

Proper transcription of the important cell cycle regulatory gene *ctrA* of *Caulobacter crescentus* is dependent on its chromosomal location, too (Reisenauer and Shapiro, 2002). In alphaproteobacteria, the CtrA phosphorelay regulatory system is widely conserved (Brilli et al., 2010; Panis et al., 2015). We recently found that its key regulatory components are concentrated proximal to *ori* and *ter* in the Rhodobacterales *Dinoroseobacter shibae* and *Rhodobacter capsulatus* and, in contrast to *C. crescentus*, the *ctrA* gene itself is located close to *ter* in both organisms (Koppenhöfer et al., 2019).

In this perspective article, we will first provide a brief overview of the CtrA phosphorelay and its role in controlling the cell cycle and other traits in different bacteria. We will focus on how changes in DNA methylation during replication and the formation of a phosphorylation gradient in predivisional cells influence the regulatory system. Then, we will show that chromosomal location of the regulatory genes is conserved to varying degrees within alphaproteobacterial orders and differs among them. We propose that one consequence of the differing gene locations might be altered timing of expression during the cell cycle. Understanding how their positioning shapes the functionality of the CtrA phosphorelay and associated genes might help to explain the evolution of distinct roles in different alphaproteobacterial orders.

CtrA AND CELL CYCLE CONTROL IN ALPHAPROTEOBACTERIA

Replicating only once per cell division, the dimorphic bacterium *C. crescentus* displays a eukaryotic-like cell cycle (**Figure 1A**). The growth-arrested flagellated cell (G1 phase) can transform into a stalked cell that replicates (S phase) and prepares for division (G2 phase) into two physiologically different daughter cells. The old, stalked cell can directly undergo the next round of replication while the new, flagellated cell remains in a growth-arrested state (Degnen and Newton, 1972). Key to the replication-coupled differentiation is the orchestration of gene expression by an array of interconnected regulatory circuits (reviewed by Frandi and Collier, 2019) among which the CtrA phosphorelay takes a leading role (Laub et al., 2000, 2002).

As part of the regulation of cell division events, autophosphorylation of the transmembrane histidine kinase CckA results in phosphorylation of the phosphotransferase ChpT, which in turn phosphorylates the response regulator CtrA (Biondi et al., 2006). Phosphorylated CtrA then activates differentiation-specific genes and inhibits replication initiation by blocking *ori* from binding by the replication initiator DnaA (Quon et al., 1998; Siam and Marczynski, 2000). Replication rounds are controlled via the directed proteolysis of several transcriptional regulators, including fast degradation of CtrA in the stalked cell, mediated by ClpX, in order to enable initiation of replication (reviewed by Jenal, 2009). The chromosomal position influences transcription of *ctrA* and CtrA targets through DNA methylation and establishment of a CtrA phosphorylation gradient, respectively (detailed in the following sections and **Figure 1**).

The *C. crescentus* chromosome is linearly ordered in the cell with the *ori* located at the stalked or flagellated pole and *ter* oriented at the pole where the division plane will form (Yildirim and Feig, 2018). The CcrM methyltransferase specifically methylates the adenosine in GANTC palindromic motifs. Expression of *ccrM* is restricted to the transition from S to G2 phase (Zweiger et al., 1994; Laub et al., 2000). Thus, newly replicated DNA stays hemi-methylated until replication has finished (**Figure 1A**). Expression of *ori*-proximal *ctrA* is controlled by two different promoters (**Figure 1B**). Promoter P1 gets activated in the early S phase and CtrA then triggers its own expression from promoter P2 while inhibiting expression from P1 in pre-divisional cells and the flagellated daughter cell (Domian et al., 1999). Activation of P1 requires hemi-methylation of an upstream GANTC site, and thus the replication fork has to pass the *ctrA* locus for this promoter to be active (Reisenauer and Shapiro, 2002; Mohapatra et al., 2020). Activity of P1 is highest when the respective motif on the coding strand is methylated (Mohapatra et al., 2020). If *ctrA* is moved closer to *ter*, the P1-associated GANTC motif remains in the fully methylated state longer and the resulting delay of *ctrA* transcription leads to elongated flagellated daughter cells. The transcription factor responsible for P1 regulation is GcrA, which is active exclusively in S phase cells and oscillates with CtrA activity (Holtzendorff et al., 2004). GcrA preferentially binds and activates promoters

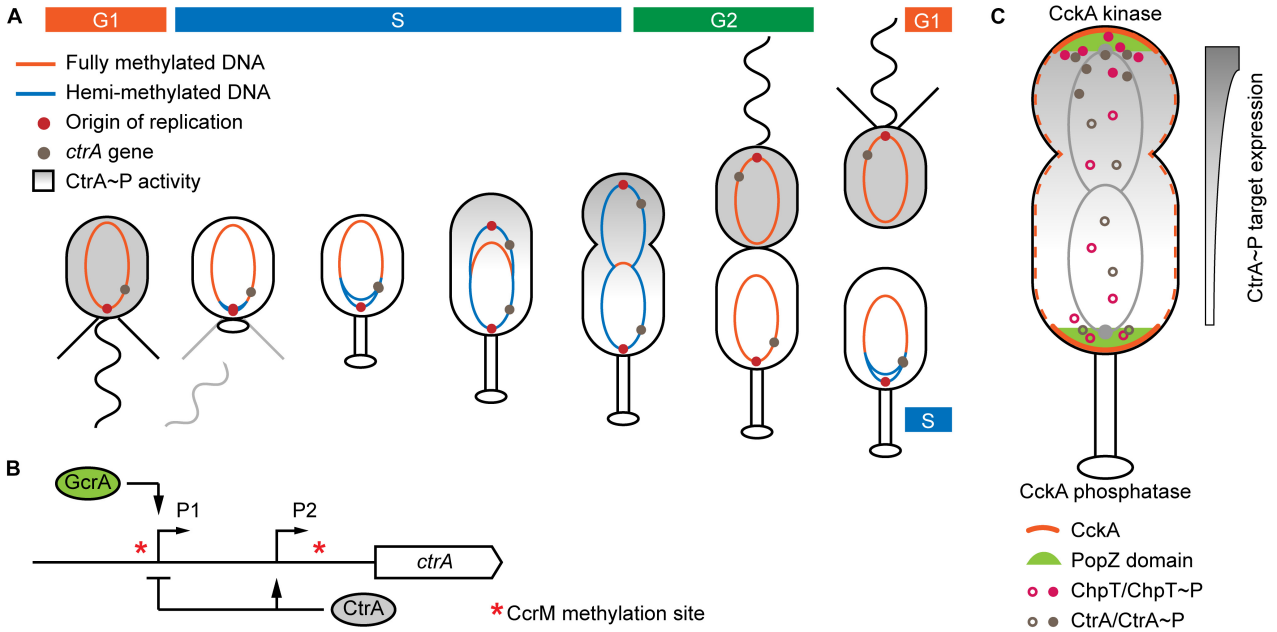


FIGURE 1 | Mechanisms of *C. crescentus* differentiation for which chromosomal localization matters. **(A)** Changes of chromosome methylation state and CtrA activity during the *C. crescentus* cell cycle. Newly replicated DNA stays hemi-methylated during the S phase allowing *ctrA* transcription to be activated. CtrA activity is restricted to the late S phase and the flagellated daughter cell. **(B)** Control of *ctrA* expression. Hemi-methylated P1 is activated by GcrA. Phosphorylated CtrA inhibits expression from P1 and activates expression from P2. **(C)** Establishment of a CtrA activity gradient through localized phosphorylation/dephosphorylation. The protein environment at the new pole triggers kinase functionality of CckA. The PopZ microdomain ensures proximity of phosphorelay components. Panel A inspired by Panis et al. (2015) and panel C inspired by Lasker et al. (2020).

carrying fully or hemi-methylated GANTC motifs (Fioravanti et al., 2013; Haakonsen et al., 2015). Replication-controlled methylation also affects *ftsZ* expression, which encodes the divisome Z-ring protein. In this case the promoter is most active in the fully methylated state (Gonzalez and Collier, 2013). The regulatory function of CcrM is probably conserved broadly in alphaproteobacteria as GANTC motifs are enriched in intergenic regions on the vast majority of chromosomes (Gonzalez et al., 2014).

CckA is dispersed throughout the inner membrane, but concentrates at the cell poles in pre-divisional cells (Angelastro et al., 2010). It acts as a kinase at the new cell pole and as a phosphatase elsewhere. The switch in enzymatic activity is controlled by interaction with different sets of proteins (Tsokos et al., 2011). Essential for triggering the kinase activity of CckA are its homo-oligomerization and direct interaction with the pseudo-kinase DivL, both concentrated at the cell poles (Mann and Shapiro, 2018). Recently, Lasker et al. (2020) demonstrated the formation of diffusion-limiting microdomains at the cell poles that ensure close proximity of CckA, ChpT and CtrA in order to allow efficient phosphotransfer (**Figure 1C**). The polar localization of phosphorylating and dephosphorylating enzymatic chains ensures the formation of a CtrA activity gradient from the flagellated to the stalked pole in pre-divisional cells. When a promoter that is regulated exclusively by CtrA was repositioned on the chromosome, its expression decreased along the *ori-ter* axis, in accordance with the increasing distance from the flagellated cell pole (Lasker et al., 2020).

The core components of the CtrA phosphorelay are highly conserved within the alphaproteobacteria and connected to accessory regulatory systems that are often restricted to specific orders (Brilli et al., 2010). In particular, most genes of the polarity module (van Teeseling and Thanbichler, 2020) essential for dimorphic development of *C. crescentus* are found only in members of the Caulobacterales and Rhizobiales orders, an exception being the more widely conserved *divL* gene. The CtrA regulon also differs among orders (Brilli et al., 2010; Panis et al., 2015). Flagellar genes are controlled by CtrA in all studied orders, including the early branching Rhodospirillales, leading to the hypothesis that regulation of motility was the primordial role of CtrA and cell cycle control was acquired later (Greene et al., 2012). Transcriptional activation of the DNA repair machinery, observed in several species, might also be a more ancient function of CtrA (Poncin et al., 2018).

Construction of *ctrA* knockout mutants failed or they showed severe growth defects in the Caulobacterales *C. crescentus* (Quon et al., 1996; Christen et al., 2011; Guzzo et al., 2020) and *Hyphomonas neptunium* (Leicht et al., 2020) as well as the Rhizobiales *Sinorhizobium meliloti* (Barnett et al., 2001) and *Brucella abortus* (Bellefontaine et al., 2002), but has no negative effects on growth or viability in the Sphingomonadales *Sphingomonas melonis* (Francez-Charlot et al., 2015), Rhodospirillales *Rhodospirillum centenum*

(Bird and MacKrell, 2011) and *Magnetospirillum magneticum* (Greene et al., 2012), as well as the Rhodobacterales members studied so far (Miller and Belas, 2006; Mercer et al., 2010; Zan et al., 2013; Wang et al., 2014; Hernández-Valle et al., 2020). While the influence of CtrA on replication has been demonstrated in *Sphingomonas melonis* and some Rhodobacterales, these bacteria lack a strict dimorphic lifestyle or polar growth that has been demonstrated for the Caulobacterales and Rhizobiales. These findings led to the hypothesis that essentiality of *ctrA* arose during the evolution of a lifestyle that couples differentiation to reproduction (van Teeseling and Thanbichler, 2020). An intermediate step might have been a non-essential influence on replication and cell division.

CONSERVED LOCATION OF CtrA-ASSOCIATED GENES IN ALPHAPROTEOBACTERIAL ORDERS

Given that chromosomal localization influences expression of cell cycle-controlling genes in the model alphaproteobacterium *C. crescentus* and knowing that genes key to this process are conserved within this class, we asked if the location of these genes shows a pattern and if there are differences among orders in which the CtrA phosphorelay is an essential regulator of replication-coupled differentiation and those in which it is not. We used a dataset of 179 closed genomes from five alphaproteobacterial orders for which Ori-Finder (Gao and Zhang, 2008) unambiguously identified *ori* and identified the orthologs of CtrA phosphorelay and associated genes using Proteinortho (Lechner et al., 2011). Only one representative strain was selected for each species to avoid species overrepresentation bias. Detailed analysis steps can be found in the **Supplementary Material**. The analyzed genomes from the Rhodospirillales, Sphingomonadales, Rhodobacterales, Caulobacterales and Rhizobiales are listed in **Supplementary Table 1**, and the analyzed genes are listed in **Supplementary Tables 2, 3**.

Figure 2A summarizes the localization analysis with the upper panel showing one representative genome and the lower panel showing the frequency of the respective genes within 20 segments on the *ori-ter* axis. The orders are arranged phylogenetically with the earliest branching Rhodospirillales at the left and the latest branching Rhizobiales on the right (Muñoz-Gómez et al., 2019). In almost all analyzed genomes the *parAB* genes were located close to *ori* with some exceptions in the Rhodospirillales, Rhodobacterales, and Rhizobiales. This is in accordance with previous studies that found the *par* locus predominantly conserved close to *ori* (Livny et al., 2007). The location of the other analyzed genes and the respective conservation differed among the orders.

We observed no conserved localization of the genes examined in the Rhodospirillales with the exceptions of *parAB* and *clpX*, which tended to be located closer to *ter* (**Figure 2A**). Surprisingly, and in contrast to the other orders, *ctrA*, *gcrA* and *ccrM* were identified in 72–82% of the genomes, whereas *chpT*, *cckA* and *divL* were identified in only 27–36% (**Supplementary Table 1**). This might indicate that the selection pressure to maintain these genes is lower in Rhodospirillales than in the other orders. However, CckA and DivL are modular proteins, therefore their architecture might have evolved differently in this order and ChpT is a small protein that also shows greater divergence within orders (Brilli et al., 2010; Mercer et al., 2012), making definitive identification of homologs more difficult. Similarly, no clear distribution patterns of the genes analyzed were observed in the Sphingomonadales genomes. The only exceptions were *cckA* and *chpT*, which tended to be localized near *ter* (**Figure 2A**). Of note, in the Rhodospirillales and Sphingomonadales, the *dnaA* gene was not conserved close to *ori*, in contrast to the other three orders.

In the Caulobacterales the phosphorelay genes *ctrA*, *cckA* and *chpT* as well as *divL* showed conserved localization in the half of the chromosome closer to *ori*. The *clpX* gene was highly conserved in proximity to *ter*. By contrast, *ccrM* and *gcrA* were predominantly found midway between *ori* and *ter*. In the Rhizobiales, localization of *cckA* was conserved in the "lower half" of the chromosome with a peak close to *ter* while *chpT* and *ctrA* were preferentially located midway between *ori* and *ter*. The genes *clpX*, *ccrM* and *gcrA* showed similar trends as in the Caulobacterales. Interestingly, in many of the Rhizobiales genomes we identified two *divL* homologs, one of which was conserved near *ori* (**Supplementary Table 2**).

In the Rhodobacterales a clear preference for *ter*-proximal localization was observed for *cckA*, *chpT* and *ctrA*, while *ccrM* was preferentially located close to *ori*. Like in the Rhizobiales, several genomes contained two paralogs of *divL* that were located mostly near *ori* (**Supplementary Table 2**). We also analyzed the chromosomal position of other genes that are part of the CtrA regulon in this order and that regulate the gene transfer agent (GTA) gene cluster (**Supplementary Table 3**). The direct activator of the GTA cluster *gafA* (Fogg, 2019) and its neighbor (Dshi_1585 in *D. shibae*) were located in proximity to *ter* while the *rbaVW* genes that encode part of a partner-switching phosphorelay system (Mercer and Lang, 2014) were preferentially found close to *ori* (**Supplementary Figure 1**). Interestingly, the CtrA-controlled genes (Koppenhöfer et al., 2019) that are part of the DNA uptake and recombination machinery (*lexA*, *recA*, and *comECFM*) also showed a conserved location pattern.

As (hemi)-methylation is an important factor in the regulation of *ctrA* expression in *C. crescentus* we determined the number of GANTC motifs 300 bp upstream of the *ctrA* homologs in all orders (**Figure 2B**). All putative *ctrA* promotors contained at least one and up to five or six CcrM methylation sites in the Sphingomonadales and Caulobacterales, respectively. In 65% and 50% of all putative Rhodospirillales and Rhizobiales *ctrA* promoters, respectively, we identified the GANTC motif. The lowest number was found in Rhodobacterales where only 45% of all promoter regions contained this motif.

DISCUSSION

Here, we evaluated whether or not key regulators associated with the CtrA phosphorelay have conserved chromosomal locations.

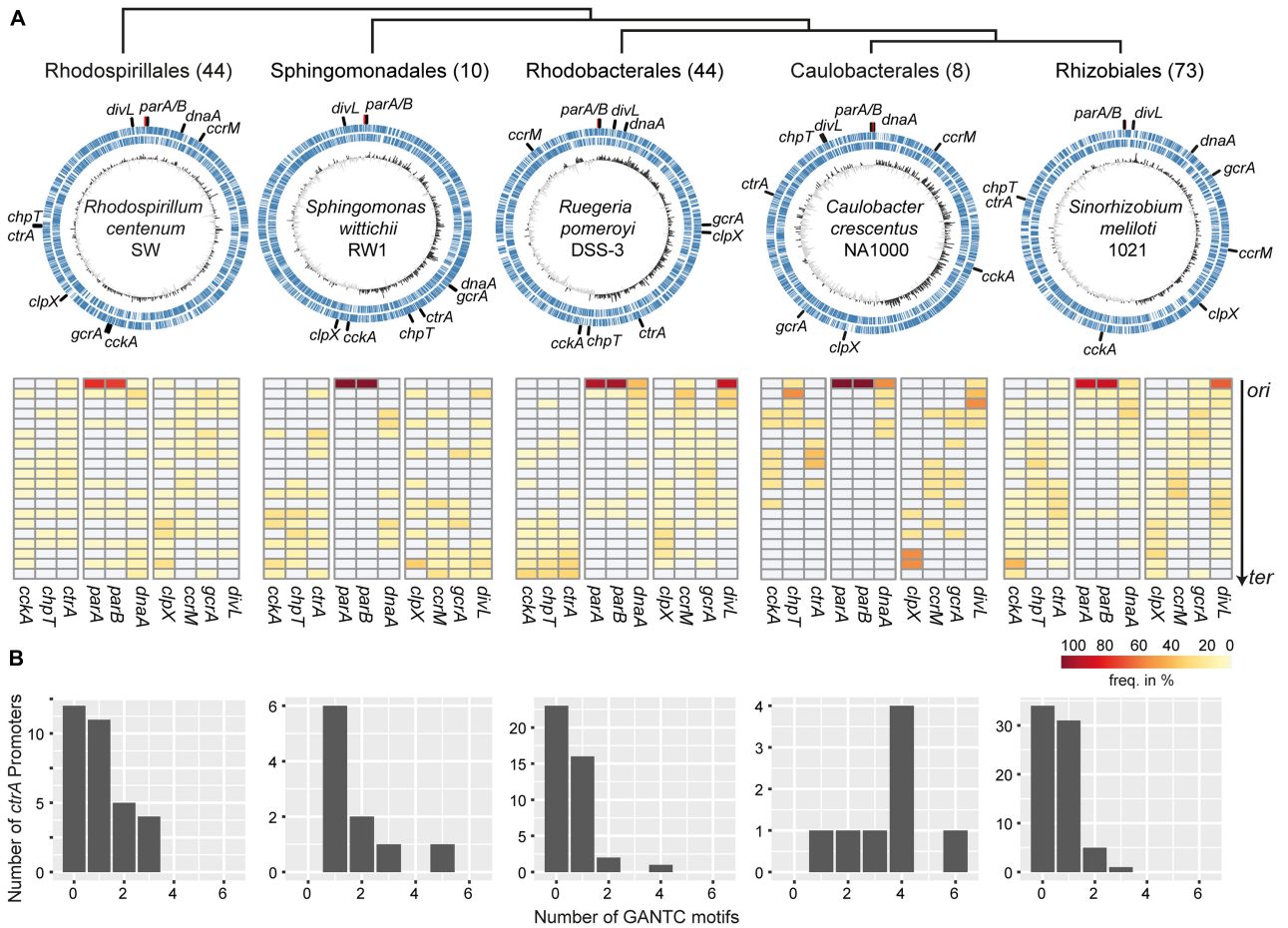


FIGURE 2 | Chromosomal localization of CtrA phosphorelay component genes and methylation of the *ctrA* promoter region in alphaproteobacteria. **(A)** Alphaproteobacterial orders are arranged according to Muñoz-Gómez et al. (2019). The numbers of genomes per order are in brackets. Upper panel: representative chromosomes for each order with positions of regulators marked. Lower panel: percentage of regulator genes within 20 segments of the chromosomes of the particular order oriented along the *ori-ter* axis. **(B)** Distribution of the number of CcrM methylation motifs in the *ctrA* promoter regions for each order.

The number of genomes available to analyze was small for the orders Caulobacterales and Sphingomonadales, leaving the possibility of a bias in our study. The employed Ori-Finder tool returned several possible *ori* positions for a considerable number of genomes that we excluded from further analysis. We found that the *parAB* genes might serve as a good anchor for manual curation of the *ori* position. The locations of *ori* and *ter* can also be identified experimentally by sequencing DNA from growing cultures when there will be a coverage gradient decreasing from start toward end of replication (Skovgaard et al., 2011; Jung et al., 2019). This could be considered for all future genome sequencing projects.

Despite the limitations, we could identify localization patterns in all orders except for the early branching Rhodospirillales in which the conservation of the CtrA phosphorelay was also lower than in the other orders. Particularly striking was the strong conservation of the phosphorelay genes near *ter* on the Rhodobacterales chromosomes. This conserved localization is also remarkable because core genes in this order show very distinct location patterns among different species (Kopejtka et al., 2019). Localization near *ter* and the low occurrence of GANTC motifs in the *ctrA* promoter might indicate that replication-mediated changes of the state of DNA methylation do not play a major role in regulation of gene expression in this order. On the other hand, establishment of a CtrA phosphorylation gradient might indeed also play a role in Rhodobacterales. The bifunctionality of CckA as a kinase and phosphatase has recently been demonstrated for *R. capsulatus* (Farrera-Calderon et al., 2020).

In some Rhodobacterales the CtrA phosphorelay is integrated into quorum sensing (QS) regulation (Leung et al., 2013; Zan et al., 2013; Wang et al., 2014). CtrA-mediated QS communication induces subpopulation-specific responses, most notably the "decision" of a small number of cells to produce GTAs (Ding et al., 2019; Koppenhöfer et al., 2019). A loss of the CtrA phosphorelay genes is not lethal, the bacteria just resemble a "silent" population. The location of the phosphorelay genes close to *ter* might indicate that communication-induced

differentiation is uncoupled from replication and cell division. It is also tempting to speculate that the location of genes controlling GTA expression at the opposite poles of the chromosome ensures repression of GTA production during replication, similar to spore formation in *B. subtilis* (Narula et al., 2015). Indeed, no DNA packaging bias along the *ori-ter* axis has been observed for GTAs, which would be expected if they are produced in replicating cells (Hynes et al., 2012; Tomasch et al., 2018). In the Rhizobiales and Caulobacterales, however, most of the essential CtrA phosphorelay genes are located toward the upper half of the chromosome. This might result in their activation during replication as observed for *ctrA* of *C. crescentus* (Reisenauer and Shapiro, 2002), leading to an interconnected essentiality of reproduction and physiological differentiation. Most essential *C. crescentus* genes are concentrated near *ori* or *ter* (Christen et al., 2011). It would be interesting to see if this pattern is conserved in other species with a pronounced dimorphic lifestyle.

In conclusion, our analysis suggests selection pressure to fix the position of CtrA phosphorelay and associated genes in different chromosomal regions depending on their involvement in different cell physiological processes. This is particularly evident in the Rhodobacterales, Caulobacterales and Rhizobiales. Understanding the underlying evolutionary forces will require both comparative genomic analysis and experimental data beyond what is currently available for a limited number of established model organisms. Our analysis concentrated on the core components of the CtrA phosphorelay but could be expanded to include more accessory regulators and CtrA targets in the different orders. It would also be interesting to identify highly related strains where recent chromosome rearrangements have led to different positions of genes of interest. In *Pseudomonas aeruginosa* a large-scale chromosome inversion resulted in large gene expression and physiological differences between two strains (Irvine et al., 2019). Similarly, analyzing the consequences of relocating genes, as has been done for *C. crescentus* and several other organisms, is a promising experimental approach for understanding the effects of chromosome positioning on gene regulation (Reisenauer and Shapiro, 2002; Gonzalez and Collier, 2013; Soler-Bistué et al., 2015; Lasker et al., 2020).

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 authors.

AUTHOR CONTRIBUTIONS

JT designed the study. SK analyzed the data. JT and SK visualized the data. JT, SK, and ASL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Research in ASL's group is funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada (RGPIN-2017-04636). SK was partially supported by funding from the Memorial University of Newfoundland School of Graduate Studies. Publication of this article has been financed by the open access fund of the Helmholtz-Center for Infection Research (HZI), Braunschweig, Germany.

ACKNOWLEDGMENTS

We thank the reviewers for their critical and constructive assessment of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.662907/full#supplementary-material

REFERENCES

Angelastro, P. S., Sliusarenko, O., and Jacobs-Wagner, C. (2010). Polar Localization of the CckA Histidine Kinase and Cell Cycle Periodicity of the Essential Master Regulator CtrA in Caulobacter crescentus. *J. Bacteriol.* 192, 539–552. doi: 10.1128/JB.00985-09

Barnett, M. J., Hung, D. Y., Reisenauer, A., Shapiro, L., and Long, S. R. (2001). A homolog of the CtrA cell cycle regulator is present and essential in Sinorhizobium meliloti. *J. Bacteriol.* 183, 3204–3210. doi: 10.1128/JB.183.10.3204-3210.2001

Bellefontaine, A. F., Pierreux, C. E., Mertens, P., Vandenhaute, J., Letesson, J. J., and De Bolle, X. (2002). Plasticity of a transcriptional regulation network among alpha-*proteobacteria* is supported by the identification of CtrA targets in Brucella abortus. *Mol. Microbiol.* 43, 945–960. doi: 10.1046/j.1365-2958.2002.02777.x

Biondi, E. G., Reisinger, S. J., Skerker, J. M., Arif, M., Perchuk, B. S., Ryan, K. R., et al. (2006). Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* 444, 899–904. doi: 10.1038/nature05321

Bird, T. H., and MacKrell, A. (2011). A CtrA homolog affects swarming motility and encystment in Rhodospirillum centenum. *Arch. Microbiol.* 193, 451–459. doi: 10.1007/s00203-011-0676-y

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., et al. (2010). The diversity and evolution of cell cycle regulation in alpha-*proteobacteria*: A comparative genomic analysis. *BMC Syst. Biol.* 4:52. doi: 10.1186/1752-0509-4-52

Christen, B., Abeliuk, E., Collier, J. M., Kalogeraki, V. S., Passarelli, B., Coller, J. A., et al. (2011). The essential genome of a bacterium. *Mol. Syst. Biol.* 7, 528. doi: 10.1038/msb.2011.58

Couturier, E., and Rocha, E. P. C. (2006). Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Mol. Microbiol.* 59, 1506–1518. doi: 10.1111/j.1365-2958.2006.05046.x

Degnen, S. T., and Newton, A. (1972). Chromosome replication during development in Caulobacter crescentus. *J. Mol. Biol.* 64, 671–680. doi: 10.1016/0022-2836(72)90090-3

Ding, H., Grüll, M. P., Mulligan, M. E., Lang, A. S., and Beatty, J. T. (2019). Induction of Rhodobacter capsulatus Gene Transfer Agent Gene Expression

Is a Bistable Stochastic Process Repressed by an Extracellular Calcium-Binding RTX Protein Homologue. *J. Bacteriol* 201, e430–e419. doi: 10.1128/JB.00430-19 *e00430-19,

Domian, I. J., Reisenauer, A., and Shapiro, L. (1999). Feedback control of a master bacterial cell-cycle regulator. *Proc. Natl. Acad. Sci.* 96, 6648–6653. doi: 10.1073/pnas.96.12.6648

Farrera-Calderon, R. G., Pallegar, P., Westbye, A. B., Wiesmann, C., Lang, A. S., and Beatty, J. T. (2020). The CckA-ChpT-CtrA phosphorelay controlling Rhodobacter capsulatus gene transfer agent (RcGTA) production is bi-directional and regulated by cyclic-di-GMP. *J. Bacteriol* 203, e525–e520. doi: 10.1128/JB.00525-20 *e00525-20,

Fioravanti, A., Fumeaux, C., Mohapatra, S. S., Bompard, C., Brilli, M., Frandi, A., et al. (2013). DNA Binding of the Cell Cycle Transcriptional Regulator GcrA Depends on N6-Adenosine Methylation in Caulobacter crescentus and Other Alpha*proteobacteria*. *PLoS Genet* 9:e1003541. doi: 10.1371/journal.pgen.1003541

Fogg, P. C. M. (2019). Identification and characterization of a direct activator of a gene transfer agent. *Nat. Commun* 10, 595. doi: 10.1038/s41467-019-08526-1

Francez-Charlot, A., Kaczmarczyk, A., and Vorholt, J. A. (2015). The branched CcsA/CckA-ChpT-CtrA phosphorelay of Sphingomonas melonis controls motility and biofilm formation. *Mol. Microbiol* 97, 47–63. doi: 10.1111/mmi.13011

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

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

Gonzalez, D., and Collier, J. (2013). DNA methylation by CcrM activates the transcription of two genes required for the division of Caulobacter crescentus. *Mol. Microbiol.* 88, 203–218. doi: 10.1111/mmi.12180

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

Greene, S. E., Brilli, M., Biondi, E. G., and Komeili, A. (2012). Analysis of the CtrA pathway in magnetospirillum reveals an ancestral role in motility in alpha*proteobacteria*. *J. Bacteriol.* 194, 2973–2986. doi: 10.1128/JB.00170-12

Guzzo, M., Castro, L. K., Reisch, C. R., Guo, M. S., and Laub, M. T. (2020). A CRISPR Interference System for Efficient and Rapid Gene Knockdown in Caulobacter crescentus. *mBio* 11, doi: 10.1128/mBio.02415-19 *Q,

Haakonsen, D. L., Yuan, A. H., and Laub, M. T. (2015). The bacterial cell cycle regulator GcrA is a σ70 cofactor that drives gene expression from a subset of methylated promoters. *Genes Dev.* 29, 2272–2286. doi: 10.1101/gad.270660.115

Hernández-Valle, J., Sanchez-Flores, A., Poggio, S., Dreyfus, G., and Camarena, L. (2020). The CtrA Regulon of Rhodobacter sphaeroides Favors Adaptation to a Particular Lifestyle. *J. Bacteriol* 202, doi: 10.1128/JB.00678-19 *Q,

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

Hynes, A. P., Mercer, R. G., Watton, D. E., Buckley, C. B., and Lang, A. S. (2012). DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the Rhodobacter capsulatus gene transfer agent. *RcGTA. Mol. Microbiol.* 85, 314–325. doi: 10.1111/j.1365-2958.2012.08113.x

Irvine, S., Bunk, B., Bayes, H. K., Spröer, C., Connolly, J. P. R., Six, A., et al. (2019). Genomic and transcriptomic characterization of *Pseudomonas aeruginosa* small colony variants derived from a chronic infection model. *Microb. Genomics* 5, e000262. doi: 10.1099/mgen.0.000262

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

Jung, A., Raßbach, A., Pulpetta, R. L., van Teeseling, M. C. F., Heinrich, K., Sobetzko, P., et al. (2019). Two-step chromosome segregation in the stalked budding bacterium Hyphomonas neptunium. *Nat. Commun.* 10, 1–16. doi: 10.1038/s41467-019-11242-5

Kopejtka, K., Lin, Y., Jakubovièová, M., Koblížek, M., and Tomasch, J. (2019). Clustered Core- and Pan-Genome Content on Rhodobacteraceae Chromosomes. *Genome Biol. Evol.* 11, 2208–2217. doi: 10.1093/gbe/evz138

Koppenhöfer, S., Wang, H., Scharfe, M., Kaever, V., Wagner-Döbler, I., and Tomasch, J. (2019). Integrated Transcriptional Regulatory Network of Quorum Sensing. *Replication Control, and SOS Response in Dinoroseobacter shibae. Front. Microbiol.* 10, 803. doi: 10.3389/fmicb.2019.00803

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

Lasker, K., von Diezmann, L., Zhou, X., Ahrens, D. G., Mann, T. H., Moerner, W. E., et al. (2020). Selective sequestration of signalling proteins in a membraneless organelle reinforces the spatial regulation of asymmetry in Caulobacter crescentus. *Nat. Microbiol.* 5, 418–429. doi: 10.1038/s41564-019-0647-7

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

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

Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12, doi: 10.1186/1471-2105-12-124 *Q,

Leicht, O., van Teeseling, M. C. F., Panis, G., Reif, C., Wendt, H., Viollier, P. H., et al. (2020). Integrative and quantitative view of the CtrA regulatory network in a stalked budding bacterium. *PLoS Genet.* 16:e1008724. doi: 10.1371/journal.pgen.1008724

Leung, M. M., Brimacombe, C. A., and Beatty, J. T. (2013). Transcriptional regulation of the Rhodobacter capsulatus response regulator CtrA. *Microbiol. U. K.* 159, 96–106. doi: 10.1099/mic.0.062349-0

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

Mann, T. H., and Shapiro, L. (2018). Integration of cell cycle signals by multi-PAS domain kinases. *Proc. Natl. Acad. Sci. U. S. A.* 115, E7166–E7173. doi: 10.1073/pnas.1808543115

Mercer, R. G., Callister, S. J., Lipton, M. S., Pasa-Tolic, L., Strnad, H., Paces, V., et al. (2010). Loss of the Response Regulator CtrA Causes Pleiotropic Effects on Gene Expression but Does Not Affect Growth Phase Regulation in Rhodobacter capsulatus. *J. Bacteriol.* 192, 2701–2710. doi: 10.1128/JB.00160-10

Mercer, R. G., and Lang, A. S. (2014). Identification of a predicted partner-switching system that affects production of the gene transfer agent RcGTA and stationary phase viability in Rhodobacter capsulatus. *BMC Microbiol* 14, doi: 10.1186/1471-2180-14-71 *Q,

Mercer, R. G., Quinlan, M., Rose, A. R., Noll, S., Beatty, J. T., and Lang, A. S. (2012). Regulatory systems controlling motility and gene transfer agent production and release in Rhodobacter capsulatus. *FEMS Microbiol. Lett.* 331, 53–62. doi: 10.1111/j.1574-6968.2012.02553.x

Miller, T. R., and Belas, R. (2006). Motility is involved in Silicibacter sp. TM1040 interaction with dinoflagellates. *Environ. Microbiol.* 8, 1648–1659. doi: 10.1111/j.1462-2920.2006.01071.x

Mohapatra, S. S., Fioravanti, A., Vandame, P., Spriet, C., Pini, F., Bompard, C., et al. (2020). Methylation-dependent transcriptional regulation of crescentin gene (creS) by GcrA in Caulobacter crescentus. *Mol. Microbiol.* 114, 127–139. doi: 10.1111/mmi.14500

Muñoz-Gómez, S. A., Hess, S., Burger, G., Lang, B. F., Susko, E., Slamovits, C. H., et al. (2019). An updated phylogeny of the Alpha*proteobacteria* reveals that the parasitic Rickettsiales and Holosporales have independent origins. *eLife* 8, e42535. doi: 10.7554/eLife.42535

Narula, J., Kuchina, A., Lee, D. Y. D., Fujita, M., Süel, G. M., and Igoshin, O. A. (2015). Chromosomal Arrangement of Phosphorelay Genes Couples Sporulation and DNA Replication. *Cell* 162, 328–337. doi: 10.1016/j.cell.2015.06.012

Panis, G., Murray, S. R., and Viollier, P. H. (2015). Versatility of global transcriptional regulators in alpha-*Proteobacteria*: from essential cell cycle control to ancillary functions. *FEMS Microbiol. Rev.* 39, 120–133. doi: 10.1093/femsre/fuu002

Poncin, K., Gillet, S., and De Bolle, X. (2018). Learning from the master: targets and functions of the CtrA response regulator in Brucella abortus and other alpha-*proteobacteria*. *FEMS Microbiol. Rev.* 42, 500–513. doi: 10.1093/femsre/fuy019

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.* 95, 120–125. doi: 10.1073/pnas.95.1.120

Reisenauer, A., and Shapiro, L. (2002). DNA methylation affects the cell cycle transcription of the CtrA global regulator in Caulobacter. *EMBO J.* 21, 4969–4977. doi: 10.1093/emboj/cdf490

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. (2004). The replication-related organization of bacterial genomes. *Microbiology* 150, 1609–1627. doi: 10.1099/mic.0.26974-0

Siam, R., and Marczynski, G. T. (2000). Cell cycle regulator phosphorylation stimulates two distinct modes of binding at a chromosome replication origin. *EMBO J.* 19, 1138–1147. doi: 10.1093/emboj/19.5.1138

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

Slager, J., and Veening, J.-W. (2016). Hard-Wired Control of Bacterial Processes by Chromosomal Gene Location. *Trends Microbiol.* 24, 788–800. doi: 10.1016/j.tim.2016.06.003

Sobetzko, P., Travers, A., and Muskhelishvili, G. (2012). Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc. Natl. Acad. Sci. U. S. A* 109, doi: 10.1073/pnas.1108229109 *Q,

Soler-Bistué, A., Mondotte, J. A., Bland, M. J., Val, M.-E., Saleh, M.-C., and Mazel, D. (2015). Genomic Location of the Major Ribosomal Protein Gene Locus Determines *Vibrio cholerae* Global Growth and Infectivity. *PLoS Genet.* 11:e1005156. doi: 10.1371/journal.pgen.1005156

Tomasch, J., Wang, H., Hall, A. T. K., Patzelt, D., Preusse, M., Petersen, J., et al. (2018). Packaging of Dinoroseobacter shibae DNA into gene transfer agent particles is not random. *Genome Biol. Evol* 10, doi: 10.1093/gbe/evy005 *Q,

Touchon, M., and Rocha, E. P. C. (2016). Coevolution of the organization and structure of prokaryotic genomes. *Cold Spring Harb. Perspect. Biol* 8, doi: 10.1101/cshperspect.a018168 *Q,

Tsokos, C. G., Perchuk, B. S., and Laub, M. T. (2011). A Dynamic Complex of Signaling Proteins Uses Polar Localization to Regulate Cell-Fate Asymmetry in Caulobacter crescentus. *Dev. Cell* 20, 329–341. doi: 10.1016/j.devcel.2011.01.007

van Teeseling, M. C. F., and Thanbichler, M. (2020). Generating asymmetry in a changing environment: cell cycle regulation in dimorphic alpha*proteobacteria*. *Biol. Chem* 1, doi: 10.1515/hsz-2020-0235 *Q,

Wang, H., Ziesche, L., Frank, O., Michael, V., Martin, M., Petersen, J., et al. (2014). The CtrA phosphorelay integrates differentiation and communication in the marine alphaproteobacterium Dinoroseobacter shibae. *BMC Genomics* 15:130. doi: 10.1186/1471-2164-15-130

Yildirim, A., and Feig, M. (2018). High-resolution 3D models of Caulobacter crescentus chromosome reveal genome structural variability and organization. *Nucleic Acids Res.* 46, 3937–3952. doi: 10.1093/nar/gky141

Zan, J., Heindl, J. E., Liu, Y., Fuqua, C., and Hill, R. T. (2013). The CckA-ChpT-CtrA phosphorelay system is regulated by quorum sensing and controls flagellar motility in the marine sponge symbiont Ruegeria sp. KLH11. *PloS One* 8:e66346. doi: 10.1371/journal.pone.0066346

Zweiger, G., Marczynski, G., and Shapiro, L. (1994). A Caulobacter DNA Methyltransferase that Functions only in the Predivisional Cell. *J. Mol. Biol.* 235, 472–485. doi: 10.1006/jmbi.1994.1007

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: 07 May 2021
doi: 10.3389/fmicb.2021.626874

Distinct H_2O_2-Scavenging System in *Yersinia pseudotuberculosis*: KatG and AhpC Act Together to Scavenge Endogenous Hydrogen Peroxide

Fen Wan[1†], Xue Feng[2†], Jianhua Yin[3] and Haichun Gao[2*]*

[1] College of Laboratory Medicine, Hangzhou Medical College, Hangzhou, China, [2] Institute of Microbiology and College of Life Sciences, Zhejiang University, Hangzhou, China, [3] College of Biotechnology and Bioengineering, Zhenjiang University of Technology, Hangzhou, China

OPEN ACCESS

Edited by:
Jianping Xie,
Southwest University, China

Reviewed by:
Paula Maria Tribelli,
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
Xihui Shen,
Northwest A and F University, China
Hui Wang,
Nanjing Agricultural University, China

***Correspondence:**
Fen Wan
fenw@hmc.edu.cn
Haichun Gao
haichung@zju.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: 07 November 2020
Accepted: 22 March 2021
Published: 07 May 2021

Citation:
Wan F, Feng X, Yin J and Gao H (2021) Distinct H_2O_2-Scavenging System in Yersinia pseudotuberculosis: KatG and AhpC Act Together to Scavenge Endogenous Hydrogen Peroxide. Front. Microbiol. 12:626874. doi: 10.3389/fmicb.2021.626874

To colonize in the digestive tract of animals and humans, *Yersinia pseudotuberculosis* has to deal with reactive oxygen species (ROS) produced by host cells and microbiota. However, an understanding of the ROS-scavenging systems and their regulation in this bacterium remains largely elusive. In this study, we identified OxyR as the master transcriptional regulator mediating cellular responses to hydrogen peroxide (H_2O_2) in *Y. pseudotuberculosis* through genomics and transcriptomics analyses. OxyR activates transcription of diverse genes, especially the core members of its regulon, including those encoding catalases, peroxidases, and thiol reductases. The data also suggest that sulfur species and manganese may play a particular role in the oxidative stress response of *Y. pseudotuberculosis*. Among the three H_2O_2-scavenging systems in *Y. pseudotuberculosis*, catalase/peroxidase KatE functions as the primary scavenger for high levels of H_2O_2; NADH peroxidase *a*lkyl *hydroperoxide reductase* (AhpR) and catalase KatG together are responsible for removing low levels of H_2O_2. The simultaneous loss of both AhpC (the peroxidatic component of AhpR) and KatG results in activation of OxyR. Moreover, we found that AhpC, unlike its well-characterized *Escherichia coli* counterpart, has little effect on protecting cells against toxicity of organic peroxides. These findings provide not only novel insights into the structural and functional diversity of bacterial H_2O_2-scavenging systems but also a basic understanding of how *Y. pseudotuberculosis* copes with oxidative stress.

Keywords: *Yersinia*, OxyR, AhpC, catalase, oxidative stresss response

INTRODUCTION

Oxidative stress caused by reactive oxygen species (ROS), including superoxide (O_2^-), hydrogen peroxide (H_2O_2), and hydroxyl radical ($OH^·$), is inevitable to all organisms that respire oxygen (Li et al., 2016). These strong oxidants and/or radicals could damage virtually all biomolecules, such as nucleic acids, proteins, and lipids (Imlay, 2013). Naturally, detoxification of ROS is extremely critical for survival of diverse bacteria, pathogens in particular, because the host cells release ROS as a deadly weapon to defend against bacterial infections (Fang, 2004). Among ROS, H_2O_2 not only can be formed very rapidly endogenously (for example, 15 $\mu M\ s^{-1}$ in *Escherichia coli*) but also

enters into cells nearly freely, where it reacts with Fe^{2+} to generate most deadly hydroxyl radicals (Li et al., 2016). Because of these features, most bacteria employ multiple enzymes, including catalases, various peroxidases, and rubrerythrin, to keep the intracellular concentrations of H_2O_2 at safe limit (nanomolar levels; Mishra and Imlay, 2012).

Catalases are principal enzymes that protect bacterial cells against H_2O_2 stress by catalyzing H_2O_2 into water and oxygen (Mishra and Imlay, 2012). Two types of catalases are present in bacteria: one is bifunctional (or called catalase-peroxidase, HPI, e.g., *E. coli* KatG), with both catalytic and peroxidatic activities, and the other is monofunctional (HPII, e.g., *E. coli* KatE), with only catalytic activity (Mishra and Imlay, 2012). Although it is well recognized that catalases serve as the primary scavenger of H_2O_2, they generally work best with H_2O_2 at high levels (millimolar levels; Mishra and Imlay, 2012). One of the best characterized peroxidases is NADH peroxidase AhpR (named from *a*lkyl *h*ydro*p*eroxide *r*eductase), which could decompose both H_2O_2 and organic peroxides (OPs) in diverse bacteria (Jacobson et al., 1989; Niimura et al., 1995). Compared to catalase, AhpR is more efficient in scavenging low-level H_2O_2 and regarded as the primary scavenger of endogenous H_2O_2 (Seaver and Imlay, 2001). AhpR typically consists of two cytoplasmic proteins encoded by a single operon, the peroxidase component AhpC and its cognate reductase AhpF: the former is a bacterial representative of typical 2-Cys peroxiredoxins and the latter is a flavoprotein with NADH:disulfide oxidoreductase activity (Poole, 2005). AhpC reduces H_2O_2 by oxidizing the two conservative cysteines to form an intermolecular disulfide bond, which is reactivated by AhpF using NADH as reducing equivalent (Perkins et al., 2015). Atypical AhpR systems, in which AhpF is absent, have been found in some bacteria (Bryk et al., 2000; Baker et al., 2001). For instance, in *Helicobacter pylori*, AhpC is reduced by a thioredoxin/thioredoxin reductase (TrxB) system, and in *Mycobacterium tuberculosis*, AhpD reduces AhpC with electrons from NADH through dihydrolipoamide dehydrogenase and dihydrolipoamide succinyltransferase (Baker et al., 2001; Jaeger et al., 2004). Because of the essential roles of catalase and AhpR in H_2O_2-scavenging, bacterial strains lacking both together display a drastically elevated sensitivity to H_2O_2 and carry an apparent aerobic growth defect (Cosgrove et al., 2007; Ezraty et al., 2017).

In many bacteria, OxyR is a primary transcriptional regulator that mediates cellular response to oxidative stress (Imlay, 2013; Fu et al., 2015). As a LysR-family DNA-binding protein, OxyR senses and responds to H_2O_2 stress *via* formation of an intramolecular disulfide bond between two conserved cysteine residues [Cys 199 and Cys 208 within *E. coli* OxyR (*Ec*OxyR); Zheng et al., 1998]. The regulatory mode of OxyR varies among bacteria. For example, in *E. coli*, *Salmonella enterica* serovar Typhimurium (*S.* Typhi), and many other bacteria, OxyR functions as an activator only for major H_2O_2-scavenging proteins, such as catalases and AhpR (Imlay, 2008). However, it can also function in a dual-control manner (both a repressor and an activator for certain genes) in *Neisseria* and *Shewanella oneidensis* or as a repressor only in *Corynebacterium diphtheriae* (Ieva et al., 2008; Kim and Holmes, 2012; Jiang et al., 2014). OxyR proteins now have been generally regarded as a global regulator implicated in diverse biological processes, but the core members of their regulons are consistently constituted by operons responsible for H_2O_2 degradation such as catalases and peroxidases, iron-sequestering proteins, and thioredoxin and glutathione antioxidant systems (Imlay, 2015).

Yersinia pseudotuberculosis is a Gram-negative enteric pathogen that often causes self-limiting gastrointestinal disorders such as enteritis, diarrhea, and mesenteric lymphadenitis to animals and humans (Brady et al., 2020). During infection, environmental stress and host immunity reactions in human guts can introduce an increase in levels of ROS that *Y. pseudotuberculosis* cells encounter, thus requiring a sophisticated regulation of genes to reduce the intracellular H_2O_2 level to a safe line (Knaus et al., 2017). It has been reported that *Y. pseudotuberculosis* is able to combat oxidative stress by an unconventional way: it imports zinc to mitigate ROS by secreting a zincophore *via* the type VI secretion system (T6SS; Wang et al., 2016, 2020). However, how this bacterium copes with oxidative stress has not been systematically investigated and thus is poorly understood.

In this study, we carried out the transcriptomics analysis of *Y. pseudotuberculosis* YPIII in response to exogenous H_2O_2. We identified an OxyR analog in YPIII and found that its regulon members are highly conserved, including those encoding H_2O_2-scavenging enzymes, iron-sequestering proteins, and thiol-reducing systems. Our data support that OxyR of YPIII functions in an activator-only mode, and its absence causes a plating defect on LB agar, a result of insufficient production of H_2O_2-scavenging enzymes, including KatE, KatG, and AhpC. We further showed that KatE functions as the primary scavenger for H_2O_2, but both KatG and AhpC together play an essential role in decomposing low levels of H_2O_2. The simultaneous loss of both KatG and AhpC results in activation of OxyR, which in turn upregulates KatE production, conferring cells' enhanced resistance to H_2O_2. The findings presented here reveal a novel mechanism through which bacterial cells differentially exploit individual components of the H_2O_2-scavenging repertoire to increase fitness in harsh living environments.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Culture Conditions

All bacterial strains and plasmids used in this study are listed in **Table 1**. All chemicals were obtained from Sigma (Shanghai, China) unless otherwise noted. For genetic manipulation, *E. coli* and *Y. pseudotuberculosis* were grown in LB (containing 1% tryptone, 0.5% yeast extract, and 0.5% NaCl) under aerobic conditions at 37 and 26°C. When needed, the following chemicals were added to the growth medium: 2,6-diaminopimelic acid (DAP), 0.3 mM; ampicillin, 50 μg/ml; kanamycin, 50 μg/ml; gentamycin, 15 μg/ml; and streptomycin, 100 μg/ml.

TABLE 1 | Bacterial strains and plasmids used in this study.

Strain or plasmid	Description	Source or references
Strains		
E. coli		
DH5α	Host strain for cloning	Lab stock
WM3064	ΔdapA, donor strain for conjugation	W. Metcalf, UIUC
S. oneidensis		
MR-1	Wild type	
HG1070	ΔkatB derived from MR-1	Jiang et al., 2014
HG0958-1070	ΔkatBΔahpC derived from MR-1	Feng et al., 2020
Y. pseudotuberculosis		
YPIII	Wild type	Rosqvist et al., 1988
YPK_RS20585	ΔoxyR derived from YPIII	This study
YPK_RS14285	ΔkatE derived from YPIII	This study
YPK_RS17025	ΔkatG derived from YPIII	This study
YPK_RS16370	ΔahpC derived from YPIII	This study
YPK17025-16370	ΔkatGΔahpC derived from YPIII	This study
YPK14285-17025	ΔkatEΔkatG derived from YPIII	This study
YPK3CAT	ΔkatEΔkatGΔahpC derived from YPIII	This study
Plasmids		
pHGM01	$Ap^rGm^rCm^r$,att-based suicide vector	Jin et al., 2013
pHG101	Promoterless vector for complentation	Wu et al., 2011
pHGEI01	Km^r, integrative lacZ reporter vector	Fu et al., 2014
pHGEI01-katE	pHGEI01 containing the katE promoter	This study
pHGEI01-katG	pHGEI01 containing the katG promoter	This study
pHGEI01-ahpC	pHGEI01 containing the ahpC promoter	This study

Transcriptomics Analysis

For transcriptomics analysis, cell samples were prepared as described previously (Jiang et al., 2014). In brief, cultures of the mid-exponential phase (∼0.4 of OD_{600}, the same throughout the study) were subjected to the H_2O_2 treatment. H_2O_2 was added to a final concentration of 0.5 mM, and cells were collected before and 5 min after the addition. Cells were pelletted at 14,000 rpm for 30 s at room temperature and frozen immediately in liquid nitrogen. Three biological replicates under each condition were prepared. Total RNA was extracted with RNeasy Mini Kit (QIAGEN, Alameda, United States), and RNA sequencing (RNA-seq) analysis was conducted as before (Gao et al., 2004; Sun et al., 2020). Data analysis was conducted according to standard procedures used previously (Sun et al., 2020). Genes that passed statistical analysis by analysis of variance (ANOVA; $p < 0.05$) with Benjamini–Hochberg false discovery rate multiple-testing correction and showed two-fold difference between the H_2O_2-treated and untreated control samples were discussed in the study. NCBI SRA accession number is PRJNA671546 for raw transcriptomics analysis data.

Real-Time Quantitative RT-PCR

Quantitative reverse transcription-PCR (qRT-PCR) was performed to verify the expression of key OxyR regulon members with an ABI7300 96-well qRT-PCR system (Applied Biosystems) as described previously (Jiang et al., 2014). The expression of each gene was determined from four replicas in a single real-time qRT-PCR experiment. The cycle threshold (C_T) values for each gene of interest were averaged and normalized against the C_T value of the 16s rRNA gene, whose abundance was constant during the exponential phase. Relative abundance (RA) of each gene was standardized to the C_T values of the 16s rRNA gene using the equation $RA = 2^{-\Delta CT}$.

Mutagenesis, Complementation of Mutant Strains

Yersinia *pseudotuberculosis* in-frame deletion strains were constructed by the *att*-based Fusion PCR method as described previously (Jin et al., 2013). In brief, two fragments flanking the target gene were generated by PCR with primers containing *attB* and the gene-specific sequence, which were linked by a linker sequence *via* second round of PCR. The fusion fragments were integrated into plasmid pHGM01 by site-specific recombination using Gateway BP clonase II enzyme mix (Invitrogen). The resulting vectors were introduced in *E. coli* WM3064 and transferred to *Y. pseudotuberculosis* by conjugation. Integration of the mutagenesis constructs into the chromosome was selected by resistance to gentamycin and confirmed by PCR. Verified *trans*-conjugants were grown in LB broth without NaCl and plated on LB supplemented with 10% sucrose. Gentamycin-sensitive and sucrose-resistant colonies were screened by PCR for deletion of the target gene. To facilitate growth of mutants, catalase (from bovine liver, Sigma) was added onto the plates. All mutations were verified by sequencing the mutated regions.

Plasmid pHG-101 was used in genetic complementation of mutants as described before (Wu et al., 2011). For complementation of genes next to their promoter, a fragment containing the gene of interest and its native promoter was amplified by PCR and cloned into pHG-101. After sequencing verification, the resulting vectors were transferred into the relevant strains *via* conjugation.

Growth and Susceptibility to H_2O_2 or *t*-BHP

The spotting assay was used to evaluate the plating defect on LB plates. Cells of the mid-exponential phase were collected by centrifugation and adjusted to 10^9 cells/ml, which was set as the undiluted (dilution factor 0). Ten-fold serial dilutions were prepared with fresh LB medium. Five microliter of each dilution was spotted onto LB plates. The plates were incubated for 24 h or longer in the dark before being photographed. All experiments were repeated at least three times.

Disk diffusion assays to test for sensitivity to oxidative stress conditions were performed with *Y. pseudotuberculosis* strains. Two hundred microliter of mid-exponential phase cultures were spread onto LB plates, 6 mm (in diameter) paper disks loaded with 10 μl H_2O_2 or *tert*-butyl hydroperoxide (*t*-BHP) of various concentrations were placed onto the bacterial lawn grown for 6 h, and plates were incubated 26°C for 16 h.

H_2O_2 Quantification

H_2O_2 at high concentrations (>100 μM) was quantified using the ferrous ion oxidation-xylenol orange method (Wolff, 1994). In brief, cells of the mid-exponential phase grown in liquid LB were collected by centrifugation, washed in PBS, and resuspended in the same buffer to an OD_{600} of 0.1. H_2O_2 was added to final concentrations indicated in the figure legends. Cells were filtered out at different time points, and elutions were assayed immediately for the remaining H_2O_2.

H_2O_2 at low concentrations (<50 μM) was quantified using Amplex red fluorescent method (Seaver and Imlay, 2001). In the presence of H_2O_2, Amplex red can be oxidized by horseradish peroxidase to the fluorescent product resorufin. To measure H_2O_2, 200 μl of samples from cells grown in MS medium was mixed with 100 μl of stock solutions for Amplex red and horseradish peroxidase prepared the same as described elsewhere (Seaver and Imlay, 2001). Fluorescence was then measured in a Synergy 2 Pro200 Multi-Detection Microplate Reader (Tecan) and converted to H_2O_2 concentration using a curve obtained from standard samples.

β-Galactosidase Activity Assay

β-Galactosidase activity assay was used to determine gene expression. The sequence in sufficient length (∼400 bp) upstream of gene of interest was amplified and inserted in front of the full-length *E. coli lacZ* gene in plasmid pHGEI01 (Fu et al., 2014). The resulting plasmid was verified by sequencing, introduced into *E. coli* WM3064, and then conjugated with relevant *Y. pseudotuberculosis* strains. Cultures of the mid-exponential phase were collected by centrifugation, washed with PBS, and treated with lysis buffer (0.25 M Tris/HCl, 0.5% Triton X-100, and pH 7.5). Extracts were collected by centrifugation and applied for enzyme assay by adding *o*-nitrophenyl-β-*D*-galactopyranoside (4 mg/ml). Changes in absorption over time were monitored at 420 nM with a Synergy 2 Pro200 Multi-Detection Microplate Reader (Tecan), and the results were presented as Miller units.

Bioinformatics and Statistical Analyses

Multiple sequence alignment was carried out with Clustal Omega (Madeira et al., 2019). Sequence logos were generated by using WebLogo (Crooks et al., 2004). Three-dimensional structures of YpAhpC were predicted using Phyre[1] (Kelley et al., 2015). The predicted structures were then visualized by software Pymol (DeLano Scientific LLC). Genome screening for OxyR-binding sites based on established weight matrixes from various bacteria was performed using regulatory sequence analysis tools (RSATs; Medina-Rivera et al., 2015). For statistical analysis, Student's *t*-test was performed for pairwise comparisons of groups, and values are presented as means ± standard deviation (SD).

[1]http://www.sbg.bio.ic.ac.uk/∼phyre/

RESULTS

Genomics Analysis of *Y. pseudotuberculosis* YPIII With Respect to Oxidative Stress Response

To identify the oxidative stress response regulator(s) in YPIII, a BLASTp search of functional analogs of established oxidative stress-responding regulators, including *Ec*OxyR and SoxRS as well as *Bacillus subtilis* PerR, against the YPIII proteome was performed. While no homolog of *E. coli* SoxRS or of *B. subtilis* PerR was found, YPIII possesses a highly confident homolog of *Ec*OxyR, YPK_RS20585 (BLASTp *E*-value = 0; identities, 88%; **Supplementary Figure 1**). Like all 2-Cys OxyRs, YPK_RS20585 contains two conserved cysteine residues (Cys199 and Cys208 within both YPK_RS20585 and *Ec*OxyR) implicated in the activation by disulfide bond formation, suggesting a possible role in the oxidative stress response of YPIII, and therefore we named it OxyR (*Yp*OxyR). However, the sequence similarities between *Yp*OxyR and dual-activity OxyRs, such as that of *S. oneidensis*, are substantially lower (*E*-value = 7e-47; identities, 33%; **Supplementary Figure 2**), implying that *Yp*OxyR might function as an activator only.

The current knowledge on bacterial oxidative stress response is most well developed in *E. coli*, whose OxyR regulon is composed of over 20 operons (Imlay, 2015). The most important and conserved *Ec*OxyR regulon members encode proteins involved in detoxification and prevention and/or repair of oxidative damage, such as catalases and peroxidases, iron-sequestering proteins, thioredoxin, and glutathione antioxidant systems. Similarly, the YPIII genome encodes two catalases (HPII KatE and HPI KatG), iron-sequestering protein Dps, and a complete set of thioredoxin and glutathione antioxidant proteins, including TrxA (thioredoxin), TrxB (thioredoxin-disulfide reductase), TrxC (thioredoxin), GrxA (glutaredoxin), YPK_RS20590 (glutathione peroxidase), and GorA (glutathione-disulfide reductase; **Table 2**).

Perhaps one of the most striking observations is that YPIII seemingly lacks the counterpart of *E. coli* AhpF, the cognate reductase for peroxidase component AhpC of AhpR (Tartaglia et al., 1990; Poole and Ellis, 1996). A BLASTp search using *E. coli* AhpCF against the YPIII proteome revealed a single putative homolog for AhpC and AhpF, YPK_RS16370 (peroxiredoxin C; *E*-value, 3e-38) and TrxB (*E*-value, 3e-42), which are not in proximity on the chromosome. Combining the fact that AhpCF is encoded by a single operon in other bacteria and TrxB is clearly the counterpart of *Ec*TrxB (*E*-value, 0) but not *Ec*AhpF, these data strongly suggest that YPIII lacks a conventional AhpF. Furthermore, the YPIII genome also lacks a gene for a homolog of Ohr, the primary enzyme that decomposes OPs. Given the involvement of AhpR in scavenging OPs, thus it is particularly

TABLE 2 | Highly regulated genes involved in the response to H_2O_2 and genes encoding analogs of OxyR regulon members.

Locus tag	Gene	OxyR regulon[a]	YpOxyR motif W[b]	Fold Change[c]	COG type[d]	Description
Top 20 up-regulated genes						
YPK_RS20590		Y		689.5	O	Glutathione peroxidase
YPK_RS20595				89.3		Dihydrolipoyl dehydrogenase
YPK_RS16760	*trxC*	Y		75.4	J	Thioredoxin TrxC
YPK_RS08095	*dps*	Y	15.4	48.3	P	Iron-sequestering protein Dps
YPK_RS17250	*cysC*			38.5	P	Adenylyl-sulfate kinase
YPK_RS22535				34.0		Hypothetical protein
YPK_RS17260	*cysD*			32.0		Sulfate adenylyltransferase subunit CysD
YPK_RS17255	*cysN*			28.3		Sulfate adenylyltransferase subunit CysN
YPK_RS14285	*katE*	Y	22.7	27.6	P	Catalase, OxyR regulon
YPK_RS13675	*grxA*	Y	12.6	22.6	S	GrxA family glutaredoxin
YPK_RS17265	*cobA*			18.0		Uroporphyrinogen-III C-methyltransferase
YPK_RS00550	*gorA*	Y		16.2	C	Glutathione-disulfide reductase
YPK_RS13735				16.2		Isopenicillin N synthase family oxygenase
YPK_RS13730				15.4		ABC transporter substrate-binding protein
YPK_RS03760				13.2		TonB-dependent receptor
YPK_RS09490				13.1		CMD domain-containing protein
YPK_RS07100	*cysA*			13.0		Sulfate/thiosulfate ABC transporter ATP-binding protein CysA
YPK_RS16765				12.7		DTW domain-containing protein
YPK_RS10680	*zwf*			12.3	G	Glucose-6-phosphate dehydrogenase
YPK_RS07190	*cysK*			11.9	G	Cysteine synthase A
Other OxyR regulon members						
YPK_RS17025	*katG*	Y	23.1	7.2	P	Catalase/peroxidase HPI
YPK_RS20585	*oxyR*	Y		4.5		DNA-binding transcriptional regulator OxyR
YPK_RS13445	*trxB*	Y	21.0	4.5	O	Thioredoxin-disulfide reductase
YPK_RS16370	*ahpC*	Y	21.1	3.5	O	Peroxiredoxin C
YPK_RS09285	*sufA*	Y		2.598	P	Fe-S cluster assembly protein

[a] *Y, OxyR regulon members in at least three bacterial genera (E. coli, S. Typhi, and Neisseria meningitidis).*
[b] *Weight values of predicted YpOxyR regulon members by using regulatory sequence analysis tool (RSAT).*
[c] *Fold changes in transcripts (the ratio of the H_2O_2-treated sample of the wild type to the control) revealed in transcriptomics analysis.*
[d] *COG: O: posttranslational modification; J: translation, ribosomal and biogenesis; P: metabolism; S: Function unknown; C: energy production; and G: carbohydrate transport and metabolism.*

important and interesting to understand the physiological role of AhpC in YPIII.

Transcriptomics Analysis of YPIII in Response to H_2O_2 Stress

In order to gain a comprehensive understanding of the cellular response of YPIII to H_2O_2, we performed an RNA-seq analysis to obtain gene expression changes resulting from the H_2O_2 treatment. Our early stress response studies suggest that the best concentration of H_2O_2 for transcriptomics analyses would be the dosage at which the agent arrests growth of cells but does show an evident killing effect on them (Gao et al., 2004; Jiang et al., 2014). To determine the concentration, the minimum inhibitory concentration (MIC) of H_2O_2 against the wild-type strain of YPIII was assessed. The results revealed MIC to be 4 mM, which is four times lower than that of *E. coli* (**Figure 1A**), indicating that YPIII is significantly more sensitive to H_2O_2 than *E. coli*. Then, impacts of H_2O_2 over a range of concentrations (under MIC) on growth and viability of actively growing cells (∼0.4 of OD_{600}) were examined. Upon addition of H_2O_2 at 0.2 mM or higher, growth paused immediately and resumed after lag periods that increase with the H_2O_2 concentrations (**Figure 1B**). Additionally, viability assays revealed that H_2O_2 at 0.5 mM or lower did not show a significant killing effect (**Supplementary Figure 3**). Based on these observations, for RNA-seq transcriptomics analysis, we collected cells before and 5 min after the addition of

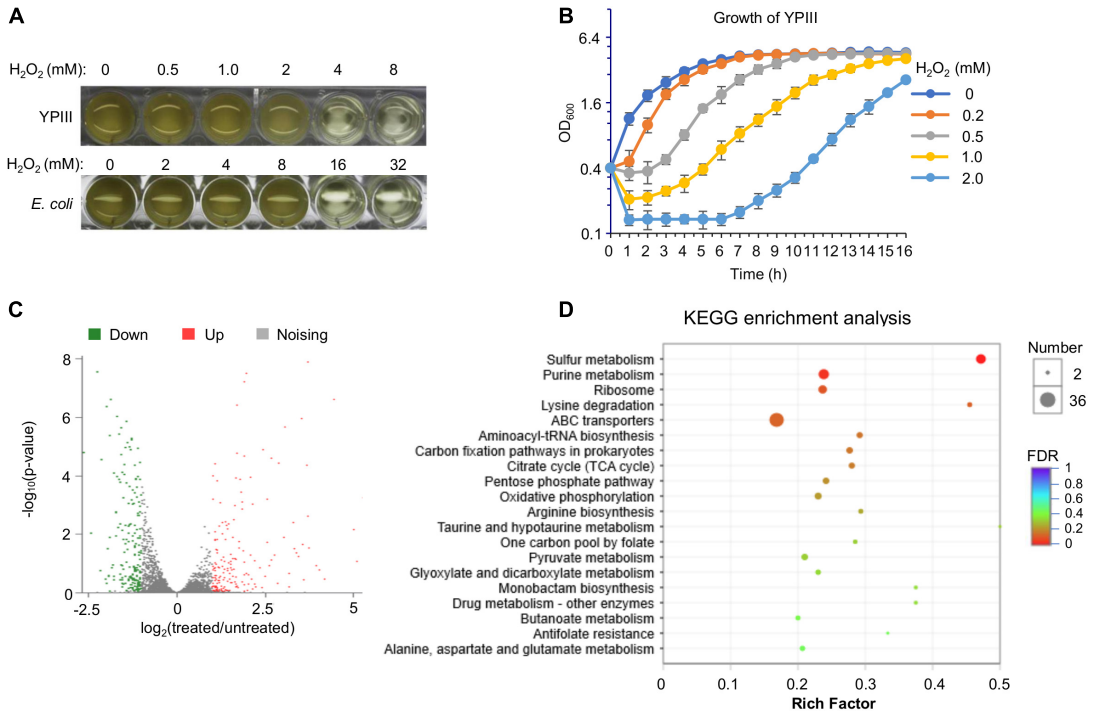


FIGURE 1 | Characteristics of *Y. pseudotuberculosis* in response to H_2O_2. **(A)** Minimum inhibitory concentration (MIC) assay. Mid-exponential phase cultures (OD_{600} of ~0.4) were used to inoculate each well to an OD_{600} of 0.01, and MIC was determined 16 h later. **(B)** Impact of H_2O_2 on growing cells. H_2O_2 was added to mid-exponential phase cultures of the wild type to the final concentrations as indicated. Growth was monitored by recording OD_{600} values. **(C)** Volcano plot of the different expression genes in wild-type cells between before and after the H_2O_2 treatment (0.5 mM). **(D)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of differentially expressed genes in wild-type cells between before and after the H_2O_2 treatment.

0.5 mM H_2O_2 as the untreated control and the treated samples, respectively.

In total, 364 genes displayed significant differences in transcription levels between the H_2O_2-treated and untreated control samples (**Supplementary Table 1**). Among these genes, 186 genes were up-regulated, while 178 genes were down-regulated (**Figure 1C**). Genes displaying significant differences in expression levels due to H_2O_2-induced oxidative stress were observed in almost every functional category, indicating that YPIII transcriptionally responds to H_2O_2 in a global scale (**Supplementary Table 1**). The high quality of the expression data was validated with a statistical analysis as previously described (Gao et al., 2004) and by real-time qRT-PCR. Eight genes were selected for analysis with the same RNA samples used in the RNA-seq based on the level and reproducibility of changes observed in the RNA-seq experiments. A high level of concordance (R^2 = 0.96) was observed between RNA-seq and real-time qRT-PCR data despite quantitative differences in the level of change (**Supplementary Figure 4**), suggesting that the RNA-seq results are an accurate reflection of the gene expression profile. The highly induced included many genes encoding proteins combating oxidative stress. In contrast, genes encoding metabolic enzymes were down-regulated in general (**Figure 1D**), an observation that reflects a paused/reduced growth rate rather than a specific response to the stress has been commonly found in other stress response studies (Gao et al., 2004; Jiang et al., 2014). Intriguingly, many ABC transporters and sulfur metabolic enzymes are among differentially expressed genes, implicating that sulfur species and certain small molecules may have a particular significance in the oxidative stress response of YPIII.

To stay focused, here we only discussed genes whose transcription was significantly induced upon the addition of H_2O_2, especially those specific to H_2O_2-induced oxidative stress (**Table 2**). With respect to OxyR regulon members, our transcriptomics data revealed that YPK_RS20590 (glutathione peroxidase), *trxC*, *dps*, *katE*, *grxA*, and *gorA* were among the top 20 most up-regulated genes in H_2O_2-treated cells (**Table 2**). Other putative OxyR regulon members that were induced but not among the top 10 included *katG*, *trxB*, *ahpC*, and *sufABC* (Fe-S cluster assembly proteins; **Table 2**). Intriguingly, the entire thioredoxin and glutathione antioxidant systems, which play a supporting role in combating oxidative stress in many bacteria (Feng et al., 2019), were found to be highly induced by H_2O_2 (for instance, YPK_RS20590 was induced 689-fold), implying that these systems may have more significant contribution in protecting YPIII cells from oxidative damage. It is also worth mentioning that the *oxyR* gene was transcribed at substantially elevated levels (4.5-fold) upon H_2O_2 stress (**Table 2**), suggesting that the regulator *per se* is also subjected to quantity control, in addition to activity transformation upon oxidation.

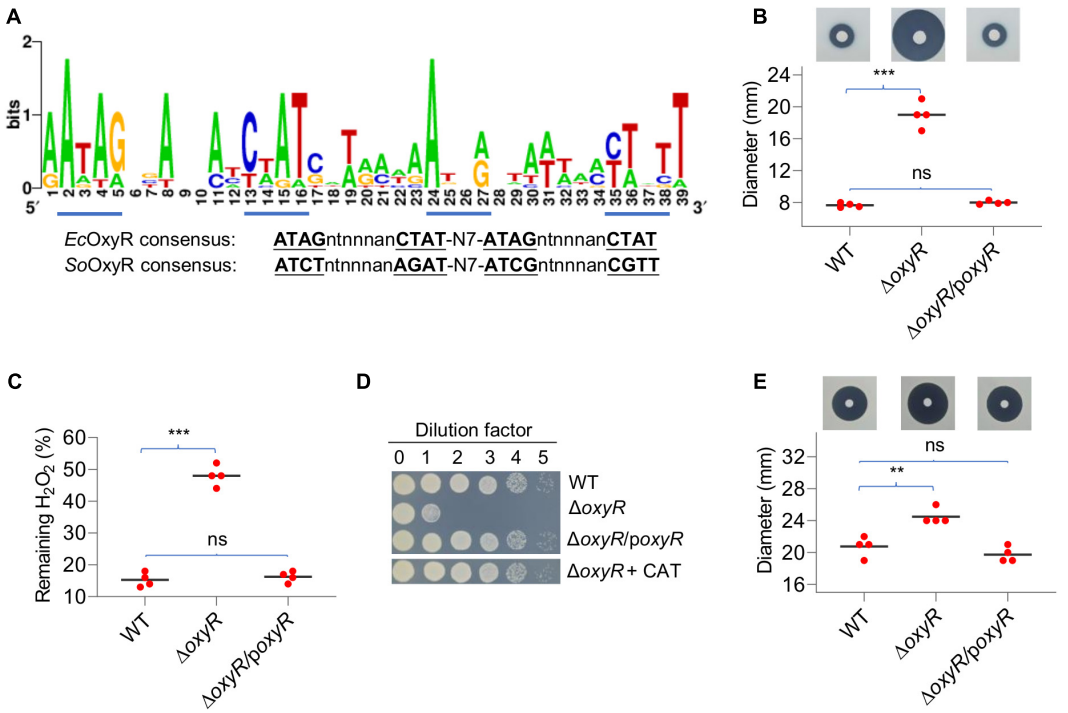


FIGURE 2 | Characteristics of the *Y. pseudotuberculosis oxyR* mutant. **(A)** OxyR-binding motifs in YPIII derived from the top 20 members of the predicted regulon. The matrix for screening the YPIII genome was generated using most conserved members of OxyR regulons of diverse bacteria. Tetranucleotide sequences are underlined based on the *E. coli* and *S. oneidensis* OxyR consensus. **(B)** Disk diffusion assay for H_2O_2. Paper disks loaded with 10 μl of 5 M H_2O_2 were placed onto bacterial lawns (pregrown for 6 h). Results shown are from 24 h after the disks were placed. The sensitivity was represented by the diameter of the inhibition zone. p*oxyR* represents a copy of *oxyR* to be expressed *in trans* for complementation. **(C)** Measuring H_2O_2-scavenging rates. Cells grown to the mid-exponential phase were collected, washed, and suspended to an OD_{600} of ∼0.4. 5 min after the addition of 0.5 mM H_2O_2, the remaining H_2O_2 was determined. The data were normalized to the initial concentration. **(D)** Plating defects of the *oxyR* mutant. Mid-exponential phase cultures were adjusted to ∼10^8 CFU/ml and diluted by dilution factor (10-fold serial dilution), and 5 μl of each dilution was spotted onto LB plates. Photos were taken 24 h after plating. CAT, catalase (2,000 Units/ml). Experiments were performed at least four times, and representative results were presented. **(E)** Disk diffusion assay for *tert*-butyl hydroperoxide (*t*-BHP). Paper disks loaded with 10 μl of 5 M *t*-BHP. In panels **B**, **C**, **E**, asterisks indicate statistically significant differences of the values compared ($n = 4$; ns, not significant; $^{**}p < 0.01$; and $^{***}p < 0.001$).

Among the remaining top 20 most significantly induced genes, those for sulfur species transport and metabolism (*cysC*, *cysD*, *cysN*, *cysA*, and *cysK*) drew special attention. In many organisms, sulfur species are critically involved in cell protection against oxidative damage (Wu et al., 2015; Peng et al., 2017; Korshunov et al., 2020). Along with highly induced transcription of *cysJ*, *cysW*, *cysT*, and *cysI* in H_2O_2-challenged cells, these data suggest that sulfur species may be critical for oxidative stress response in YPIII. In addition, we also noticed that genes encoding iron/manganese ABC transport system (*yfeA*, *yfeB*, *yfeC*, and *yfeD*) were transcribed at significantly elevated levels upon the treatment. This may not be surprising given that the intracellular Mn/Fe ratios are correlated with bacterial resistance to oxidative stress (Chandrangsu et al., 2017). To maintain the activity of vulnerable mononuclear iron proteins in the presence of H_2O_2, many bacteria replace the iron atom with manganese as the active center (Anjem et al., 2009; Imlay et al., 2019). These data concur well with the result from the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis in which sulfur metabolic pathways are among the most enriched (**Figure 1D**). Together with the missing of AhpF, these results suggest that the mechanism underlying the response of YPIII to H_2O_2-induced oxidative stress carries novel characteristics that are different from the classical mode established in *E. coli*.

Characterization of YPIII *oxyR* Mutant

In many bacteria, OxyR-binding motifs have been determined. Given sequence and function conservation, it is not surprising that these motifs are highly similar (Wan et al., 2018). To predict the operons under the direct control of *Yp*OxyR, we constructed a matrix for the OxyR-binding motif with verified OxyR-binding sequences from closely related bacterial species and used it to screen the entire genome of YPIII by RSAT (Medina-Rivera et al., 2015). In total, 38 putative *Yp*OxyR-binding sites were identified with weight values greater than 7, an arbitrary cutoff implemented in RSAT (**Supplementary Table 2**). On the top of the list are well-established OxyR regulon members, including *katG*, *katE*, *ahpC*, *trxB*, *grxA*, *dps*, and *trxC*, supporting that *Yp*OxyR recognizes a similar DNA motif (**Figure 2A**). For the remaining genes on the list, we supposed that many of them may not belong to the *Yp*OxyR regulon because weight values of

their binding motifs are relatively low (≤10). But this requires further verification.

To determine the impacts of OxyR on oxidative stress response, an *oxyR* in-frame deletion strain (Δ*oxyR*) was constructed and characterized in YPIII. In liquid LB, Δ*oxyR* grew indistinguishably from the wild type (**Supplementary Figure 5**), indicating that *Yp*OxyR is not required for normal growth. We then assessed the role of *Yp*OxyR in combating oxidative stress. Disk diffusion assays demonstrated that Δ*oxyR* was more sensitive to H_2O_2, generating inhibition zones that were substantially larger than those of the wild type (**Figure 2B**). The phenotype can be confidently attributed to the loss of OxyR given successful genetic complementation. To provide further support to this, we compared the mutant and the wild type with respect to H_2O_2 consumption. Cells at the mid-exponential phase were collected, adjusted to identical densities, and incubated with 0.2 mM H_2O_2. As shown in **Figure 2C**, the Δ*oxyR* strain degraded H_2O_2 at a rate much lower than the wild type did; 14% and 49% of H_2O_2 remained 2 min after the reaction started for the wild type and the mutant, respectively. These results indicate that the *oxyR* deletion greatly impaired the ability of YPIII to remove exogenous H_2O_2. Furthermore, we also confirmed that the Δ*oxyR* strain carried a plating defect on LB agar plates (**Figure 2D**), a common phenotype of the *oxyR* mutants in bacteria such as *E. coli* and *S. oneidensis* because of compromised H_2O_2-scavenging capacity (Jiang et al., 2014). As exogenous catalase completely suppressed the plate defect of the Δ*oxyR* strain, it is reasonable to propose that the same mechanism is responsible for the YPIII *oxyR* mutant (Jiang et al., 2014; Shi et al., 2015).

In many bacteria lacking Ohr, such as *E. coli* and *S.* Typhi, OxyR also mediates cellular response to OPs (Imlay, 2015). Clearly, this is also true in the case of YPIII because the *oxyR* mutant exhibited significantly increased sensitivity to *t*-BHP, a widely used representative OP (**Figure 2E**). All together, these data conclude that YPIII employs transcriptional regulator OxyR to mediate an oxidative stress response.

KatE Is the Primary Catalase in YPIII Positively Regulated by OxyR

In general, multiple catalases are encoded in a bacterial genome, but only one exhibits predominant H_2O_2-scavenging activity *in vivo* (Mishra and Imlay, 2012). To test whether this also holds true in YPIII, we assessed the contribution of catalases to decomposition of H_2O_2. Strains lacking *katE* and *katG* were generated and characterized with the disk diffusion assay and the H_2O_2-decomposing assay. As shown in **Figure 3A**, the *katE* mutant (Δ*katE*) was hypersensitive to H_2O_2, generating inhibition zones that were approximately two and a half times larger than those of the wild type. On the contrary, the difference in zone sizes between the wild type and Δ*katG* was insignificant. Consistent with the results in disk diffusion assay, the *katE* deletion resulted in severely impaired ability to decompose H_2O_2, with nearly 60% of H_2O_2 remaining 2 min after the reaction started (**Figure 3B**). In the case of Δ*katG*, the assay revealed that this mutation slightly but still significantly compromised the ability of YPIII to decompose H_2O_2 when compared with that of Δ*katE*, validating that KatG is also functioning in YPIII.

Impacts of *Yp*OxyR on the expression of some of its regulon members predicted by the *in silico* analysis, including four high-confident (weight value > 12; *trxB*, *dps*, *katE*, and *grxA*) and two low-confident genes (weight value ≈9; *acnA* and *zwf*; **Table 2** and **Supplementary Table 2**), were assessed by qRT-PCR. In cells of Δ*oxyR* prepared the same as for transcriptomics analysis, we found that none of the four high-confident genes were responsive to the H_2O_2 treatment, whereas two control genes, *yfeC* and *tauA*, which lack a predicted OxyR-binding motif, were upregulated upon exposure to H_2O_2 as in the wild type (**Figure 3C**). In the case of *acnA* and *zwf*, however, we found that the loss of OxyR negatively affected but not completely abolished their transcription in response to H_2O_2 (**Figure 3C**). Apparently, these data support that the predicted DNA motifs largely determine OxyR-dependent transcription.

Given the central role of catalases in protecting cells from H_2O_2 damage, we further compared the expression of *katE* and *katG* genes in the wild type and Δ*oxyR* strains by measuring activities of *katE* and *katG* promoters with an integrative *lacZ* reporter system (**Figure 3D**). In line with the transcriptomics data, the expression of *katE* and *katG* increased nearly six-fold in the wild-type cells when challenged by H_2O_2. While in the Δ*oxyR* strain, both genes were no longer responsive to H_2O_2, being expressed at levels even lower than those observed in the untreated wild-type cells. In summary, these data validate that expression of both KatE and KatG is under the positive control of OxyR, and KatE functions as the primary catalase to cope with H_2O_2 stress.

AhpC Possesses Scavenging Activity Against H_2O_2 but Not Organic Peroxides

The genomics analysis as presented above indicates that YPIII possesses an atypical AhpR, whose AhpF is missing. Like *katE* and *katG*, the *ahpC* gene is positively regulated by OxyR (**Figure 3C**), and therefore, the protein is likely produced when OxyR is activated. Given the importance of AhpR in scavenging both H_2O_2 and OPs, we reasoned that AhpC may be still functional in YPIII because it can also be reduced by TrxB and/or glutathione (GSH) reductase (Gor), albeit low in efficiency (Feng et al., 2020). To test this, we knocked out the *ahpC* gene from mutants lacking KatE, KatG, or both. The removal of AhpC alone had no detectable impact on H_2O_2 consumption (**Figure 4A**), an expected result because of the dominance of catalase in H_2O_2 degradation (Mishra and Imlay, 2012). In the absence of KatE, the effect of the AhpC loss became significant, indicating that AhpC does have some capacity of H_2O_2 degradation. Compared to Δ*katE*Δ*ahpC*, the loss of both catalases (Δ*katE*Δ*katG*) significantly impaired the ability of YPIII to decompose H_2O_2 (**Figure 4A**). Furthermore, additional removal of AhpC from Δ*katE*Δ*katG* nearly completely abolished the H_2O_2-decomposing capacity of YPIII. Therefore, these data conclude that all of KatE, KatG, and AhpC are capable of decomposing H_2O_2 in YPIII. Intriguingly, we found that

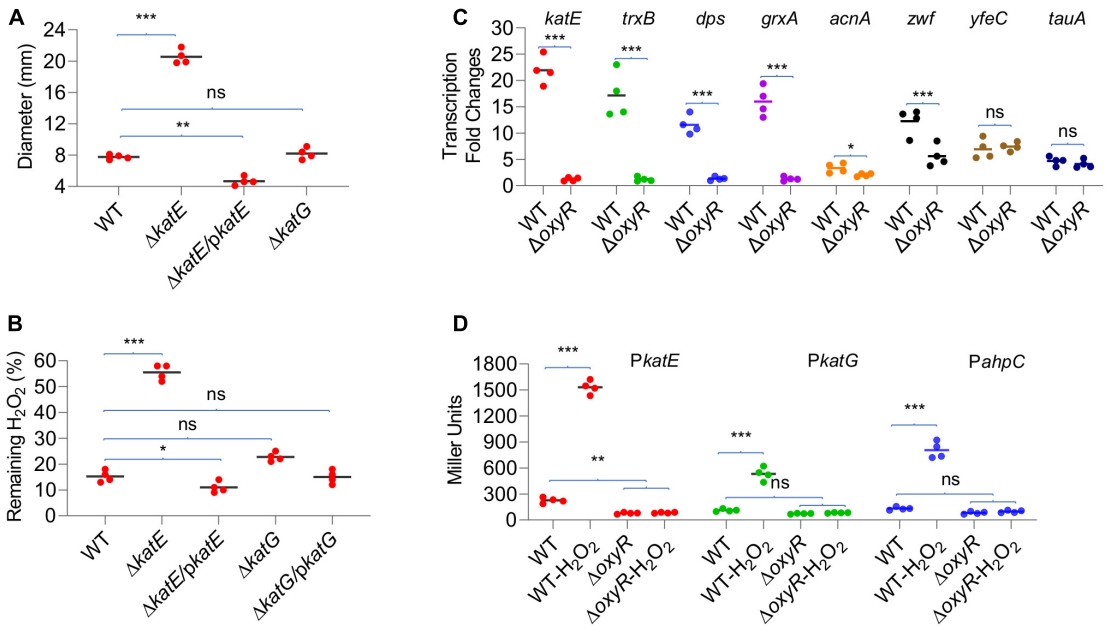


FIGURE 3 | Role of catalases in decomposition of H_2O_2. **(A)** Disk diffusion assay of catalase mutants. **(B)** Measuring H_2O_2-scavenging rate of catalase mutants. **(C)** qRT-PCR analysis of transcription differences upon *oxyR* deletion. Genes with high-confident (*katE*, *trxB*, *dps*, and *grxA*) and low-confident (*acnA* and *zwf*) OxyR-binding motifs and without (*yfeC* and *tauA*) were examined. Signal intensities were processed, and the fold changes [values of the H_2O_2-treated wild type (WT) and Δ*oxyR*/values of the untreated WT] were calculated according to the method described in *Materials and Methods*. **(D)** Impacts of OxyR on the expression of *katE*, *katG*, and *ahpC*. Cells of mid-exponential phase before and 20 min after the H_2O_2 treatment were harvested for measuring the β-galactosidase assays using an integrative *lacZ* reporter. The activity of β-galactosidase represents the activity of indicated promoters. In all panels, asterisks indicate statistically significant differences of the values compared ($n = 4$; ns, not significant; $^*p < 0.05$; $^{**}p < 0.01$; and $^{***}p < 0.001$).

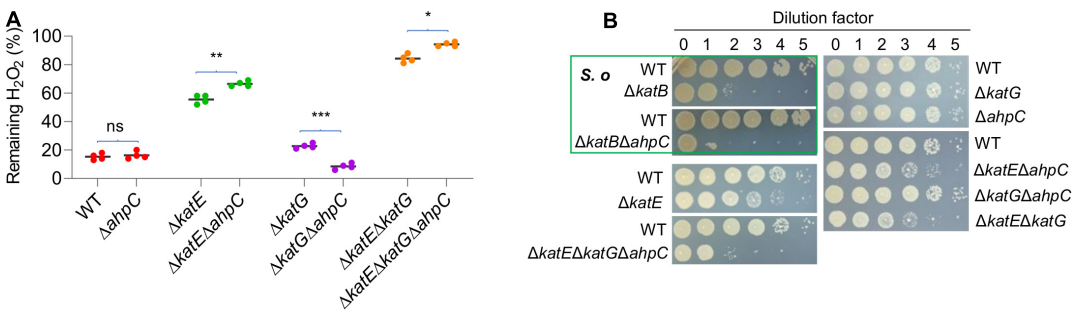


FIGURE 4 | Role of AhpC in decomposition of H_2O_2. **(A)** Measuring the scavenging rate of H_2O_2 in indicated strains. Asterisks indicate statistically significant differences of the values compared ($n = 4$; ns, not significant; $^*p < 0.05$; $^{**}p < 0.01$; and $^{***}p < 0.001$). **(B)** Plating defects of the indicated strains. *S. oneidensis* represents *S. oneidensis*, which is shown for comparison. Strains outside the green box are YPIII. Experiments were performed at least four times, and representative results were presented.

the strain lacking both KatG and AhpC had stronger H_2O_2-decomposing capacity than the wild type. All of these data were verified by genetic complementation, in which one of the missing genes was expressed *in trans* in deletion mutants (**Supplementary Figure 6**).

To provide additional evidence, we then examined if the removal of catalases and/or AhpC would cause YPIII a plating defect, as this is also the case with *S. oneidensis* (Shi et al., 2015). Similar to an *S. oneidensis katB* (encoding dictating catalase) mutant, the Δ*katE* strain of YPIII exhibited a plating defect, albeit much less severe (**Figure 4B**). In contrast, the removal of either *katG* or *ahpC* did not introduce a significant defect, supporting that KatE alone is sufficient to protect cells from the killing of H_2O_2 generated spontaneously on the plates (Shi et al., 2015). Compared to the Δ*katE* strain, the additional removal of either *katG* or *ahpC* only marginally deteriorated the defect (**Figure 4B**). However, the triple mutant (Δ*katE*Δ*katG*Δ*ahpC*) exhibited significantly further lowered viability, resembling the scenario observed with the *S. oneidensis* Δ*katB*Δ*ahpC* strain. Given that *S. oneidensis* AhpC alone is responsible for cleaning endogenous H_2O_2, these data suggest that YPIII would require both KatG and AhpC to be a functional equivalence.

AhpC has been identified as a primary OP scavenger in many bacteria lacking Ohr, such as *E. coli*, *B. subtilis*, and *S.* Typhi (Niimura et al., 1995; Antelmann et al., 1996). We therefore predicted that AhpC is likely involved in combating against OPs in YPIII. To test this, we compared the abilities of the wild type and strains lacking catalase and/or AhpC to scavenge *t*-BHP. Results revealed that the removal of any of the genes under test (*katE*, *katG*, and *ahpC*) or all together did not introduce a significant difference in YPIII susceptibility to *t*-BHP (**Supplementary Figure 7A**). Quantification of the remaining *t*-BHP in the reaction also showed that the *t*-BHP-scavenging capacity of YPIII was not affected by any of these mutations or combined (**Supplementary Figure 7B**). In summary, these data suggest that AhpC in YPIII is able to decompose H_2O_2 but not OPs.

Loss of AhpC and KatG Together Activates OxyR

The data presented thus far have concluded that in YPIII, KatE is the primary catalase in combating H_2O_2-induced oxidative stress, whereas both KatG and AhpC contribute. It has been established that catalase functions to scavenge high concentrations of H_2O_2, whereas AhpC is a more kinetically efficient scavenger of trace H_2O_2 (Seaver and Imlay, 2001). To determine whether the YPIII H_2O_2-scavenging enzymes act in a similar manner, we measured the rates at which cells decomposed low (2 μM) and high (150 μM) concentrations of H_2O_2. The result showed that the *katE* mutant scavenged 2 μM H_2O_2 significantly more rapidly than Δ*katG* and Δ*ahpC* strains did (**Figure 5A**), indicating that in wild type, both of KatG and AhpC carry out the decomposition of low-dose H_2O_2 mostly. Interestingly, a Δ*katG*Δ*ahpC* strain exhibited a faster rate in reducing 2 μM H_2O_2 than strains without either of these two proteins. Given that the additional removal of KatE totally disabled the Δ*katG*Δ*ahpC* strain, this observation implies that the loss of KatG and AhpC induces the expression of KatE. On the contrary, with 150 μM H_2O_2, the *katE* mutant worked very poorly (**Figure 5B**), an expected result from a strain without the primary catalase. While both Δ*katG* and Δ*ahpC* strains showed normal scavenging activity, the Δ*katG*Δ*ahpC* strain decomposed 150 μM H_2O_2 nearly two times faster than did the wild type, supporting an increased production of KatE.

In *E. coli*, the removal of AhpC results in H_2O_2 accumulation, leading to activation of OxyR and thereby increased catalase production (Seaver and Imlay, 2001). Although this is apparently not the case in YPIII, the data presented above imply that the loss of both KatG and AhpC is likely able to stimulate the production of KatE. To test this, we assessed the effects of depletion of KatG, AhpC, and both on KatE expression. As shown in **Figure 5C**, the *katE* gene was expressed indistinguishably in the wild-type, Δ*ahpC*, and Δ*katG* strains, and the same scenario was observed in H_2O_2-treated cells. These observations indicated that the loss of either AhpC or KatG alone does not significantly influence KatE production, eliminating the possibility that neither its absence is sufficient to activate OxyR in YPIII. Conversely, in the Δ*katG*Δ*ahpC* cells grown under normal aerobic conditions, we found that the *katE* promoter activity increased 3.7 times (**Figure 5C**). Despite this, activity of the *katE* promoter can be elevated further under the treatment of H_2O_2 (**Figure 5C**), implying that the oxidative stress risen from the loss of KatG and AhpC probably activates only a portion of OxyR molecules, which is consistent with our earlier proposal that the ratio between oxidized and reduced OxyR is in a dynamic equilibrium (Wan et al., 2018, Wan et al., 2019).

DISCUSSION

Resistance to oxidative stress resulting from ubiquitous ROS generated both endogenously and exogenously belongs to key virulent factors of bacterial pathogens. To deal with H_2O_2-induced oxidative stress, Gram-negative bacteria widely employ OxyR to sense the oxidant and mediate transcription of their regulon in a concerted manner (Imlay, 2013). Although regarded as a pleiotropic regulator, OxyR has evolved to consistently possess operons that are involved in H_2O_2 decomposition and damage control as the core member of regulons (Imlay, 2015). This is seemingly the case in *Y. pseudotuberculosis* YPIII. Our result in **Supplementary Table 2** predicted that all H_2O_2-scavenging enzymes, iron-sequestering proteins, and thioredoxin and glutathione antioxidant systems are encoded by operons whose promoter regions contain most conserved OxyR-binding motifs with the highest similarities. OxyR of YPIII, highly similar to its *E. coli* counterpart in terms of sequence identity (88%) and the almost identical binding motifs, exhibits stimulating activity only (Zheng et al., 1998). This gains support from our transcriptomics data, which reveal elevated transcription of all predicted OxyR regulon members in the H_2O_2-treated cells.

In addition to OxyR regulon members, genes for sulfur species transport and metabolism and iron–manganese transport system are among the top up-regulated in the H_2O_2-treated cells. Sulfur species, especially hydrogen sulfide (H_2S) and cysteine, are well recognized as important factors in bacterial oxidative stress response. These reductants have been shown to play a critical role in the detoxification of H_2O_2 in the periplasm (Ohtsu et al., 2010; Ezraty et al., 2017). Intracellular sulfur homeostasis should be carefully maintained because H_2S and cysteine in vast excess promote oxidative damages by inhibiting catalases (Park and Imlay, 2003; Wu et al., 2015). Consistently, all components of thiol-based antioxidant systems of YPIII are most highly induced upon exposure to H_2O_2, conceivably requiring an active metabolism and fast shuttling of sulfur species (Feng et al., 2019). In parallel, genes for iron/manganese transport system are highly induced by exogenous H_2O_2. In bacteria, many enzymes using iron as a cofactor may become inactive due to the loss of the metal upon oxidative stress (Anjem et al., 2009; Anjem and Imlay, 2012). A common way to overcome this is to replace iron with manganese, and hence more manganese is imported when cells are challenged by H_2O_2 (Imlay et al., 2019). Although further investigations are needed, all of these observations suggest that sulfur species and manganese are important for YPIII to combat oxidative stress.

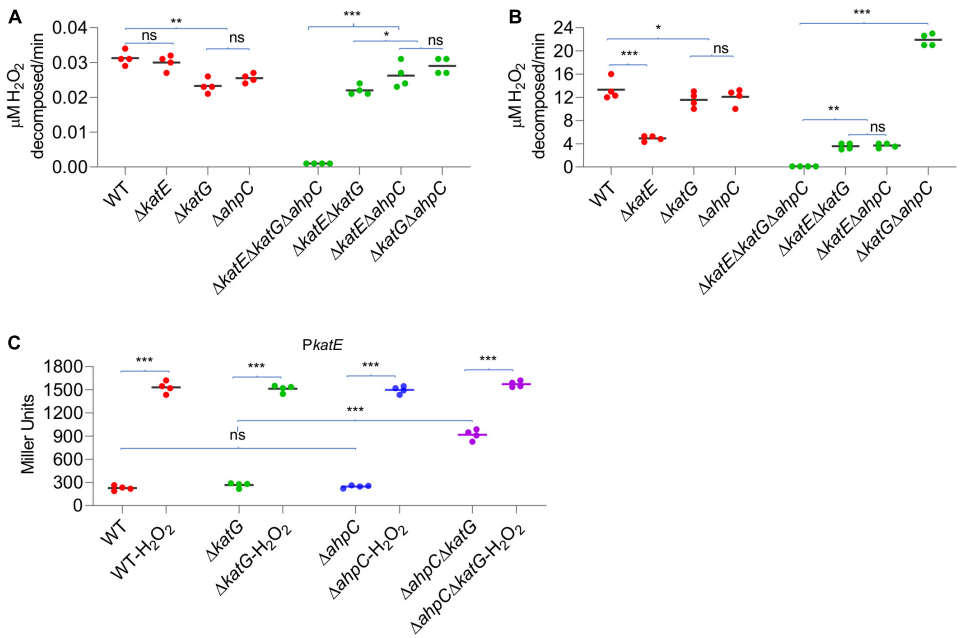


FIGURE 5 | Simultaneous loss of AhpC and KatG increases H_2O_2 resistance. Efficiencies of AhpC, KatE, and KatG at different H_2O_2 concentrations. H_2O_2 was added at a final concentration of 2 μM **(A)** and 150 μM **(B)** to cultures of YPIII strains indicated. 2 min after addition of H_2O_2, the H_2O_2 concentration was measured. **(C)** KatE expression in strains indicated before and after the H_2O_2 treatment. In all panels, asterisks indicate statistically significant differences of the values compared ($n = 4$; ns, not significant; *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$).

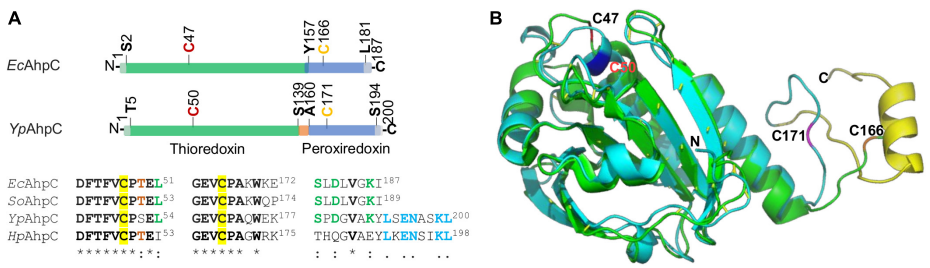


FIGURE 6 | Comparative analysis of AhpC proteins. **(A)** The domain structure of *Ec*AhpC and *Yp*AhpC. Both proteins are composed of two domains, thioredoxin (green) and peroxiredoxin (blue), with peroxidatic Cys47 (or Cys50) and resolving Cys166 (or Cys171) in red and yellow, respectively. Sequence alignment of four representative AhpCs shown below demonstrates the conservation of regions covering two active Cys residues (identical residues are in bold with star mark) and the *C*-terminal tail region. Conserved residues in three out of four AhpCs are shown in red and green; the identical residues in *Yp*AhpC and *Hp*AhpC are shown in blue. **(B)** Structural comparison of *Yp*AhpC and *Ec*AhpC (PDB accession number 4o5r). Shown is a superimposition of *Yp*AhpC (cyan) and *Ec*AhpC (green). The α helix formed by the *C*-terminal tail of *Yp*AhpC is shown in yellow. Active Cys residues are labeled.

Like *E. coli*, YPIII contains two catalases: HPII KatE and HPI KatG (Mishra and Imlay, 2012). Although both KatE and KatG are highly induced upon H_2O_2 treatment, we first anticipate that YPIII uses KatG as the predominant H_2O_2-scavenging enzyme because *E. coli* does so (Seaver and Imlay, 2001). Unexpectedly, it is KatE that confers the H_2O_2-decomposing ability at high concentrations in YPIII. Instead, KatG exhibits a much minor but still significant contribution in scavenging high-dose H_2O_2. More importantly, KatG shows a strong scavenging activity toward low concentrations of H_2O_2, a role that is exclusively played by AhpC in *E. coli*. Despite this, given that *Yp*OxyR becomes activated in the absence of both AhpC and KatG, but not of either one, it is clear that both KatG and AhpC, while neither is sufficient, is able to scavenge low concentrations of H_2O_2 in YPIII. It is well established that AhpC allows *E. coli* cells to degrade low concentrations of H_2O_2 because of its high kinetic efficiency (k_{cat}/K_M, $\sim 10^{-8}$ M^{-1} s^{-1}; Mishra and Imlay, 2012). We believe that this is also likely true in the case of YPIII KatG. But the result of impaired H_2O_2-scavenging ability of Δ*katG* and Δ*ahpC* at low levels in YPIII implies that YPIII KatG possesses a kinetic efficiency similar to, or at least not significantly lower than, that of AhpC, the catalytic activity that is substantially greater than those of KatE and KatG in *E. coli* whose activities are similar to each other ($\sim 10^{-6}$ M^{-1} s^{-1}; Obinger et al., 1997; Hillar et al., 2000). We are working to test this notion.

Bacterial AhpC is regarded as the founding member of the peroxiredoxin family, and its homologs appear to be more widely distributed than catalase/peroxidase (Chae et al., 1994). The AhpR complex is highly conserved among diverse bacteria, with respect to not only amino acid sequence but also gene organization (Feng et al., 2020). To date, all bacteria that employ OxyR as the master H_2O_2-responding regulator are equipped with an *ahp* operon encoding both AhpC and its cognate reductase AhpF, including *M. tuberculosis* in which AhpF is replaced by AhpD to reduce AhpC (Baker et al., 2001; Jaeger et al., 2004). Apart from that, AhpC without a co-transcribed cognate reductase is found in *H. pylori*, a gastric pathogen that represents an extreme example in terms of oxidative stress response machinery because it lacks homologs of the oxidative stress regulators present in other bacteria, including OxyR, SoxR, SoxS, and PerR (Baker et al., 2001; Wang et al., 2006). A sequence alignment reveals that AhpC of *Y. pseudotuberculosis* is more closely related to typical AhpCs than to the *H. pylori* AhpC: there are significantly more identical residues in AhpCs of *E. coli*, *S. oneidensis*, and *Y. pseudotuberculosis* (**Supplementary Figure 8**). Despite this, we also observed that the linear arrangement of AhpC gene is highly conserved in *Yersinia* species (**Supplementary Figure 9**), a phenomenon that is quite rare in other bacteria. Therefore, it would be interesting to demystify how evolution has honed this particular phenomenon.

Although the reductase for YPIII AhpC remains to be determined, we may get clues from *H. pylori* AhpC in terms of their differences in sequence and structure from the typical ones (with AhpF as the cognate reductase). Reduction of *H. pylori* AhpC is carried out by thioredoxin TrxA, which is reduced by thioredoxin reductase TrxR (also called TrxB) using NADPH as the electron donor (Wang et al., 2006). A common feature that *Yp*AhpC and *Hp*AhpC share is that they have an additional eight-residue segment in the *C*-terminal end, which is missing in the typical AhpC proteins (**Figure 6A** and **Supplementary Figure 8**). In fact, the last few residues in the *C* terminus of AhpC have been shown to be particularly critical in maintaining an enzymatically active AhpR complex *in vitro* (Dip et al., 2014a). This *C*-terminal tail of AhpC is crucial in complexation of AhpC with AhpF, TrxR, or glutathione-disulfide reductase GorA (Dip et al., 2014a; Feng et al., 2020). Coincidently, earlier studies of certain *E. coli* and *S. oneidensis* AhpC mutants have established that certain mutations, especially in the *C*-terminal region, can transform AhpC to a surprisingly malleable enzyme, reducible not only by Trx system but also by Grx system (Faulkner et al., 2008; Feng et al., 2020). Therefore, we speculate that the extra eight residues at the *C*-terminal end of YPIII AhpC may affect the specificity of the protein for its reducing partner.

The difference in the *C*-terminal region may also provide an explanation for the finding that YPIII AhpC does not have the ability to reduce OPs. It has been shown that catalytic efficiency of AhpC with small hydroperoxides, especially H_2O_2, is substantially higher (∼100-fold) than that with bulky, tertiary hydroperoxides, such as *t*-BHP (Parsonage et al., 2008). For both types of oxidants, the catalytic reaction adopts a binding and releasing mechanism that enables the assembly of AhpC and the reducing partner to undergo efficient catalytic cycles of transferring electrons without compromising the catalytic turnover rate (Parsonage et al., 2015). A prerequisite for such a mechanism is the correct fitting of the *C*-terminal tail of AhpC into the "arch-like" groove on the binding surface of the reducing partner, resulting in formation of an active-site pocket, to which small peroxides are more accessible (Dip et al., 2014b). The additional *C*-terminal tail of YPIII AhpC may completely prevent large OPs from entering the active-site pocket (**Figure 6B**). Compared to the *E. coli* AhpC structure (PDB accession number 4o5r; Dip et al., 2014b), *Y. pseudotuberculosis* AhpC has an α-helix that is formed by the extra residues. Unfortunately, how this α-helix folds remains unknown because the available *H. pylori* AhpC structure is truncated without this tail (PDB accession number 1zof; Papinutto et al., 2005; **Supplementary Figure 10**). But one may imagine that the extra *C*-terminus α-helix may hinder the reduction of OPs. Given that *Yp*AhpC, unlike typical AhpCs, is not sufficient to scavenge endogenous H_2O_2, we speculate that the α-helix also compromises the efficiency of H_2O_2 reduction. If this holds, the difference in the kinetic efficiencies for H_2O_2 between AhpC and catalases of YPIII is likely significantly smaller than that between the *E. coli* counterparts.

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 in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

FW and HG designed and supported the research. FW, XF, and JY performed the research and analyzed the data. FW, XF, and HG wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (31900021 to FW, 31600041 to JY, and 31930003 and 41976087 to HG).

ACKNOWLEDGMENTS

We thank Yao Wang for providing us with the *Y. pseudotuberculosis* YPIII wild type.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.626874/full#supplementary-material

REFERENCES

Anjem, A., and Imlay, J. A. (2012). Mononuclear iron enzymes are primary targets of hydrogen peroxide stress. *J. Biol. Chem.* 287, 15544–15556. doi: 10.1074/jbc.m111.330365

Anjem, A., Varghese, S., and Imlay, J. A. (2009). Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Mol. Microbiol.* 72, 844–858. doi: 10.1111/j.1365-2958.2009.06699.x

Antelmann, H., Engelmann, S., Schmid, R., and Hecker, M. (1996). General and oxidative stress responses in *Bacillus subtilis*: cloning, expression, and mutation of the alkyl hydroperoxide reductase operon. *J. Bacteriol.* 178, 6571–6578. doi: 10.1128/jb.178.22.6571-6578.1996

Baker, L. M., Raudonikiene, A., Hoffman, P. S., and Poole, L. B. (2001). Essential thioredoxin-dependent peroxiredoxin system from *Helicobacter* pylori: genetic and kinetic characterization. *J. Bacteriol.* 183, 1961–1973. doi: 10.1128/jb.183.6.1961-1973.2001

Brady, M. F., Yarrarapu, S. N. S., and Anjum, F. (2020). "Yersinia pseudotuberculosis," in *StatPearls*, (Treasure Island (FL): StatPearls Publishing). Available online at: https://www.ncbi.nlm.nih.gov/books/NBK430717/

Bryk, R., Griffin, P., and Nathan, C. (2000). Peroxynitrite reductase activity of bacterial peroxiredoxins. *Nature* 407, 211–215. doi: 10.1038/35025109

Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G., and Rhee, S. G. (1994). Cloning and sequencing of thiol-specific antioxidant from mammalian brain: alkyl hydroperoxide reductase and thiol-specific antioxidant define a large family of antioxidant enzymes. *Proc. Natl. Acad. Sci. U S A.* 91, 7017–7021. doi: 10.1073/pnas.91.15.7017

Chandrangsu, P., Rensing, C., and Helmann, J. D. (2017). Metal homeostasis and resistance in bacteria. *Nat. Rev. Microbiol.* 15, 338–350. doi: 10.1038/nrmicro.2017.15

Cosgrove, K., Coutts, G., Jonsson, I.-M., Tarkowski, A., Kokai-Kun, J. F., Mond, J. J., et al. (2007). Catalase (KatA) and alkyl hydroperoxide reductase (AhpC) have compensatory roles in peroxide stress resistance and are required for survival, persistence, and nasal colonization in *Staphylococcus aureus*. *J. Bacteriol.* 189, 1025–1035. doi: 10.1128/jb.01524-06

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Dip, P. V., Kamariah, N., Nartey, W., Beushausen, C., Kostyuchenko, V. A., Ng, T.-S., et al. (2014a). Key roles of the *Escherichia coli* AhpC C-terminus in assembly and catalysis of alkylhydroperoxide reductase, an enzyme essential for the alleviation of oxidative stress. *Biochim. Biophys. Acta* 1837, 1932–1943. doi: 10.1016/j.bbabio.2014.08.007

Dip, P. V., Kamariah, N., Subramanian Manimekalai, M. S., Nartey, W., Balakrishna, A. M., Eisenhaber, F., et al. (2014b). Structure, mechanism and ensemble formation of the alkylhydroperoxide reductase subunits AhpC and AhpF from *Escherichia coli*. *Acta Crystallogr D* 70, 2848–2862. doi: 10.1107/s1399004714019233

Ezraty, B., Gennaris, A., Barras, F., and Collet, J. F. (2017). Oxidative stress, protein damage and repair in bacteria. *Nat. Rev. Microbiol.* 15, 385–396. doi: 10.1038/nrmicro.2017.26

Fang, F. C. (2004). Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nat. Rev. Microbiol.* 2, 820–832. doi: 10.1038/nrmicro1004

Faulkner, M. J., Veeravalli, K., Gon, S., Georgiou, G., and Beckwith, J. (2008). Functional plasticity of a peroxidase allows evolution of diverse disulfide-reducing pathways. *Proc. Natl. Acad. Sci. U S A.* 105, 6735–6740. doi: 10.1073/pnas.0801986105

Feng, X., Guo, K., and Gao, H. (2020). Plasticity of the peroxidase AhpC links multiple substrates to diverse disulfide-reducing pathways in Shewanella oneidensis. *J. Biol. Chem.* 295, 11118–11130. doi: 10.1074/jbc.ra120.014010

Feng, X., Sun, W., Kong, L., and Gao, H. (2019). Distinct roles of Shewanella oneidensis thioredoxin in regulation of cellular responses to hydrogen and organic peroxides. *Appl. Environ. Microbiol.* 85, e01700–e01719.

Fu, H., Jin, M., Ju, L., Mao, Y., and Gao, H. (2014). Evidence for function overlapping of CymA and the cytochrome bc1 complex in the *Shewanella oneidensis* nitrate and nitrite respiration. *Environ. Microbiol.* 16, 3181–3195. doi: 10.1111/1462-2920.12457

Fu, H., Yuan, J., and Gao, H. (2015). Microbial oxidative stress response: novel insights from environmental facultative anaerobic bacteria. *Arch. Biochem. Biophys.* 584, 28–35. doi: 10.1016/j.abb.2015.08.012

Gao, H., Wang, Y., Liu, X., Yan, T., Wu, L., Alm, E., et al. (2004). Global transcriptome analysis of the heat shock response of Shewanella oneidensis. *J. Bacteriol.* 186, 7796–7803. doi: 10.1128/jb.186.22.7796-7803.2004

Hillar, A., Peters, B., Pauls, R., Loboda, A., Zhang, H., Mauk, A. G., et al. (2000). Modulation of the activities of catalase-peroxidase HPI of *Escherichia coli* by site-directed mutagenesis. *Biochemistry* 39, 5868–5875. doi: 10.1021/bi0000059

Ieva, R., Roncarati, D., Metruccio, M. M. E., Seib, K. L., Scarlato, V., and Delany, I. (2008). OxyR tightly regulates catalase expression in *Neisseria meningitidis* through both repression and activation mechanisms. *Mol. Microbiol.* 70, 1152–1165. doi: 10.1111/j.1365-2958.2008.06468.x

Imlay, J. A. (2008). Cellular defenses against superoxide and hydrogen peroxide. *Annu. Rev. Biochem.* 77, 755–776. doi: 10.1146/annurev.biochem.77.061606.161055

Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* 11, 443–454. doi: 10.1038/nrmicro3032

Imlay, J. A. (2015). Transcription factors that defend bacteria against reactive oxygen species. *Annu. Rev. Microbiol.* 69, 93–108. doi: 10.1146/annurev-micro-091014-104322

Imlay, J. A., Sethu, R., and Rohaun, S. K. (2019). Evolutionary adaptations that enable enzymes to tolerate oxidative stress. *Free Radical. Biol. Med.* 140, 4–13. doi: 10.1016/j.freeradbiomed.2019.01.048

Jacobson, F. S., Morgan, R., Christman, M., and Ames, B. (1989). An alkyl hydroperoxide reductase from *Salmonella typhimurium* involved in the defense of DNA against oxidative damage. purification and properties. *J. Biol. Chem.* 264, 1488–1496. doi: 10.1016/s0021-9258(18)94214-6

Jaeger, T., Budde, H., Flohé, L., Menge, U., Singh, M., Trujillo, M., et al. (2004). Multiple thioredoxin-mediated routes to detoxify hydroperoxides in *Mycobacterium tuberculosis*. *Arch. Biochem. Biophys.* 423, 182–191. doi: 10.1016/j.abb.2003.11.021

Jiang, Y., Dong, Y., Luo, Q., Li, N., Wu, G., and Gao, H. (2014). Protection from oxidative stress relies mainly on derepression of OxyR-dependent KatB and Dps in *Shewanella oneidensis*. *J. Bacteriol.* 196, 445–458. doi: 10.1128/jb.01077-13

Jin, M., Jiang, Y., Sun, L., Yin, J., Fu, H., Wu, G., et al. (2013). Unique organizational and functional features of the cytochrome c maturation system in *Shewanella oneidensis*. *PLoS One* 8:e75610. doi: 10.1371/journal.pone.0075610

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocols* 10, 845–858. doi: 10.1038/nprot.2015.053

Kim, J.-S., and Holmes, R. (2012). Characterization of OxyR as a negative transcriptional regulator that represses catalase production in *Corynebacterium diphtheriae*. *PLoS One* 7:e31709. doi: 10.1371/journal.pone.0031709

Knaus, U. G., Hertzberger, R., Pircalabioru, G. G., Yousefi, S. P. M., Branco, and dos Santos, F. (2017). Pathogen control at the intestinal mucosa – H2O2 to the rescue. *Gut Microbes* 8, 67–74. doi: 10.1080/19490976.2017.1279378

Korshunov, S., Imlay, K. R. C., and Imlay, J. A. (2020). Cystine import is a valuable but risky process whose hazards *Escherichia coli* minimizes by inducing a cysteine exporter. *Mol. Microbiol.* 113, 22–39. doi: 10.1111/mmi.14403

Li, R., Jia, Z., and Trush, M. A. (2016). Defining ROS in biology and medicine. *React. Oxygen Species* 1, 9–21.

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641.

Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J. A., Delerce, J., et al. (2015). RSAT 2015: regulatory sequence analysis tools. *Nucleic Acids Res.* 43, W50–W56.

Mishra, S., and Imlay, J. (2012). Why do bacteria use so many enzymes to scavenge hydrogen peroxide? *Arch Biochem. Biophys.* 525, 145–160. doi: 10.1016/j.abb.2012.04.014

Niimura, Y., Poole, L. B., and Massey, V. (1995). Amphibacillus xylanus NADH oxidase and *Salmonella typhimurium* alkyl-hydroperoxide reductase flavoprotein components show extremely high scavenging activity for both alkyl hydroperoxide and hydrogen peroxide in the presence of *S. typhimurium* alkyl-hydroperoxide reductase 22-kDa protein component. *J. Biol. Chem.* 270, 25645–25650. doi: 10.1074/jbc.270.43.25645

Obinger, C., Maj, M., Nicholls, P., and Loewen, P. (1997). Activity, peroxide compound formation, and heme d synthesis in *Escherichia coli* HPII catalase. *Arch. Biochem. Biophys.* 342, 58–67. doi: 10.1006/abbi.1997.9988

Ohtsu, I., Wiriyathanawudhiwong, N., Morigasaki, S., Nakatani, T., Kadokura, H., and Takagi, H. (2010). The L-cysteine/L-cystine shuttle system provides reducing equivalents to the periplasm in *Escherichia coli*. *J. Biol. Chem.* 285, 17479–17487. doi: 10.1074/jbc.m109.081356

Papinutto, E., Windle, H. J., Cendron, L., Battistutta, R., Kelleher, D., and Zanotti, G. (2005). Crystal structure of alkyl hydroperoxide-reductase (AhpC) from *Helicobacter pylori*. *Biochim. Biophys. Acta* 1753, 240–246. doi: 10.1016/j.bbapap.2005.09.001

Park, S., and Imlay, J. A. (2003). High levels of intracellular cysteine promote oxidative DNA damage by driving the fenton reaction. *J. Bacteriol.* 185, 1942–1950. doi: 10.1128/jb.185.6.1942-1950.2003

Parsonage, D., Karplus, P. A., and Poole, L. B. (2008). Substrate specificity and redox potential of AhpC, a bacterial peroxiredoxin. *Proc. Natl. Acad. Sci. U S A.* 105, 8209–8214. doi: 10.1073/pnas.0708308105

Parsonage, D., Nelson, K. J., Ferrer-Sueta, G., Alley, S., Karplus, P. A., Furdui, C. M., et al. (2015). Dissecting peroxiredoxin catalysis: separating binding, peroxidation, and resolution for a bacterial AhpC. *Biochemistry* 54, 1567–1575. doi: 10.1021/bi501515w

Peng, H., Zhang, Y., Palmer, L. D., Kehl-Fie, T. E., Skaar, E. P., Trinidad, J. C., et al. (2017). Hydrogen sulfide and reactive sulfur species impact proteome s-sulfhydration and global virulence regulation in *Staphylococcus aureus*. *ACS Infect Dis.* 3, 744–755. doi: 10.1021/acsinfecdis.7b00090

Perkins, A., Nelson, K. J., Parsonage, D., Poole, L. B., and Karplus, P. A. (2015). Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends Biochem. Sci.* 40, 435–445. doi: 10.1016/j.tibs.2015.05.001

Poole, L. B. (2005). Bacterial defenses against oxidants: mechanistic features of cysteine-based peroxidases and their flavoprotein reductases. *Arch. Biochem. Biophys.* 433, 240–254. doi: 10.1016/j.abb.2004.09.006

Poole, L. B., and Ellis, H. R. (1996). Flavin-dependent alkyl hydroperoxide reductase from *Salmonella typhimurium*. purification and enzymatic activities of overexpressed AhpF and AhpC proteins. *Biochemistry* 35, 56–64. doi: 10.1021/bi951887s

Rosqvist, R., Skurnik, M., and Wolf-Watz, H. (1988). Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations. *Nature* 334, 522–525. doi: 10.1038/334522a0

Seaver, L. C., and Imlay, J. A. (2001). Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J. Bacteriol.* 183, 7173–7181. doi: 10.1128/jb.183.24.7173-7181.2001

Shi, M., Wan, F., Mao, Y., and Gao, H. (2015). Unraveling the mechanism for the viability deficiency of *Shewanella oneidensis* oxyR null mutant. *J. Bacteriol.* 197, 2179–2189. doi: 10.1128/jb.00154-15

Sun, Y., Meng, Q., Zhang, Y., and Gao, H. (2020). Derepression of bkd by the FadR loss dictates elevated production of BCFAs and isoleucine starvation. *Biochim. Biophys. Acta* 1865:158577. doi: 10.1016/j.bbalip.2019.158577

Tartaglia, L. A., Storz, G., Brodsky, M. H., Lai, A., and Ames, B. N. (1990). Alkyl hydroperoxide reductase from *Salmonella typhimurium*. sequence and homology to thioredoxin reductase and other flavoprotein disulfide oxidoreductases. *J. Biol. Chem.* 265, 10535–10540. doi: 10.1016/s0021-9258(18)86980-0

Wan, F., Kong, L., and Gao, H. (2018). Defining the binding determinants of *Shewanella oneidensis* OxyR: implications for the link between the contracted OxyR regulon and adaptation. *J. Biol. Chem.* 293, 4085–4096. doi: 10.1074/jbc.ra117.001530

Wan, F., Yin, J., Sun, W., and Gao, H. (2019). Oxidized OxyR up-regulates *ahpCF*expression to suppress plating defects of *oxyR*- and catalase-deficient strains. *Front. Microbiol.* 10:439. doi: 10.3389/fmicb.2019.00439

Wang, G., Alamuri, P., and Maier, R. J. (2006). The diverse antioxidant systems of *Helicobacter pylori*. *Mol. Microbiol.* 61, 847–860. doi: 10.1111/j.1365-2958.2006.05302.x

Wang, T., Yang, X., Gao, F., Zhao, C., Kang, Y., Wang, Y., et al. (2016). Zinc acquisition via ZnuABC in *Yersinia pseudotuberculosis* facilitates resistance to oxidative stress. *Ann. Microbiol.* 66, 1189–1197. doi: 10.1007/s13213-016-1205-7

Wang, Z., Wang, T., Cui, R., Zhang, Z., Chen, K., Li, M., et al. (2020). HpaR, the repressor of aromatic compound metabolism, positively regulates the expression of T6SS4 to resist oxidative stress in *Yersinia pseudotuberculosis*. *Front. Microbiol.* 11:705. doi: 10.3389/fmicb.2020.00705

Wolff, S. P. (1994). Ferrous ion oxidation in presence of ferric ion indicator xylenol orange for measurement of hydroperoxides. *Methods Enzymol.* 233, 182–189. doi: 10.1016/s0076-6879(94)33021-2

Wu, G., Wan, F., Fu, H., Li, N., and Gao, H. (2015). A matter of timing: contrasting effects of hydrogen sulfide on oxidative stress response in *Shewanella oneidensis*. *J. Bacteriol.* 197, 3563–3572. doi: 10.1128/jb.00603-15

Wu, L., Wang, J., Tang, P., Haijiang, C., and Gao, H. (2011). Genetic and molecular characterization of flagellar assembly in *Shewanella oneidensis*. *PLoS One* 6:e21479. doi: 10.1371/journal.pone.0021479

Zheng, M., Åslund, F., and Storz, G. (1998). Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* 279, 1718–1722. doi: 10.1126/science.279.5357.1718

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: 01 June 2021
doi: 10.3389/fmicb.2021.662317

Transcriptional Activity of the Bacterial Replication Initiator DnaA

*Inoka P. Menikpurage, Kristin Woo and Paola E. Mera**

Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL, United States

In bacteria, DnaA is the most conserved DNA replication initiator protein. DnaA is a DNA binding protein that is part of the AAA+ ATPase family. In addition to initiating chromosome replication, DnaA can also function as a transcription factor either as an activator or repressor. The first gene identified to be regulated by DnaA at the transcriptional levels was *dnaA*. DnaA has been shown to regulate genes involved in a variety of cellular events including those that trigger sporulation, DNA repair, and cell cycle regulation. DnaA's dual functions (replication initiator and transcription factor) is a potential mechanism for DnaA to temporally coordinate diverse cellular events with the onset of chromosome replication. This strategy of using chromosome replication initiator proteins as regulators of gene expression has also been observed in archaea and eukaryotes. In this mini review, we focus on our current understanding of DnaA's transcriptional activity in various bacterial species.

Keywords: DnaA, replication initiation, transcription factor, cell cycle, sporulation, DNA repair, autoregulation

OPEN ACCESS

Edited by:
Jianping Xie,
Southwest University, China

Reviewed by:
Justine Collier,
University of Lausanne, Switzerland
Rodrigo Reyes,
McGill University, Canada
Julie Anne Maupin-Furlow,
University of Florida, United States

***Correspondence:**
Paola E. Mera
pmera@illinois.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *01 February 2021*
Accepted: *19 April 2021*
Published: *01 June 2021*

Citation:
Menikpurage IP, Woo K and Mera PE (2021) Transcriptional Activity of the Bacterial Replication Initiator DnaA. Front. Microbiol. 12:662317. doi: 10.3389/fmicb.2021.662317

INTRODUCTION

DnaA is a multifunctional protein that can serve as a master regulator in bacteria. DnaA is composed of four structural domains with some species-specific variations: (I) protein–protein interaction domain, (II) linker domain, (III) AAA+ ATPase domain, and (IV) DNA binding (helix–turn–helix) domain (Fujikawa et al., 2003; Duderstadt and Berger, 2013) (**Figure 1A**). The two functions of DnaA that are best understood are as an initiator of chromosome replication and as a transcription factor. DnaA and/or the onset of chromosome replication have been linked to cell size regulation and chromosome segregation; however, details about these links remain unclear (Lobner-Olesen et al., 1989; Hill et al., 2012; Mera et al., 2014). As a replication initiator, DnaA opens the origin of replication (*ori*) by binding at specific DNA sequences referred to as DnaA boxes (Fuller and Kornberg, 1983; Bramhill and Kornberg, 1988). This function has been extensively studied, and mechanistic questions continue to be articulated (recent reviews; Hansen and Atlung, 2018; Frandi and Collier, 2019; Leonard et al., 2019; Ozaki, 2019; Kohiyama, 2020). In this mini review, we focus on the transcriptional activity of DnaA and its role in modulating various cellular events (**Figure 1B**). DnaA has been shown to have a global transcriptional impact because it regulates the expression levels of other global regulators involved in cell cycle progression and developmental processes (**Figure 2**). The specific genes found in DnaA's transcriptional regulon vary significantly depending on the bacterial species. To begin, we will discuss the ability of DnaA to autoregulate its levels as a way to provide context for the importance of its transcriptional activity.

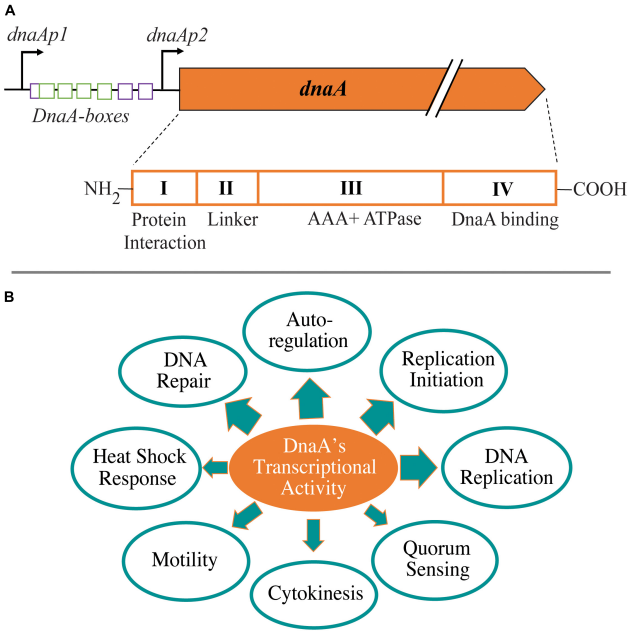


FIGURE 1 | Cellular impact from transcriptional activity of DnaA. **(A)** Schematic of the promoter region of *dnaA* in *Escherichia coli* and conserved domains of the DnaA protein. Within the promoter region, purple boxes depict ATP-specific DnaA boxes and green boxes have no nucleotide-bound DnaA specificity. **(B)** Cellular processes from different bacterial species shown to be influenced by DnaA's transcriptional activity. Thickness of arrow is an approximate representation of the amount of evidence that exists for that particular process.

AUTOREGULATION OF *dnaA* EXPRESSION

The ability of DnaA to autoregulate its transcription was discovered at a time when the direct role of DnaA in chromosome replication was not yet clear. Early characterizations of temperature-sensitive *dnaA* mutants of *Escherichia coli* revealed that the copy number of *ori* decreased, while DnaA levels increased simultaneously at nonpermissive temperatures, thus accurately predicting that DnaA had a positive involvement in the onset of chromosome replication and a negative involvement in regulating its own synthesis (Hanna and Carl, 1975; Hansen and Rasmussen, 1977). DnaA's ability to repress its own transcription was later confirmed using *in vitro* and *in vivo* analyses. *In vitro*, DnaA was shown to directly bind *dnaA*'s promoter region using gel shift assays and DNase footprinting methods (Braun et al., 1985; Wang and Kaguni, 1987). *In vivo*, two strategies were used to show that DnaA represses its own transcription. First, increased cellular levels of DnaA (expressed from inducible promoters) were shown to decrease the activity of the *dnaA* promoter, and second, decreased DnaA levels able to bind the *dnaA* promoter (by adding *dnaA* binding sites on a plasmid that titrated DnaA levels away) were shown to increase the activity of the *dnaA* promoter (Atlung et al., 1985; Braun et al., 1985; Kucherer et al., 1986; Hansen et al., 1987).

In *E. coli*, the expression of *dnaA* is regulated from two promoters: *dnaAp1* and *dnaAp2* ((Hansen E. B. et al., 1982; Hansen F. G. et al., 1982) (**Figure 1A**). DnaA can repress its own expression from both promoters with *dnaAp2* being a threefold stronger promoter than *dnaAp1* (Atlung et al., 1984, 1985; Braun et al., 1985; Kucherer et al., 1986; Chiaramello and Zyskind, 1990). A 9-mer DnaA box with the consensus sequence TTATCCACA was first identified between *dnaAp1* and *dnaAp2* explaining the mechanism of autoregulation (Hansen F. G. et al., 1982). Intriguingly, the elimination of this DnaA box on *dnaA*'s promoter did not fully eliminate the autoregulation of *dnaA* expression, leading to the hypothesis of indirect mechanisms (Polaczek and Wright, 1990; Smith et al., 1997). The mystery was solved by the identification of DnaA-ATP-specific DnaA boxes (AGATCT) that were involved in full repression of *dnaA* expression (Speck et al., 1999). In addition to the autoregulation by DnaA, the transcription of *E. coli dnaA* is regulated at multiple levels and by various other proteins, including DNA methylation, second messenger ppGpp, growth rate, SeqA, Fis, IciA (ArgP), and QseB (Braun and Wright, 1986; Kucherer et al., 1986; Campbell and Kleckner, 1990; Chiaramello and Zyskind, 1990; Polaczek and Wright, 1990; Zyskind and Smith, 1992; Lu et al., 1994; Froelich et al., 1996; Lee et al., 1996; Hansen and Atlung, 2018; Riber and Lobner-Olesen, 2020; Wu et al., 2021).

Besides *E. coli*, autoregulation by DnaA directly binding at its promoter region has been confirmed in *Bacillus subtilis* (Ogura et al., 2001), *Vibrio harveyi* (Berenstein et al., 2002), *Streptomyces lividans* (Zakrzewska-Czerwinska et al., 1994; Jakimowicz et al., 2000), and in slow-growing mycobacteria (Salazar et al., 2003). In *Pseudomonas putida*, 12 DnaA boxes were identified on its *dnaA* promoter region (Fujita et al., 1989). However, when the levels of DnaA from *P. putida* were increased (by inducible expression from plasmid), repression of *dnaA* expression was not observed in *P. putida*. Interestingly, the overexpression of *dnaA* (only the open reading frame) from *E. coli* under the same broad host range plasmid construct in *P. putida* does result in repression of *dnaA*'s transcription (Ingmer and Atlung, 1992). The function, if any, of the 12 DnaA boxes in *P. putida dnaA* promoter remains to be determined. Likewise, in *Caulobacter crescentus*, the role that DnaA plays in its transcriptional regulation remains unclear. Transcription of *dnaA* in *C. crescentus* has been proposed to be regulated by methylation status of the promoter region and by a *cis*-acting element upstream of the -35 promoter region (Collier et al., 2007; Cheng and Keiler, 2009; Felletti et al., 2019; Frandi and Collier, 2019).

NUCLEOTIDE SWITCH REGULATES THE TRANSCRIPTIONAL ACTIVITY OF DnaA

Both activities of DnaA (transcription factor and replication initiator) can be modulated using an ATP-dependent molecular switch. In *E. coli*, DnaA can be found inside the cell tightly bound to ATP (K_d 30 nM) or bound to ADP (K_d 100 nM) (Sekimizu et al., 1987). The protein Hda (homologous to DnaA) and chromosomal loci *datA* promote hydrolysis of DnaA–ATP

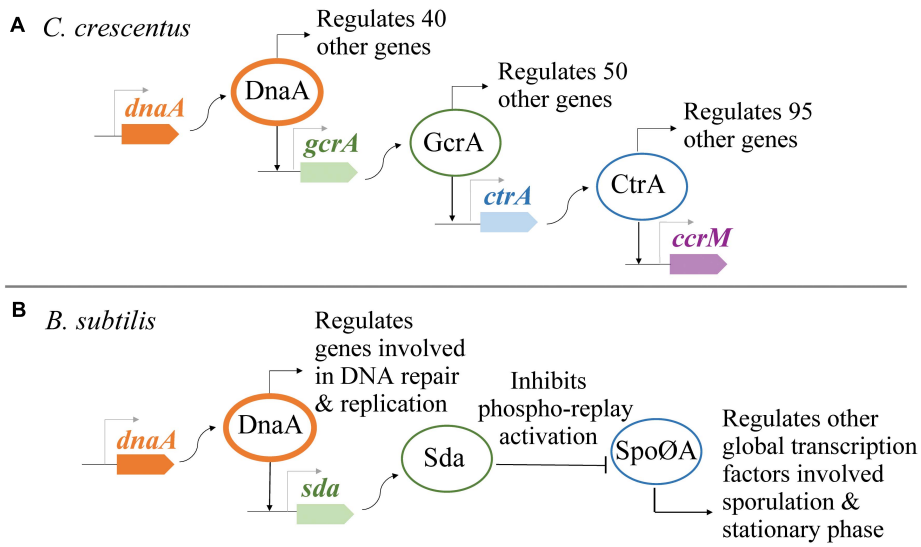


FIGURE 2 | DnaA regulates the transcription of other global regulators. **(A)** In *Caulobacter crescentus*, DnaA is part of a genetic network responsible for regulating the forward progression of the cell cycle. **(B)** In *Bacillus subtilis*, DnaA indirectly inhibits the phosphorylation of Spo0A, a global regulator of sporulation and entrance to stationary phase.

to DnaA–ADP in *E. coli* (Kato and Katayama, 2001; Kasho and Katayama, 2013). The cellular ratio of DnaA–ATP to DnaA–ADP changes as the cell progresses over the cell cycle. Although DnaA–ATP and DnaA–ADP bind *ori*, only DnaA–ATP can open *ori* and initiate replication in *E. coli* (Sekimizu et al., 1987; Mizushima et al., 1996). As a transcription factor, the nucleotide switch has been shown to turn DnaA into a stronger repressor or stronger activator depending on the bound nucleotide. The ability of DnaA's transcriptional activity to be regulated with this switch was first identified in the autoregulation of *dnaA*'s transcription in *E. coli* (Speck et al., 1999). DnaA–ATP is able to fully repress the expression of *dnaA*, whereas DnaA–ADP can only repress up to 40%. By distinguishing the nucleotide bound to DnaA, four new DnaA boxes were identified in the promoter region of *dnaA* with specificity for DnaA–ATP and potential cooperative binding (Speck et al., 1999).

The second example of this switch involves the synthesis of DNA substrates in *E. coli*. Ribonucleotide reductase (RNR) encoded by *nrdAB* catalyzes the reduction of ribonucleotides to deoxyribonucleotides (Thelander and Reichard, 1979). In *E. coli*, DnaA regulates the expression of *nrdAB* by binding the promoter region that contains three DnaA boxes: two boxes do not have specificity for the nucleotide-bound DnaA, and one box is specific for DnaA–ATP (Tuggle and Fuchs, 1986; Augustin et al., 1994; Jacobson and Fuchs, 1998; Olliver et al., 2010). DnaA can activate or repress the expression of *nrdAB* based on the nucleotide bound to DnaA and based on the levels of DnaA. For instance, cells expressing DnaA variants that are deficient for ATP binding or hyperactive for ATPase activity display an increased expression of *nrdAB* (Gon et al., 2006; Babu et al., 2017). Thus, high levels of DnaA-ATP repress the transcription of *nrdAB* presumably by precluding RNA polymerase from binding the promoter region (Olliver et al., 2010). Conversely, DnaA–ADP or low levels of DnaA–ATP bound at the high-affinity DnaA boxes in *nrdAB* promoter region can activate *nrdAB* transcription by stabilizing the RNA polymerase–DNA complex (Augustin et al., 1994; Olliver et al., 2010). Although the expression of RNR coincides with the onset of chromosome replication (Sun and Fuchs, 1992), DnaA seems to only regulate the expression levels and not necessarily the timing of *nrdAB* expression (Sun et al., 1994; Olliver et al., 2010). Interestingly, the expression of *dnaA* and *nrdAB* in *E. coli* share two other regulators besides DnaA: Fis and IciA (Augustin et al., 1994; Froelich et al., 1996; Lee et al., 1996; Han et al., 1998). In *B. subtilis* and *C. crescentus*, DnaA has also been proposed to act as a transcriptional regulator of genes encoding RNR (Goranov et al., 2005; Hottes et al., 2005), although no detailed analyses has been performed to determine if DnaA's nucleotide switch is also involved.

In *C. crescentus*, DnaA-ADP (not DnaA-ATP) has been proposed to activate the transcription of a set of three essential genes encoding proteins involved in cell cycle regulation: FtsZ (tubulin-like protein essential for cell division), MipZ (inhibitor of FtsZ polymerization), and GcrA (global transcription factor) (Fernandez-Fernandez et al., 2011). The expression of the hyperactive replication initiator variant DnaA^{R357A} resulted in *C. crescentus* cells over-initiating replication (Fernandez-Fernandez et al., 2011; Wargachuk and Marczynski, 2015). The corresponding variant in *E. coli* (DnaA^{R334A}) was shown to bind ATP but was unable to hydrolyze ATP (Nishida et al., 2002). In *C. crescentus*, DnaA^{R357A} variant was shown to lose its ability to activate the transcription of *ftsZ*, *mipZ*, and *gcrA* (Hottes et al., 2005; Fernandez-Fernandez et al., 2011). In *B. subtilis*, most chromosomal regions that bind DnaA displayed higher specificity for DnaA-ATP over DnaA-ADP *in vitro* (Smith and Grossman, 2015), suggesting a potential wide usage of DnaA's nucleotide switch to modulate the transcriptional activity of

DnaA. However, the mechanism(s) that regulate differences in DNA binding specificity between DnaA–ATP and DnaA–ADP remain unclear.

DnaA'S GLOBAL TRANSCRIPTIONAL REGULATION

Identifying the complete transcriptional regulon of DnaA is complex due to the essential function of DnaA as a replication initiator and the role that replication initiation plays in the progression of the cell cycle. Work on *C. crescentus* and *B. subtilis* utilized innovative approaches to separate DnaA's transcriptional activity from its function as a replication initiator (Hottes et al., 2005; Washington et al., 2017). One major finding from these whole-cell transcriptional analyses is that the transcriptional activity of DnaA can have a global effect on the cell. Aside from DnaA regulating the expression of various important genes, DnaA has also been found to be part of genetic networks where DnaA (directly or indirectly) regulates the expression of other transcription factors that are themselves global regulators (**Figure 2**). A global analysis of DnaA's transcriptional activity remains to be determined for *E. coli*.

To identify the transcriptional regulon of DnaA in *C. crescentus*, Hottes et al. (2005) took advantage of the ability to synchronize the *Caulobacter* cell cycle in a strain with *dnaA*'s expression regulated from an inducible promoter. DnaA-dependent and DnaA-independent changes were identified by comparing transcriptional profiles of cells grown expressing *dnaA* with cells whose *dnaA* expression was delayed. This study focused on DnaA-dependent changes of genes whose expression increased during the transition from G1 to S phase. Of the 40 genes identified to be DnaA-dependent, 13 genes included DnaA boxes on their promoter region, and three genes (*gcrA*, *ftsZ*, and *podJ*) were shown *in vitro* to have promoter regions with affinity for purified His_6DnaA. Notably, one of the DnaA-dependent genes identified was the gene encoding for GcrA, the global regulator of *Caulobacter* cell cycle (Holtzendorff et al., 2004) (**Figure 2**). The role of DnaA in the progression of the cell cycle has been predicted to be widely conserved among alpha proteobacteria (Panis et al., 2015). In the two alpha proteobacteria model systems *C. crescentus* (Hottes et al., 2005; Collier et al., 2006) and the plant symbiont *Sinorhizobium meliloti* (De Nisco et al., 2014), DnaA has been shown to be a key component of a closed genetic network that drives the progression through G1–S–G2.

In *B. subtilis*, the transcriptional regulon of DnaA has been characterized using various *in vitro* and *in vivo* high-throughput analyses (Goranov et al., 2005; Ishikawa et al., 2007; Cho et al., 2008; Breier and Grossman, 2009; Hoover et al., 2010; Smith and Grossman, 2015). To isolate the transcriptional activity of DnaA from DNA replication, Washington et al. (2017) eliminated DnaA's essential role in replication by using a strain that initiates chromosome replication from the plasmids *oriN* by the plasmids replication initiator RepN. This *B. subtilis* strain has its native *ori* and the operon flanking it (*dnaA–dnaN*) knocked out. The expressions of *dnaA* and *dnaN* were engineered at different loci regulated by two different inducible promoters. Using this system, 91% of 339 total number of genes that displayed DnaA-dependent transcriptional regulation were shown to be indirectly regulated by DnaA via Sda. DnaA regulates the transcription of *sda* by directly binding at *sda*'s promoter region (Burkholder et al., 2001; Ishikawa et al., 2007; Breier and Grossman, 2009). Sda is an inhibitor of the phosphorelay that ultimately activates Spo0A, a global regulator of sporulation and stationary phase gene expression (**Figure 2B**) (Burkholder et al., 2001; Rowland et al., 2004; Whitten et al., 2007). Consistent with previous analyses, Washington et al. identified eight sets of genes that are directly regulated by DnaA: *dnaA-dnaN*, *sda*, *yqeG-M*, *ywlC*, *ywcI-sacT*, *vpr*, *yyzF-yydABCD*, and *trmEF-rsmG-noc* (Burkholder et al., 2001; Ogura et al., 2001; Ishikawa et al., 2007; Washington et al., 2017).

TRANSCRIPTIONAL ACTIVITY OF DnaA AND ITS ROLE IN DNA REPLICATION

Replication Initiation

Aside from opening the double-stranded *ori* region, DnaA has been proposed to regulate the initiation of replication in three other ways that are dependent on its transcriptional activity. First, DnaA regulates the levels of the replication initiator by modulating its own expression in various bacterial species (as discussed in previous sections). Second, in *C. crescentus*, DnaA regulates the levels of the active form of the replication initiator (DnaA–ATP) by regulating the expression of *hdaA* (Collier and Shapiro, 2009). HdaA represses DnaA's activity as a replication initiator by promoting the hydrolysis of DnaA–ATP to DnaA–ADP (Collier and Shapiro, 2009). Third, DnaA's transcriptional activity can promote the replication of λ-derived plasmids via a potential direct interaction between DnaA and the β-subunit of RNA polymerase (RpoB) (Szalewska-Palasz et al., 1998b). In *E. coli*, DnaA's binding at *ori*-λ is necessary for transcriptional activation of *ori*-λ and also for efficient expression of λ replication initiator proteins (Szalewska-Palasz et al., 1998a).

DNA Replication

DnaA as a transcription factor regulates the levels of substrates for DNA synthesis and of components of the replisome (molecular machinery required for chromosome replication). In *E. coli*, DnaA has been shown to repress the expression of the *gua* operon involved in purine biosynthesis (Tesfa-Selase and Drabble, 1992). The role of DnaA in regulating expression levels of *nrdAB* (involved in the last step of deoxynucleotide biosynthesis) has been shown in *E. coli* and proposed in *C. crescentus* (Tuggle and Fuchs, 1986; Augustin et al., 1994; Jacobson and Fuchs, 1998; Hottes et al., 2005). In *E. coli* and *B. subtilis*, DnaA regulates the transcription of *dnaN* (DNA polymerase III, β-subunit) commonly found in the same operon downstream of *dnaA* (Hansen F. G. et al., 1982; Ogura et al., 2001; Berenstein et al., 2002; Washington et al., 2017). In *C. crescentus*, predicted DnaA boxes are found upstream of *dnaQ* (DNA polymerase III, β-subunit) and *dnaB* (DNA helicase), both of

which were identified as part of the DnaA transcriptional regulon by delaying *dnaA* transcription (Hottes et al., 2005).

DNA Repair

DNA repair plays an important role in DNA replication due to the relatively frequent replication-fork arrests that occur in bacteria even when growing under normal conditions (Cox et al., 2000). DnaA has been proposed to serve as a regulator for maintaining the integrity of the genome in response to DNA damage (Wurihan et al., 2018). In stationary phase, DnaA activates the transcription of *polA* (Quinones et al., 1997), the DNA polymerase in *E. coli* involved in DNA replication and DNA repair (Sharma and Smith, 1987; Savic et al., 1990; Quinones et al., 1997). DnaA along with LexA (global regulator of the SOS response) were shown to co-regulate the transcription of two key genes in the SOS regulon (Wurihan et al., 2018): *uvrB* (component of Nucleotide Excision Repair NER system) and *recN* [involved in repair of double-stranded breaks (Uranga et al., 2017)]. The expression of *uvrB* had previously been predicted to be regulated by DnaA based on DnaA boxes found upstream of *uvrB* (van den Berg et al., 1985; Arikan et al., 1986). In *B. subtilis*, the transcriptional activity of DnaA has been linked to the RecA-independent response to DNA stress (Goranov et al., 2005; Washington et al., 2017).

DnaA BEYOND CHROMOSOME REPLICATION

Cytokinesis

Maintaining the integrity of the chromosome after each cell division requires exquisite coordination between chromosome replication and cytokinesis. In *E. coli*, the expression levels of the gene encoding FtsZ oscillate over the cell cycle reaching the highest levels at the same time when chromosome replication initiates (Garrido et al., 1993). The cell cycle-dependent expression of *ftsZ* and the identification of three DnaA boxes found upstream of *ftsZ* (within *ftsQA*) initially suggested DnaA's involvement in *ftsZ* expression; however, the three DnaA boxes were later shown not to play a role in *ftsZ* transcription (Masters et al., 1989; Garrido et al., 1993; Smith et al., 1996). A more recent potential connection between DnaA's transcriptional activity and cytokinesis involves MioC in *E. coli* in that DnaA regulates the transcription of *mioC* that is located next to *ori* (Lother et al., 1985; Rokeach and Zyskind, 1986; Nozaki et al., 1988). The protein MioC has been proposed to promote cell division independent of *ori*'s replication and segregation (Lies et al., 2015). However, *mioC* mutants display only a moderate cell division defect. In *C. crescentus*, DnaA was shown to bind *ftsZ*'s promoter region *in vitro* and its *in vivo* expression to be dependent on DnaA's nucleotide switch (Hottes et al., 2005; Fernandez-Fernandez et al., 2011). In *B. subtilis*, DnaA was shown to bind the promoter region of *ftsL* (encodes membrane-associated Z-ring protein) *in vivo* using chromatin immunoprecipitation assays (Goranov et al., 2005). So far, the transcriptional activity of DnaA has not been shown to have a major impact with the timing of cytokinesis in any bacterial species. Our current understanding of the mechanism(s) that coordinate the timing of replication initiation with cytokinesis remains limited.

Motility, Quorum Sensing, and Heat Shock Response

The transcriptional activity of DnaA has been linked to other developmental processes. The first observation connecting DnaA to motility was in a temperature-sensitive *dnaA* mutant of *E. coli*. When grown at its permissive temperature, this mutant strain expressed significantly lower levels of flagellin that resulted in loss of motility (Mizushima et al., 1994). DnaA's regulation was later shown to not directly affect the expression of flagellin but rather indirectly via the expression of the gene (*flhD*) encoding the flagellar transcriptional regulator FlhD (Mizushima et al., 1997). Notably, the quorum sensing regulators QseB and QseC in *E. coli* have been shown to regulate the expression of *dnaA* and *flhD* (Sperandio et al., 2002; Wu et al., 2021). In the nitrogen-fixing symbiont *S. meliloti*, the transcriptional activity of DnaA has been proposed to coordinate growth phase with quorum sensing. DnaA was shown to activate the transcription of *nurR* (encodes a LuxR-like solo regulator) by binding the promoter region during exponential growth and high nutrient availability (McIntosh et al., 2019). NurR activates the transcription of *sinR*, which encodes the major regulator of *N*-acyl-homoserine lactones (AHL) production (Calatrava-Morales et al., 2018). The transcriptional activity of DnaA has also been connected to *E. coli* heat shock response. DnaA was shown to bind the promoter region and repress the transcription of the gene encoding the sigma factor RpoH (Wang and Kaguni, 1989).

FUTURE DIRECTIONS FOR REPLICATION INITIATORS

This review has focused on DnaA as the key chromosome replication initiator in bacteria. Replication initiators are different in the other domains of life such as archaea and eukaryotes. Interestingly, the replication initiators in archaea and in eukaryotes have also been shown to serve as regulators of gene expression. In eukaryotes, chromosome replication is initiated by the origin recognition complex (ORC) composed of six subunits (Bell and Dutta, 2002). In human cells, replication initiator proteins have been shown to regulate the transcription of genes involved in cell division (Hossain and Stillman, 2016). Replication initiation in archaea resembles the molecular machinery of ORC in eukaryotes (Makarova and Koonin, 2013). In the archaeon *Sulfolobus islandicus*, replication initiator proteins were shown to bind the promoter regions and regulate the expression of genes involved in DNA damage response (Sun et al., 2018). The ability of replication initiators to regulate gene expression in eukaryotes and archaea are a relatively recent discovery that continues to accumulate supporting data (Popova et al., 2018; Hu et al., 2020).

In this mini review, several aspects of the transcriptional activity of DnaA were discussed including autoregulation, its mechanism of regulation, and its role in chromosome replication and other key cellular events. Many exciting questions about

the transcriptional activity of DnaA remain to be answered. For instance, DnaA has been shown to act as an activator and a repressor of transcription (Messer and Weigel, 1997). Is there a common mechanism that differentiates and regulates these two opposing functions of DnaA? Furthermore, bipartite *oris* flanking the *dnaA* gene (as in *B. subtilis* and *Helicobacter pylori*; Krause et al., 1997; Donczew et al., 2012) suggest that DnaA can act as a replication initiator and transcription factor when bound at the same chromosomal locus upstream of *dnaA*. How are these two different functions of DnaA coordinated and differentiated over the cell cycle? Most of the work done on DnaA's nucleotide switch has been performed in *E. coli*. How widespread is DnaA's nucleotide switch used to regulate DnaA's transcriptional activity in other bacterial species? To conclude, the transcriptional activity of DnaA has only been characterized in a few bacterial species. Within these few, the specifics of DnaA's transcriptional regulon vary significantly. Thus, the characterization of replication initiators' roles in coordinating various cellular events will continue to provide exciting results for years to come.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank Eric Morgan for critically reading this manuscript and our funding NIH R01GM133833.

REFERENCES

Arikan, E., Kulkarni, M. S., Thomas, D. C., and Sancar, A. (1986). Sequences of the E. coli uvrB gene and protein. *Nucleic Acids Res.* 14, 2637–2650.

Atlung, T., Clausen, E. S., and Hansen, F. G. (1985). Autoregulation of the dnaA gene of *Escherichia coli* K12. *Mol. Gen. Genet.* 200, 442–450. doi: 10.1007/bf00425729

Atlung, T., Clausen, E., and Hansen, F. G. (1984). Autorepression of the dnaA gene of *Escherichia coli*. *Adv. Exp. Med. Biol.* 179, 199–207. doi: 10.1007/978-1-4684-8730-5_20

Augustin, L. B., Jacobson, B. A., and Fuchs, J. A. (1994). *Escherichia coli* Fis and DnaA proteins bind specifically to the nrd promoter region and affect expression of an nrd-lac fusion. *J. Bacteriol.* 176, 378–387. doi: 10.1128/jb.176.2.378-387.1994

Babu, V. M. P., Itsko, M., Baxter, J. C., Schaaper, R. M., and Sutton, M. D. (2017). Insufficient levels of the nrdAB-encoded ribonucleotide reductase underlie the severe growth defect of the Deltahda *E. coli* strain. *Mol. Microbiol.* 104, 377–399. doi: 10.1111/mmi.13632

Bell, S. P., and Dutta, A. (2002). DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71, 333–374.

Berenstein, D., Olesen, K., Speck, C., and Skovgaard, O. (2002). Genetic organization of the *Vibrio harveyi* DnaA gene region and analysis of the function of the *V. harveyi* DnaA protein in *Escherichia coli*. *J. Bacteriol.* 184, 2533–2538. doi: 10.1128/jb.184.9.2533-2538.2002

Bramhill, D., and Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* 52, 743–755. doi: 10.1016/0092-8674(88)90412-6

Braun, R. E., and Wright, A. (1986). DNA methylation differentially enhances the expression of one of the two *E. coli* dnaA promoters in vivo and in vitro. *Mol. Gen. Genet.* 202, 246–250. doi: 10.1007/bf00331644

Braun, R. E., O'day, K., and Wright, A. (1985). Autoregulation of the DNA replication gene dnaA in E. coli K-12. *Cell* 40, 159–169. doi: 10.1016/0092-8674(85)90319-8

Breier, A. M., and Grossman, A. D. (2009). Dynamic association of the replication initiator and transcription factor DnaA with the *Bacillus subtilis* chromosome during replication stress. *J. Bacteriol.* 191, 486–493. doi: 10.1128/jb.01294-08

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

Calatrava-Morales, N., Mcintosh, M., and Soto, M. J. (2018). Regulation mediated by N-Acyl homoserine lactone quorum sensing signals in the rhizobium-legume symbiosis. *Genes (Basel)* 9:263. doi: 10.3390/genes9050263

Campbell, J. L., and Kleckner, N. (1990). *E. coli* oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* 62, 967–979. doi: 10.1016/0092-8674(90)90271-f

Cheng, L., and Keiler, K. C. (2009). Correct timing of dnaA transcription and initiation of DNA replication requires trans translation. *J. Bacteriol.* 191, 4268–4275. doi: 10.1128/jb.00362-09

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

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

Collier, J., and Shapiro, L. (2009). Feedback control of DnaA-mediated replication initiation by replisome-associated HdaA protein in caulobacter. *J. Bacteriol.* 191, 5706–5716. doi: 10.1128/jb.00525-09

Collier, J., Mcadams, H. H., and Shapiro, L. (2007). A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17111–17116. doi: 10.1073/pnas.0708112104

Collier, J., Murray, S. R., and Shapiro, L. (2006). DnaA couples DNA replication and the expression of two cell cycle master regulators. *EMBO J.* 25, 346–356. doi: 10.1038/sj.emboj.7600927

Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J., and Marians, K. J. (2000). The importance of repairing stalled replication forks. *Nature* 404, 37–41. doi: 10.1038/35003501

De Nisco, N. J., Abo, R. P., Wu, C. M., Penterman, J., and Walker, G. C. (2014). Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3217–3224. doi: 10.1073/pnas.1400421111

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., and Berger, J. M. (2013). A structural framework for replication origin opening by AAA+ initiation factors. *Curr. Opin. Struct. Biol.* 23, 144–153. doi: 10.1016/j.sbi.2012.11.012

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

Fernandez-Fernandez, C., Gonzalez, D., and Collier, J. (2011). Regulation of the activity of the dual-function DnaA protein in *Caulobacter crescentus*. *PLoS One* 6:e26028. doi: 10.1371/journal.pone.0026028

Frandi, A., and Collier, J. (2019). Multilayered control of chromosome replication in *Caulobacter crescentus*. *Biochem. Soc. Trans.* 47, 187–196. doi: 10.1042/bst20180460

Froelich, J. M., Phuong, T. K., and Zyskind, J. W. (1996). Fis binding in the dnaA operon promoter region. *J. Bacteriol.* 178, 6006–6012. doi: 10.1128/jb.178.20.6006-6012.1996

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. doi: 10.1093/nar/gkg309

Fujita, M. Q., Yoshikawa, H., and Ogasawara, N. (1989). Structure of the dnaA region of *Pseudomonas* putida: conservation among three bacteria, *Bacillus subtilis*, *Escherichia coli* and *P. putida*. *Mol. Gen. Genet.* 215, 381–387. doi: 10.1007/bf00427033

Fuller, R. S., and Kornberg, A. (1983). Purified dnaA protein in initiation of replication at the *Escherichia coli* chromosomal origin of replication. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5817–5821. doi: 10.1073/pnas.80.19.5817

Garrido, T., Sanchez, M., Palacios, P., Aldea, M., and Vicente, M. (1993). Transcription of ftsZ oscillates during the cell cycle of *Escherichia coli*. *EMBO J.* 12, 3957–3965. doi: 10.1002/j.1460-2075.1993.tb06073.x

Gon, S., Camara, J. E., Klungsoyr, H. K., Crooke, E., Skarstad, K., and Beckwith, J. (2006). A novel regulatory mechanism couples deoxyribonucleotide synthesis and DNA replication in *Escherichia coli*. *EMBO J.* 25, 1137–1147. doi: 10.1038/sj.emboj.7600990

Goranov, A. I., Katz, L., Breier, A. M., Burge, C. B., and Grossman, A. D. (2005). A transcriptional response to replication status mediated by the conserved bacterial replication protein DnaA. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12932–12937. doi: 10.1073/pnas.0506174102

Han, J. S., Kwon, H. S., Yim, J. B., and Hwang, D. S. (1998). Effect of IciA protein on the expression of the nrd gene encoding ribonucleoside diphosphate reductase in *E. coli*. *Mol. Gen. Genet.* 259, 610–614. doi: 10.1007/s004380050854

Hanna, M. H., and Carl, P. L. (1975). Reinitiation of deoxyribonucleic acid synthesis by deoxyribonucleic acid initiation mutants of *Escherichia coli*: role of ribonucleic acid synthesis, protein synthesis, and cell division. *J. Bacteriol.* 121, 219–226. doi: 10.1128/jb.121.1.219-226.1975

Hansen, E. B., Hansen, F. G., and Von Meyenburg, K. (1982). The nucleotide sequence of the dnaA gene and the first part of the dnaN gene of *Escherichia coli* K-12. *Nucleic Acids Res.* 10, 7373–7385.

Hansen, F. G., and Atlung, T. (2018). The DnaA tale. *Front. Microbiol.* 9:319. doi: 10.3389/fmicb.2018.00319NODOI

Hansen, F. G., and Rasmussen, K. V. (1977). Regulation of the dnaA product in *Escherichia coli*. *Mol. Gen. Genet.* 155, 219–225. doi: 10.1007/bf00393163

Hansen, F. G., Hansen, E. B., and Atlung, T. (1982). The nucleotide sequence of the dnaA gene promoter and of the adjacent rpmH gene, coding for the ribosomal protein L34, of *Escherichia coli*. *EMBO J.* 1, 1043–1048. doi: 10.1002/j.1460-2075.1982.tb01294.x

Hansen, F. G., Koefoed, S., Sorensen, L., and Atlung, T. (1987). Titration of DnaA protein by oriC DnaA-boxes increases dnaA gene expression in *Escherichia coli*. *EMBO J* 6, 255–258. doi: 10.1002/j.1460-2075.1987.tb04747.x

Hill, N. S., Kadoya, R., Chattoraj, D. K., and Levin, P. A. (2012). Cell size and the initiation of DNA replication in bacteria. *PLoS Genet.* 8:e1002549. doi: 10.1371/journal.pgen.1002549

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

Hoover, S. E., Xu, W., Xiao, W., and Burkholder, W. F. (2010). Changes in DnaA-dependent gene expression contribute to the transcriptional and developmental response of *Bacillus subtilis* to manganese limitation in Luria-Bertani medium. *J. Bacteriol.* 192, 3915–3924. doi: 10.1128/jb.00210-10

Hossain, M., and Stillman, B. (2016). Opposing roles for DNA replication initiator proteins ORC1 and CDC6 in control of cyclin E gene transcription. *Elife* 5:e12785.

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, Y., Tareen, A., Sheu, Y. J., Ireland, W. T., Speck, C., Li, H., et al. (2020). Evolution of DNA replication origin specification and gene silencing mechanisms. *Nat. Commun.* 11:5175.

Ingmer, H., and Atlung, T. (1992). Expression and regulation of a dnaA homologue isolated from *Pseudomonas putida*. *Mol. Gen. Genet.* 232, 431–439. doi: 10.1007/bf00266248

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

Jacobson, B. A., and Fuchs, J. A. (1998). Multiple cis-acting sites positively regulate *Escherichia coli* nrd expression. *Mol. Microbiol.* 28, 1315–1322. doi: 10.1046/j.1365-2958.1998.00897.x

Jakimowicz, D., Majka, J., Lis, B., Konopa, G., Wegrzyn, G., Messer, W., et al. (2000). Structure and regulation of the dnaA promoter region in three *Streptomyces species*. *Mol. Gen. Genet.* 262, 1093–1102. doi: 10.1007/pl00008652

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

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

Kohiyama, M. (2020). Research on DnaA in the early days. *Res. Microbiol.* 171, 287–289. doi: 10.1016/j.resmic.2020.11.004

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

Kucherer, C., Lother, H., Kolling, R., Schauzu, M. A., and Messer, W. (1986). Regulation of transcription of the chromosomal dnaA gene of *Escherichia coli*. *Mol. Gen. Genet.* 205, 115–121.

Lee, Y. S., Kim, H., and Hwang, D. S. (1996). Transcriptional activation of the dnaA gene encoding the initiator for oriC replication by IciA protein, an inhibitor of in vitro oriC replication in *Escherichia coli*. *Mol. Microbiol.* 19, 389–396. doi: 10.1046/j.1365-2958.1996.485902.x

Leonard, A. C., Rao, P., Kadam, R. P., and Grimwade, J. E. (2019). Changing perspectives on the role of DnaA-ATP in orisome function and timing regulation. *Front. Microbiol.* 10:2009. doi: 10.3389/fmicb.2019.02009NODOI

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

Lobner-Olesen, A., Skarstad, K., Hansen, F. G., Von Meyenburg, K., and Boye, E. (1989). The DnaA protein determines the initiation mass of *Escherichia coli* K-12. *Cell* 57, 881–889. doi: 10.1016/0092-8674(89)90802-7

Lother, H., Kolling, R., Kucherer, C., and Schauzu, M. (1985). dnaA protein-regulated transcription: effects on the in vitro replication of *Escherichia coli* minichromosomes. *EMBO J.* 4, 555–560. doi: 10.1002/j.1460-2075.1985.tb03664.x

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

Makarova, K. S., and Koonin, E. V. (2013). Archaeology of eukaryotic DNA replication. *Cold Spring Harb. Perspect. Biol.* 5:a012963.

Masters, M., Paterson, T., Popplewell, A. G., Owen-Hughes, T., Pringle, J. H., and Begg, K. J. (1989). The effect of DnaA protein levels and the rate of initiation at oriC on transcription originating in the ftsQ and ftsA genes: in vivo experiments. *Mol. Gen. Genet.* 216, 475–483. doi: 10.1007/bf00334393

McIntosh, M., Serrania, J., and Lacanna, E. (2019). A novel LuxR-type solo of *Sinorhizobium meliloti*, NurR, is regulated by the chromosome replication coordinator, DnaA and activates quorum sensing. *Mol. Microbiol.* 112, 678–698. doi: 10.1111/mmi.14312

Mera, P. E., Kalogeraki, V. S., and Shapiro, L. (2014). Replication initiator DnaA binds at the *Caulobacter centromere* and enables chromosome segregation. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16100–16105. doi: 10.1073/pnas.1418989111

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

Mizushima, T., Koyanagi, R., Katayama, T., Miki, T., and Sekimizu, K. (1997). Decrease in expression of the master operon of flagellin synthesis in a dnaA46 mutant of *Escherichia coli*. *Biol. Pharm. Bull.* 20, 327–331. doi: 10.1248/bpb.20.327

Mizushima, T., Sasaki, S., Ohishi, H., Kobayashi, M., Katayama, T., Miki, T., et al. (1996). Molecular design of inhibitors of in vitro oriC DNA replication based on the potential to block the ATP binding of DnaA protein. *J. Biol. Chem.* 271, 25178–25183. doi: 10.1074/jbc.271.41.25178

Mizushima, T., Tomura, A., Shinpuku, T., Miki, T., and Sekimizu, K. (1994). Loss of flagellation in dnaA mutants of *Escherichia coli*. *J. Bacteriol.* 176, 5544–5546. doi: 10.1128/jb.176.17.5544-5546.1994

Nishida, S., Fujimitsu, K., Sekimizu, K., Ohmura, T., Ueda, T., and Katayama, T. (2002). A nucleotide switch in the *Escherichia coli* DnaA protein initiates chromosomal replication: evidnece from a mutant DnaA protein defective in regulatory ATP hydrolysis in vitro and in vivo. *J. Biol. Chem.* 277, 14986–14995. doi: 10.1074/jbc.m108303200

Nozaki, N., Okazaki, T., and Ogawa, T. (1988). In vitro transcription of the origin region of replication of the *Escherichia coli* chromosome. *J. Biol. Chem.* 263, 14176–14183. doi: 10.1016/s0021-9258(18)68202-x

Ogura, Y., Imai, Y., Ogasawara, N., and Moriya, S. (2001). Autoregulation of the dnaA-dnaN operon and effects of DnaA protein levels on replication initiation in *Bacillus subtilis*. *J. Bacteriol.* 183, 3833–3841. doi: 10.1128/jb.183.13.3833-3841.2001

Olliver, A., Saggioro, C., Herrick, J., and Sclavi, B. (2010). DnaA-ATP acts as a molecular switch to control levels of ribonucleotide reductase expression in *Escherichia coli*. *Mol. Microbiol.* 76, 1555–1571. doi: 10.1111/j.1365-2958.2010.07185.x

Ozaki, S. (2019). Regulation of replication initiation: lessons from *Caulobacter crescentus*. *Genes Genet. Syst.* 94, 183–196. doi: 10.1266/ggs.19-00011

Panis, G., Murray, S. R., and Viollier, P. H. (2015). Versatility of global transcriptional regulators in alpha-*Proteobacteria*: from essential cell cycle control to ancillary functions. *FEMS Microbiol. Rev.* 39, 120–133. doi: 10.1093/femsre/fuu002

Polaczek, P., and Wright, A. (1990). Regulation of expression of the dnaA gene in *Escherichia coli*: role of the two promoters and the DnaA box. *New Biol.* 2, 574–582.

Popova, V. V., Brechalov, A. V., Georgieva, S. G., and Kopytova, D. V. (2018). Nonreplicative functions of the origin recognition complex. *Nucleus* 9, 460–473. doi: 10.1080/19491034.2018.1516484

Quinones, A., Wandt, G., Kleinstauber, S., and Messer, W. (1997). DnaA protein stimulates polA gene expression in *Escherichia coli*. *Mol. Microbiol.* 23, 1193–1202. doi: 10.1046/j.1365-2958.1997.2961658.x

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.* 114, 906–919. doi: 10.1111/mmi.14531

Rokeach, L. A., and Zyskind, J. W. (1986). RNA terminating within the *E. coli* origin of replication: stringent regulation and control by DnaA protein. *Cell* 46, 763–771. doi: 10.1016/0092-8674(86)90352-1

Rowland, S. L., Burkholder, W. F., Cunningham, K. A., Maciejewski, M. W., Grossman, A. D., and King, G. F. (2004). Structure and mechanism of action of Sda, an inhibitor of the histidine kinases that regulate initiation of sporulation in *Bacillus subtilis*. *Mol. Cell* 13, 689–701. doi: 10.1016/s1097-2765(04)00084-x

Salazar, L., Guerrero, E., Casart, Y., Turcios, L., and Bartoli, F. (2003). Transcription analysis of the dnaA gene and oriC region of the chromosome of *Mycobacterium smegmatis* and *Mycobacterium bovis* BCG, and its regulation by the DnaA protein. *Microbiology (Reading)* 149, 773–784. doi: 10.1099/mic.0.25832-0

Savic, D. J., Jankovic, M., and Kostic, T. (1990). Cellular role of DNA polymerase I. *J. Basic Microbiol.* 30, 769–784.

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

Sharma, R. C., and Smith, K. C. (1987). Role of DNA polymerase I in postreplication repair: a reexamination with *Escherichia coli* delta polA. *J. Bacteriol.* 169, 4559–4564. doi: 10.1128/jb.169.10.4559-4564.1987

Smith, J. L., and Grossman, A. D. (2015). In vitro whole genome DNA binding analysis of the bacterial replication initiator and transcription factor DnaA. *PLoS Genet.* 11:e1005258. doi: 10.1371/journal.pgen.1005258

Smith, R. W., Mcateer, S., and Masters, M. (1996). The coupling between ftsZ transcription and initiation of DNA replication is not mediated by the DnaA-boxes upstream of ftsZ or by DnaA. *Mol. Microbiol.* 21, 361–372. doi: 10.1046/j.1365-2958.1996.6431365.x

Smith, R. W., Mcateer, S., and Masters, M. (1997). Autoregulation of the *Escherichia coli* replication initiator protein, DnaA, is indirect. *Mol. Microbiol.* 23, 1303–1315. doi: 10.1046/j.1365-2958.1997.3121675.x

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

Sperandio, V., Torres, A. G., and Kaper, J. B. (2002). Quorum sensing *Escherichia coli* regulators B and C (QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in *E. coli*. *Mol. Microbiol.* 43, 809–821. doi: 10.1046/j.1365-2958.2002.02803.x

Sun, L., and Fuchs, J. A. (1992). *Escherichia coli* ribonucleotide reductase expression is cell cycle regulated. *Mol. Biol. Cell* 3, 1095–1105.

Sun, L., Jacobson, B. A., Dien, B. S., Srienc, F., and Fuchs, J. A. (1994). Cell cycle regulation of the *Escherichia coli* nrd operon: requirement for a cis-acting upstream AT-rich sequence. *J. Bacteriol.* 176, 2415–2426. doi: 10.1128/jb.176.8.2415-2426.1994

Sun, M., Feng, X., Liu, Z., Han, W., Liang, Y. X., and She, Q. (2018). An Orc1/Cdc6 ortholog functions as a key regulator in the DNA damage response in Archaea. *Nucleic Acids Res.* 46, 6697–6711. doi: 10.1093/nar/gky487

Szalewska-Palasz, A., Lemieszek, E., Pankiewicz, A., Wegrzyn, A., Helinski, D. R., and Wegrzyn, G. (1998a). *Escherichia coli* dnaA gene function and bacteriophage lambda replication. *FEMS Microbiol. Lett.* 167, 27–32. doi: 10.1016/s0378-1097(98)00363-2

Szalewska-Palasz, A., Wegrzyn, A., Blaszczak, A., Taylor, K., and Wegrzyn, G. (1998b). DnaA-stimulated transcriptional activation of orilambda: *Escherichia coli* RNA polymerase beta subunit as a transcriptional activator contact site. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4241–4246. doi: 10.1073/pnas.95.8.4241

Tesfa-Selase, F., and Drabble, W. T. (1992). Regulation of the gua operon of *Escherichia coli* by the DnaA protein. *Mol. Gen. Genet.* 231, 256–264. doi: 10.1007/bf00279799

Thelander, L., and Reichard, P. (1979). Reduction of ribonucleotides. *Annu. Rev. Biochem.* 48, 133–158. doi: 10.1146/annurev.bi.48.070179.001025

Tuggle, C. K., and Fuchs, J. A. (1986). Regulation of the operon encoding ribonucleotide reductase in *Escherichia coli*: evidence for both positive and negative control. *EMBO J.* 5, 1077–1085. doi: 10.1002/j.1460-2075.1986.tb04325.x

Uranga, L. A., Reyes, E. D., Patidar, P. L., Redman, L. N., and Lusetti, S. L. (2017). The cohesin-like RecN protein stimulates RecA-mediated recombinational repair of DNA double-strand breaks. *Nat. Commun.* 8:15282.

van den Berg, E. A., Geerse, R. H., Memelink, J., Bovenberg, R. A., Magnee, F. A., and Van De Putte, P. (1985). Analysis of regulatory sequences upstream of the E. coli uvrB gene; involvement of the DnaA protein. *Nucleic Acids Res.* 13, 1829–1840.

Wang, Q. P., and Kaguni, J. M. (1987). Transcriptional repression of the dnaA gene of *Escherichia coli* by dnaA protein. *Mol. Gen. Genet.* 209, 518–525. doi: 10.1007/bf00331158

Wang, Q. P., and Kaguni, J. M. (1989). dnaA protein regulates transcriptions of the rpoH gene of *Escherichia coli*. *J. Biol. Chem.* 264, 7338–7344. doi: 10.1016/s0021-9258(18)83238-0

Wargachuk, R., and Marczynski, G. T. (2015). The *Caulobacter crescentus* homolog of DnaA (HdaA) also regulates the proteolysis of the replication initiator protein DnaA. *J. Bacteriol.* 197, 3521–3532. doi: 10.1128/jb.00460-15

Washington, T. A., Smith, J. L., and Grossman, A. D. (2017). Genetic networks controlled by the bacterial replication initiator and transcription factor DnaA in *Bacillus subtilis*. *Mol. Microbiol.* 106, 109–128. doi: 10.1111/mmi.13755

Whitten, A. E., Jacques, D. A., Hammouda, B., Hanley, T., King, G. F., Guss, J. M., et al. (2007). The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition. *J. Mol. Biol.* 368, 407–420. doi: 10.1016/j.jmb.2007.01.064

Wu, D., Baigalmaa, L., Yao, Y., Li, G., Su, M., Fan, L., et al. (2021). The *Escherichia coli* QseB/QseC signaling is required for correct timing of replication initiation and cell motility. *Gene* 773:145374. doi: 10.1016/j.gene.2020.145374

Wurihan, Gezi, Brambilla, E., Wang, S., Sun, H., Fan, L., et al. (2018). DnaA and LexA proteins regulate transcription of the uvrB gene in *Escherichia coli*: the role of DnaA in the control of the SOS regulon. *Front. Microbiol.* 9:1212. doi: 10.3389/fmicb.2018.01212NODOI

Zakrzewska-Czerwinska, J., Nardmann, J., and Schrempf, H. (1994). Inducible transcription of the dnaA gene from *Streptomyces lividans* 66. *Mol. Gen. Genet.* 242, 440–447. doi: 10.1007/bf00281794

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: 07 June 2021
doi: 10.3389/fmicb.2021.677812

Membrane Stress Caused by Unprocessed Outer Membrane Lipoprotein Intermediate Pro-Lpp Affects DnaA and Fis-Dependent Growth

*Digvijay Patil[1], Dan Xun[1], Markus Schueritz[1], Shivani Bansal[2], Amrita Cheema[2], Elliott Crooke[1,2] and Rahul Saxena[1]**

[1] Department of Biochemistry and Molecular and Cellular Biology, Georgetown University Medical Center, Washington, DC, United States, [2] Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington DC, United States

OPEN ACCESS

Edited by:
Morigen Morigen,
Inner Mongolia University, China

Reviewed by:
Julia Grimwade,
Florida Institute of Technology, United States
Tsutomu Katayama,
Kyushu University, Japan

***Correspondence:**
Rahul Saxena
rs426@georgetown.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: 08 March 2021
Accepted: 29 April 2021
Published: 07 June 2021

Citation:
Patil D, Xun D, Schueritz M, Bansal S, Cheema A, Crooke E and Saxena R (2021) Membrane Stress Caused by Unprocessed Outer Membrane Lipoprotein Intermediate Pro-Lpp Affects DnaA and Fis-Dependent Growth. Front. Microbiol. 12:677812. doi: 10.3389/fmicb.2021.677812

In *Escherichia coli*, repression of phosphatidylglycerol synthase A gene (*pgsA*) lowers the levels of membrane acidic phospholipids, particularly phosphatidylglycerol (PG), causing growth-arrested phenotype. The interrupted synthesis of PG is known to be associated with concomitant reduction of chromosomal content and cell mass, in addition to accumulation of unprocessed outer membrane lipoprotein intermediate, pro-Lpp, at the inner membrane. However, whether a linkage exists between the two altered-membrane outcomes remains unknown. Previously, it has been shown that *pgsA*$^{+}$ cells overexpressing mutant Lpp(C21G) protein have growth defects similar to those caused by the unprocessed pro-Lpp intermediate in cells lacking PG. Here, we found that the ectopic expression of DnaA(L366K) or deletion of *fis* (encoding Factor for Inversion Stimulation) permits growth of cells that otherwise would be arrested for growth due to accumulated Lpp(C21G). The DnaA(L366K)-mediated restoration of growth occurs by reduced expression of Lpp(C21G) *via* a σ^{E}-dependent small-regulatory RNA (sRNA), MicL-S. In contrast, restoration of growth *via fis* deletion is only partially dependent on the MicL-S pathway; deletion of *fis* also rescues Lpp(C21G) growth arrest in cells lacking physiological levels of PG and cardiolipin (CL), independently of MicL-S. Our results suggest a close link between the physiological state of the bacterial cell membrane and DnaA- and Fis-dependent growth.

Keywords: ***Escherichia coli*, DNA replication, acidic phospholipids, lipoprotein biogenesis, DnaA, Fis, MicL-S**

INTRODUCTION

In *Escherichia coli*, DnaA protein initiates chromosomal DNA replication once, and only once, per cell cycle (Bramhill and Kornberg, 1988; Marszalek and Kaguni, 1994; Baker and Bell, 1998). DnaA shows a high affinity for both ATP (K_D = 0.03 μM) and ADP (K_D = 0.1 μM) (Sekimizu et al., 1987). Early biochemical studies established the role of acidic phospholipids in the *in vitro* rejuvenation of inactive ADP-DnaA to active ATP-DnaA (Sekimizu and Kornberg, 1988;

Yung and Kornberg, 1988; Castuma et al., 1993). Limited proteolysis of DnaA in the presence of acidic phospholipids to generate functional fragments revealed a specific region of DnaA involved in membrane interaction (Garner and Crooke, 1996). Independently, the same region was found to insert into fluid acidic phospholipid bilayers when probed with a photoactivatable lipid analog (Garner et al., 1998). Cytolocalization studies involving an allelic replacement of the chromosomal copy of *dnaA* with a gene encoding GFP-DnaA demonstrated a discrete, longitudinal helical arrangement of GFP-DnaA along the cell periphery (Boeneman et al., 2009), supporting that DnaA is located along the inner leaflet of the cytoplasmic membrane.

Genetic studies on acidic phospholipid synthesis revealed that the absence of a functional gene for phosphatidylglycerol phosphate synthase, *pgsA* (Heacock and Dowhan, 1987; Xia and Dowhan, 1995), results in arrested growth due to reduced levels of phosphatidylglycerol (PG) and cardiolipin (CL). Site-directed mutagenesis of the membrane-binding region of DnaA offered a novel perspective on membrane–DnaA initiator protein interactions: a single-point mutation in DnaA at the hydrophobic face of the amphipathic helix, DnaA(L366K), allows growth in cells with depleted acidic phospholipid content (Zheng et al., 2001). DnaA(L366K) cannot serve as the only source of DnaA to initiate DNA synthesis both *in vivo* (Zheng et al., 2001) and *in vitro* (Saxena et al., 2011) and requires limiting levels of wild-type DnaA to function. Biochemical characterization of DnaA(L366K) demonstrated that it behaves like DnaA with respect to nucleotide-binding affinities, ATP hydrolysis, and specificity for PG and CL in promoting nucleotide exchange (Li et al., 2005). *In vitro* DMS-footprinting demonstrated that DnaA(L366K) fails to saturate low-affinity recognition sites I2 and I3 irrespective of its adenine nucleotide-bound state. However, in the presence of a limiting amount of wild-type DnaA, DnaA(L366K) promotes the formation of replication-competent DnaA–*oriC* complexes (Saxena et al., 2011). Moreover, it has also been shown that membrane-mediated DnaA rejuvenation is strongly cooperative with respect to DnaA membrane occupancy. *In vitro*, when DnaA remains present on dioleolylPG liposomes, it could transit between two occupancy states: high-density occupancy (state I) and low-density occupancy (state II) states (Aranovich et al., 2006, 2015). Only low-density occupancy state II is efficient in facilitating the nucleotide exchange that is presumably promoted by macromolecular crowding (Aranovich et al., 2006, 2015). DnaA(L366K) and is shown to require relatively lower amounts of membrane acidic phospholipids to facilitate nucleotide exchange (Aranovich et al., 2007).

In addition to DnaA–acidic membrane association, PG is also required for the biogenesis of outer membrane lipoprotein, Lpp (also known as Braun's lipoprotein), which is the most abundant protein in the *E. coli* cell envelope (Inouye, 1979). The maturation of pro-lipoprotein (pro-Lpp) involves a series of enzyme-catalyzed reactions starting with lipoprotein diacylglyceryl transferase (Lgt), which catalyzes the transfer of diacylglycerol (DAG) from PG to cysteine 21 of pro-Lpp (Braun, 1975; Sankaran and Wu, 1994; Mao et al., 2016). Signal peptidase II (Lsp) recognizes the diacylglycerated-cysteine and cleaves the signal sequence from pro-Lpp to form apolipoprotein (Paetzel et al., 2002). Previous work has indicated the requirement for equimolar levels of PG to Lpp in the maturation process of pro-Lpp (Suzuki et al., 2002).

In sum, existing findings suggest that inhibited growth of *E. coli* cells lacking membrane acidic phospholipids arises from either defective DnaA-mediated initiation of replication in the vicinity of the cytoplasmic membrane, faulty maturation of major outer membrane lipoprotein, or the intersection of both processes (Inouye et al., 1983; Cronan, 2003; Saxena et al., 2013). However, with respect to the *latter*, the manner by which immature pro-Lpp might affect normal, DnaA-dependent normal growth is still unclear.

The recent understanding of outer membrane homeostasis highlights RNA polymerase sigma factor RpoE (σ^{24}/σ^{E})-dependent, MicL-S small RNA-mediated regulation of endogenous Lpp levels for cell survival at the early stationary phase (Guo et al., 2014; Nicoloff et al., 2017). The extracytoplasmic function (ECF) sigma factor σ^{E} responds to outer membrane disorders in a complex signal cascade involving degradation of the inner membrane anti-sigma factor, RseA, and subsequent release of σ^{E} into the cytoplasm (Raivio and Silhavy, 2001; Chaba et al., 2007; Lima et al., 2013; Klein and Raina, 2019). Active σ^{E} transcribes multiple small RNAs, like MicA, RybB, and MicL to translationally inhibit outer membrane porins and lipoproteins (Johansen et al., 2006; Udekwu and Wagner, 2007; Guo et al., 2014). MicL-S is an 80-bp small RNA derived from MicL (308 bp), present at the 3′UTR adjacent to the gene-coding region for copper homeostasis protein (*cutC*), and targets specifically *lpp* mRNA (Guo et al., 2014). Furthermore, the capacity of a suppressor mutation in *lpp* (*lpp-2*) or deletion of *lpp* to restore growth to *pgsA*-null cells with a lower acidic phospholipid content supports the requisite for PG to allow maturation of pro-Lpp (Miyazaki et al., 1985; Kikuchi et al., 2000). Related, inhibited growth caused by expression of Lpp(C21G) substantiates the need for the thioester diacylglycerol modification required for proper Lpp maturation (Inouye et al., 1983).

The bacterial nucleoid resides in the vicinity of the cytoplasmic side of the inner membrane, which also serves as an important site for Lpp biogenesis. The nucleoid is dynamic and flexible in nature and undergoes various changes to regulate complex processes, including DNA replication, DNA recombination, and gene expression (Dorman and Deighan, 2003; Luijsterburg et al., 2006; Dillon and Dorman, 2010). The organization and structure of the bacterial nucleoid are affected by small nucleoid-associated proteins (NAPs), which bind to chromosomal DNA (R.T Dame, 2005; Dillon and Dorman, 2010). NAPs include proteins such as histone-like protein HupA, HupB (functions in DNA topology management) (Varshavsky et al., 1977), H-NS protein (reinforces negative supercoiling to facilitate DNA–protein–DNA bridges) (Dorman et al., 1999; S. Rimsky, 2004), and DNA bending–binding protein such as integration host factor IhfA, IhfB (S Khrapunov et al., 2006), and Fis (Factor for Inversion Stimulation) (Gille et al., 1991). Particularly, Fis protein, which is abundantly present during exponential growth,

binds to numerous sites on the chromosome and acts as a global transcription factor (Cho et al., 2008). *In vitro* studies indicated that Fis also binds to specific chromosomal sites present in *oriC* (Wold et al., 1996) and *DARSs* (DnaA-reactivating sequences) (Kasho et al., 2014); the latter appears to play a role in rejuvenation of ADP-DnaA to ATP-DnaA (Fujimitsu et al., 2009). Fis and IHF proteins complex with these regions and act as a pleiotropic regulator of initiation of replication to inhibit untimely initiations (Wold et al., 1996; Kasho et al., 2014), probably *via* altering of the DNA supercoiling (Margulies and Kaguni, 1998). *In vivo*, *E. coli* lacking *fis* are able to grow when cultured in minimal media; however, when cultured in rich media, the rapidly growing cells had increased mass but possessed fewer origins per cell, having initiated replication in an asynchronous manner (Flatten and Skarstad, 2013).

Although reports indicate that NAPs are needed to regulate several important housekeeping genes, whether they serve in a manner to link two membrane-associated outcomes in acidic phospholipid-deficient cells, perturbed chromosomal replication and defective biogenesis of outer membrane lipoproteins, has not been addressed. Here, we investigate whether accumulated Lpp(C21G) intermediates poison DnaA-dependent *oriC*-mediated replication initiation in *E. coli* with normal levels of acidic phospholipids, and if any auxiliary initiation factors play a role in linking the two membrane-associated processes. We also examine the possible role of the σ^E-MicL/Lpp protective loop in the ability of DnaA(L366K) to facilitate growth of cells carrying mutant Lpp(C21G) or lacking the ability to synthesize the acidic phospholipids.

MATERIALS AND METHODS

Restriction enzymes were purchased from New England Biolabs (NEB). All polymerase chain reactions, unless mentioned otherwise, were performed using Q5 High fidelity DNA polymerase (NEB). PCR primers were designed using GeneRunner program Version 6.5.52[1] Primers used in the study were custom synthesized from Integrated DNA Technologies. Various ingredients to reconstitute minimal media (M9) or Luria broth (LB) were purchased from Sigma or VWR.

Media, Bacterial Strain, and Plasmids

Minimal media supplemented with 0.4% glucose and 0.5% Casamino acids were used to grow cells both in liquid culture and in agar plates. **Supplementary Table 1** showed *E. coli* strains and the plasmids used in the study. *E. coli* strain BW25113 (lpp^+) or JW1667-5 (lpp^-) was obtained from *E. coli* genetic stock center[2] and their different derivatives (used in this study) were constructed in laboratory using plasmid-inducible lambda-red recombinase-mediated genome modification system.

Plasmid Construction

Plasmid pSC were generated by replacing pBR322 origin of replication and the ampicillin-resistance gene from plasmid pBAD24c (Guzman et al., 1995) with p15A origin of replication and tetracycline-resistant gene from the pACYC184 vector (Chang and Cohen, 1978; Rose, 1988). The *dnaA* or *dnaAL366K* allele was removed from previously described plasmids, pZL606 and pZL607 (Zheng et al., 2001), by digesting with *Nde*I and *Sty*I, and inserting into the same restriction enzyme site in pSC plasmid. An open reading frame for *fis* gene was amplified by PCR using chromosomal DNA as a template and a pair of primers (**Supplementary Table 2**) carrying *Nde*I and *Sty*I. The amplified product was inserted into the corresponding sites of plasmid pSC to place *fis* allele under arabinose-inducible promoter. Similarly, plasmid pC2 (Nakamura and Inouye, 1982, we received as a gift from Dr. Thomas Silhavy) carries *lpp*(C21G) gene cloned between *Xba*I and *Eco*RI, and placed under control of an IPTG-inducible promoter. Wild-type *lpp* or *lpp*(ΔK58) was amplified by PCR using chromosomal DNA as a template and a pair of primers containing *Xba*I and *Eco*RI restriction sites. The amplified fragments were inserted into the corresponding sites of the pC2 plasmids. The transformants were screened to select for the integration of the desired products by restriction digestion analysis, which were later confirmed by the automated Sanger DNA sequencing method (Genewiz).

Lambda-Red [λ-Red] Recombinase-Mediated Genetic Recombineering

For insertional-deletion of target genes, the relevant bacterial strains were electroporated with one of the two pKD46 (Datsenko and Wanner, 2000; Baba et al., 2006) derivatives carrying λ-red recombinase genes (*exo*, *beta*, and *gam*): pKD-sg-ack (Reisch and Prather, 2015) or pSIJ8 (Jensen et al., 2015) for the insertion of antibiotic resistance gene cassette. For the homology-directed recombination (HDR), donor DNA was designed to carry (1) 50–60 base-pair-long 5′-upstream and 3′-downstream DNA sequence of the gene to be deleted, and (2) reverse DNA sequence of the antibiotic resistance gene cassette to maintain expression of the antibiotic resistance gene through its promoter. All donor DNA for HDR with desired antibiotic cassette were amplified using plasmid pCas9cr4 (Reisch and Prather, 2015), for which primers were designed based on its gene sequence.

Cell Viability Assay

E. coli cultures were prepared by growing cells overnight (12–14 h) in M9+CAA+Glu media supplemented with appropriate antibiotics (ampicillin: 100 μg/ml, tetracycline: 12.5 μg/ml, and kanamycin: 50 μg/ml) at 37°C. Overnight grown cultures were diluted to inoculate an equal number of cells (∼10,000 cfu/ml) in fresh M9+CAA+Glu media supplemented with appropriate antibiotics in the presence of inducers. To perform growth curve analysis, optical densities (OD at 600 nm) were recorded after every 60 min (±10) for at least 12–18 h. Additionally, samples (1 ml) were drawn at different time point intervals from separate conditions to analyze cell viability and level of protein expression. For viability assays, OD (600 nm) values for cell suspensions were normalized to 0.6×10^9 cells/ml, which were then serially diluted by 10-fold dilutions up to 10^{-5}.

[1] http://www.generunner.net/

[2] https://cgsc.biology.yale.edu

A fixed volume (25 μl) of each serial dilution was plated on M9+CAA+Glu agar plates to obtain 200–500 colonies/plates. Viability was also assessed using a spotting assay in which serially diluted cell suspensions (2.5 μl) grown in the absence of inducer were spotted on M9+CAA+Glu-agar plates carrying various concentrations of isopropyl-β-D-galactopyranoside (IPTG), or L-Arabinose (Sigma), or both. One-way analysis of variance (ANOVA) with multiple comparison tests was performed to compare viability among groups using Prism 8 (GraphPad).

Immunoblotting

Based on the OD (600 nm), equal numbers of cells were pelleted by centrifugation at 16,000 × *g* for 10 min at 4°C. Cell pellets were resuspended in 30 μl of PBS and 12 μl (4×) of SDS sample buffer and subsequently boiled for 5 min before resolving proteins with 19% SDS-polyacrylamide-urea gels. Proteins were transferred to PVDF membranes and visualized by Ponceau S staining to evaluate the total protein content. Expressions of Lpp and Lpp(C21G) proteins were detected using anti-Lpp antiserum, and the expression of DnaA(L366K) was detected using the anti-DnaA antiserum. Subsequently, immunoblots were treated with secondary stabilized peroxidase-conjugated goat anti-rabbit antibody (Thermo Fisher Scientific) and processed with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) using AI600 Imager (GE Healthcare Biosciences). Experiments were performed with at least three biological replicates. ImageJ (Version 1.51) software was used to quantitate the band intensity of proteins detected by immunoblots and Ponceau S staining.

Lipidomic Analysis

For lipidomic analysis, MDL12 and MDL12Δ*fis* cells were grown to obtain 10–50 million cells/ml (cell number was determined by dilution plating). Bacterial cells were collected (13,500 rpm at 4°C for 10 min), and the cell pellets were flash-frozen and stored at −80°C until further analyzed. Frozen cell pellets were thawed on ice and resuspended in 40 μl of chilled water/methanol/IPA [35:25:40]. The mixture was subjected to freezing (plunging the tube containing mixture into dry ice for 30 s) and rapid heating (plunging tubes into a 37°C water bath for 90 s) twice before sonication at 30 kHz for 30 s mixed with 100 μl of ice-chilled isopropyl alcohol containing internal standards. Subsequently, the samples were placed on ice for 30 min and transferred to −20°C for 30 min following centrifugation at 13,500 rpm at 4°C for 20 min. Supernatants were collected to perform mass spectrometric (MS) analysis using the targeted LCMS-MS approach where they were resolved on an Xbridge Amide 3.5 μm, 4.6 × 100 mm column (Phenomenex) online with a triple quadrupole mass spectrometer (5500 QTRAP, SCIEX, United States) operating in the multiple reaction monitoring (MRM) mode. The declustering potential, collision energies, cell exit potential, and entrance potential optimized for each metabolite to obtain maximum ion intensity for parent and daughter ions *via* manual tuning in Analyst 1.6.3 software (SCIEX, United States). Signal intensities from all MRM Q1/Q3 ion pairs for the analyte were ranked to ensure the selection of the most intense precursor and fragment ion pair for MRM-based quantitation. This approach resulted in the selection of declustering potential, collision energies, cell exit potential, and entrance potential that maximized the generation of each fragment ion species. The metabolite ratios were calculated by normalizing the peak area of endogenous metabolites within samples normalized to the internal standard for every class of lipid molecule. The sample queue was randomized, and solvent blanks were injected to assess sample carryover. Pooled quality control (pooled QC) samples were injected after every eight samples to check for instrumental variation. Like pooled QC samples, Sciex standard QC plasma samples were also injected for lipidomic data analysis. For CL data acquisition, bovine heart extract was used as quality control. Data normalized to QC variance. QC normalized data and imputed MRM data were processed using MultiQuant 3.0.3 (Sciex). The relative quantification values of analytes were determined by calculating the ratio of peak areas of transitions of samples normalized to the peak area of the internal standard specific for every class. All statistical analyses were performed using Prism 8 (GraphPad). Two or more groups were compared with one-way ANOVA, and data for selective 124 lipid metabolites were represented in the heatmap. Multiple *t*-tests were performed using log-transformed values of metabolite peak ratio and compared using volcano plot. For volcano plots, the *p*-value cutoff was set at $p = 0.05$, and log2 fold change (fold difference) that corresponds to the statistical significance was set at ± 2.

RESULTS

Induction of Lpp(C21G) Expression Inhibits Growth That Is Restored by Ectopic Expression of DnaA(L366K)

E. coli cells lacking acidic membranes, particularly PG, caused growth-arrested phenotype. Insufficient levels of PG interrupt two important cellular processes, (i) DnaA-mediated process of replication initiation at the chromosomal origins (Fingland et al., 2012), and (ii) the maturation and proper sorting of outer membrane lipoprotein, Lpp (Matsuyama et al., 1995). Overexpression of a mutant DnaA, DnaA(L366K) protein, suppresses the growth defect in cells unable to synthesize PG. A previous study (Inouye et al., 1983) had demonstrated that bacterial cells carrying cellular levels of PG but overproducing mutant outer membrane lipoprotein, Lpp(C21G), which is unable to be processed by PG, also resulted in growth-arrested phenotype. Wondering whether there is a link of some fashion between the two growth-arrest phenotypes led us to assess the capacity of DnaA(L366K) to restore growth in bacterial cells exogenously overproducing Lpp(C21G). For this, we developed a two-plasmid system for the simultaneous expression of the Lpp(C21G) and DnaA(L366K) (see "Materials and Methods"). The genes for Lpp and Lpp(C21G) were cloned under dual Lpp and IPTG-inducible LacUV5 promoters (Inouye et al., 1983; **Supplementary Table 1**). Lpp or Lpp(C21G) was expressed either *via* the constitutive *lpp* promoter in the absence of

IPTG or *via* the combined *lpp* and *lac* promoters in the presence of IPTG in a *lacI*$^{+}$ (Nakamura and Inouye, 1982; **Supplementary Table 1**) genetic background. The genes for DnaA and DnaA(L366K) were cloned under control of the L-arabinose-inducible pBAD promoter.

As reported before, we confirmed that *E. coli lpp-null* cells carrying plasmid-borne mutant *lpp*(C21G) grown in the presence of inducer (as low as 50 μM IPTG) resulted in a 27.4-fold reduction ($p < 0.001$) in cell viability (**Supplementary Figures 1C,D**). The bacterial cells overproducing Lpp(C21G) when transformed with plasmid vector pSC in the absence or presence of inducer (0.2% arabinose) did not show any changes in the restricted growth (**Figures 1A,B**). The ectopic expression of DnaA(L366K), however, restored growth, with cells retaining viability ($p < 0.04$, **Figures 1C,D**). In contrast, overexpression of DnaA(WT) either alone, as reported elsewhere (Atlung et al., 1987; Pierucci et al., 1989; Flatten et al., 2015), or in conjunction with Lpp(C21G) led to poor growth (**Supplementary Figures 2A,B**). As a control, similar overexpression of DnaA(L366K) in cells expressing Lpp(WT) had no effect (**Figures 1E,F**).

Given that DnaA(L366K) restored growth when the inducer for Lpp(C21G) was present, we next questioned whether DnaA(L366K) affects the cellular levels of Lpp(C21G). Immunoblots revealed that pSC vector alone did not affect the expression of Lpp(C21G) (lanes 4–5, **Figures 2A,B**); however, overproduction of DnaA(L366K) caused an approximate sevenfold reduction in levels of Lpp(C21G) (lane 5, **Figures 2C,D**), as compared to cells expressing Lpp(C21G) alone (lane 4, **Figure 2C**). When compared to the reduction of Lpp(C21G) levels, we observed significantly less reduction in Lpp arising from overexpression of DnaA(L366K) (lanes 4–5, **Figures 2E,F**), suggesting a *lpp* mutation-specific response to DnaA(L366K).

Separately, we confirmed the differences in exogenous expression of Lpp and Lpp(C21G) in *lpp-null* cells (**Supplementary Figures 1G,H**). We observed approximately threefold higher Lpp(C21G) levels (lane 3, **Supplementary Figure 1H**) when compared with Lpp (lane 3, **Supplementary Figure 1G**) in the presence of 50 μM IPTG inducer in a manner similar to the expression of Lpp(C21G) (lane 4, **Figure 2C**; and lane 4, **Figure 2E**). At the same time, the Lpp(C21G) levels (lane 3, **Supplementary Figure 1H**) were still observed to be approximately 50-fold less than the endogenous levels of Lpp in *lpp*$^{+}$ cells (lane 6, **Supplementary Figure 1G**). These results further confirm that inhibited growth in *lpp-null* cells resulted

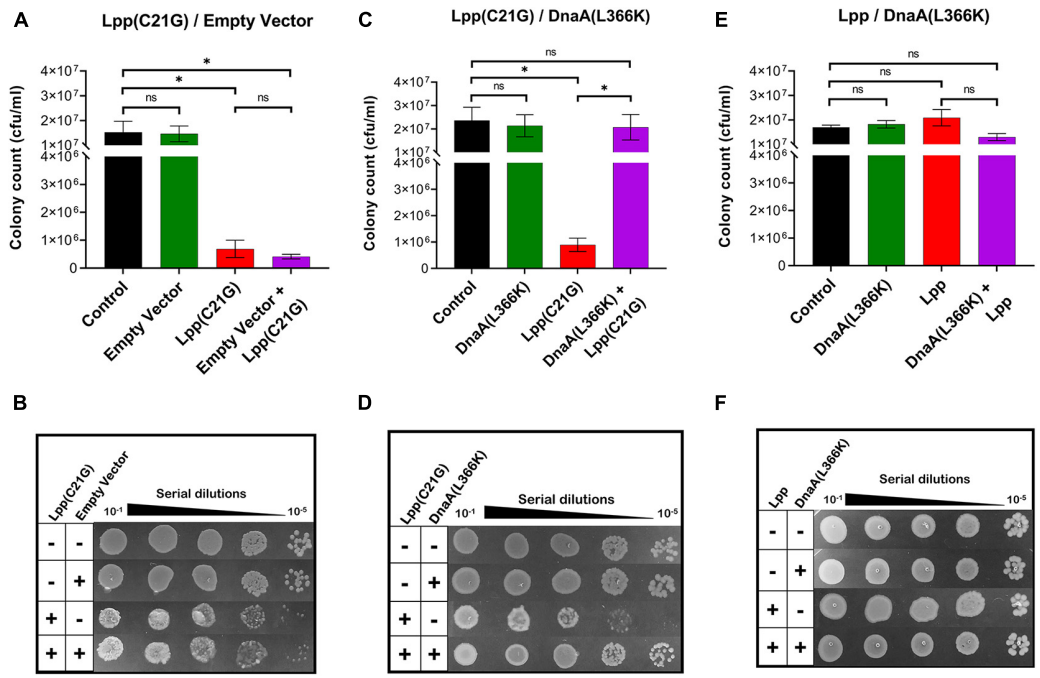


FIGURE 1 | DnaA(L366K) overexpression rescues growth in cells expressing Lpp(C21G). **(A)** Viability of *lpp-null* cells transformed with empty vector, and plasmids expressing Lpp(C21G). No inducer control (black), in the presence of 0.2% Arabinose (green) for empty vector, 50 μM IPTG (red) for Lpp(C21G), and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp(C21G) and empty vector expression in M9+CAA+Glu media. **(B)** Serial dilutions of *lpp-null* cells, capable of expressing plasmid-derived Lpp(C21G) and empty vector, on agar plates. **(C)** Viability of *lpp-null* cells expressing DnaA(L366K) and Lpp(C21G). No inducer control (black), in the presence of 0.2% Arabinose (green) for DnaA(L366K), 50 μM IPTG (red) for Lpp(C21G), and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp(C21G) and DnaA(L366K) expression in M9+CAA+Glu media. **(D)** Serial dilutions of *lpp-null* cells, capable of expressing plasmid-derived Lpp(C21G) and DnaA(L366K), on agar plates. **(E)** Viability of *lpp-null* cells expressing DnaA(L366K) and Lpp. No inducer control (black), in the presence of 0.2% Arabinose (green) for DnaA(L366K), 50 μM IPTG (red) for Lpp, and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp and DnaA(L366K) expression in M9+CAA+Glu media. **(F)** Serial dilutions of *lpp-null* cells, capable of expressing plasmid-derived Lpp and DnaA(L366K), on agar plates. Viability is expressed in cfu/ml and presented on a linear scale. Data are means ± SEM of at least three independent experiments. *$p < 0.05$, ns $p > 0.05$ in one-way ANOVA with Dunnett's multiple comparison test.

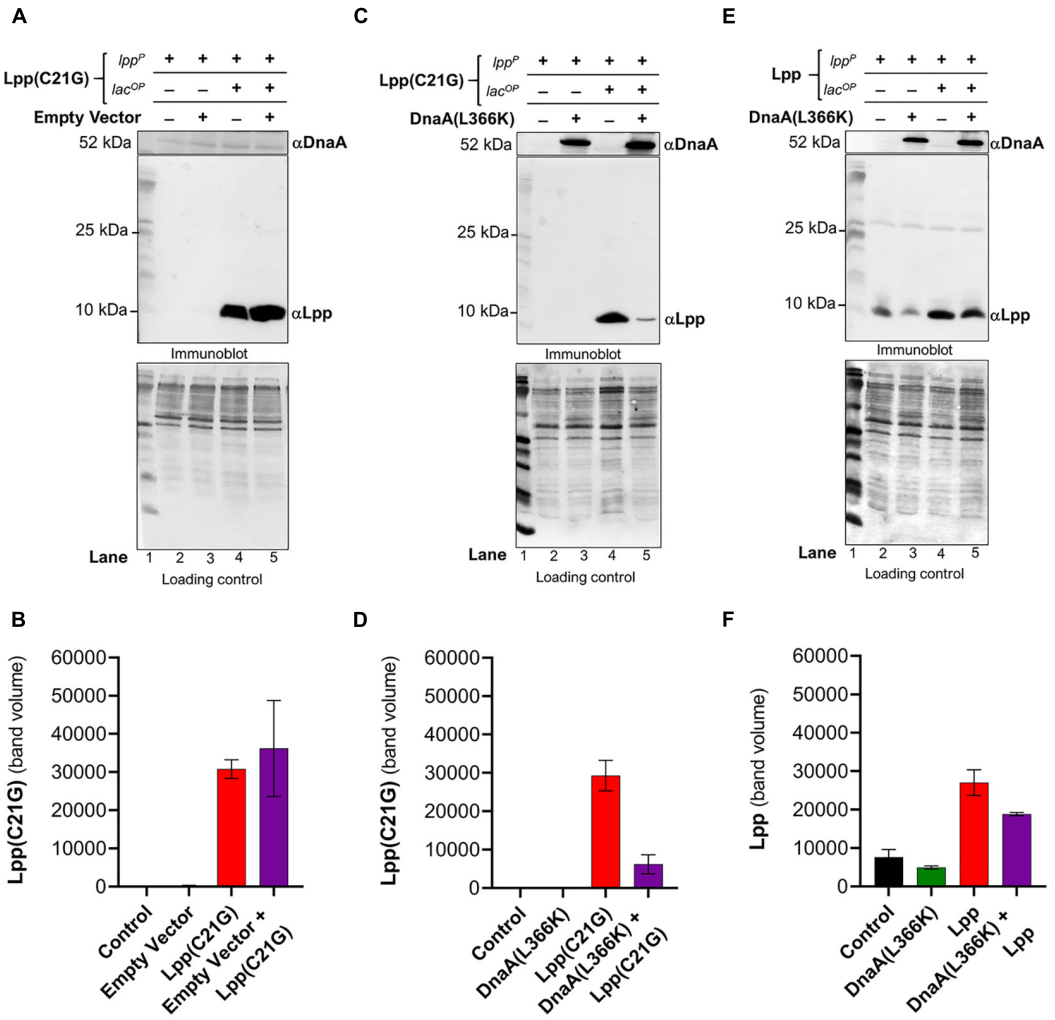


FIGURE 2 | *Ectopic* expression of DnaA(L366*K*) reduces levels of Lpp(C21*G*) even when the inducer is present. **(A,B)** Immunoblotting for the detection of DnaA and Lpp(C21G) and quantitative analysis of Lpp(C21G) levels. **(C,D)** Immunoblotting for the detection of DnaA(L366K) and Lpp(C21G) and quantitative analysis of Lpp(C21G) levels. **(E,F)** Immunoblotting for the detection of DnaA(L366K) and Lpp and quantitative analysis of Lpp levels. Lpp(C21G) and Lpp are expressed via Lpp (lpp^P) and Lac ($lacUV5^{OP}$) promoters. Loading controls for immunoblotting are ponceau-S staining for total protein normalization presented in grayscale. Data are means ± SEM of at least three independent experiments.

from the accumulation of unprocessed Lpp(C21G), but not Lpp, as shown elsewhere (Inouye et al., 1983).

Absence of the Promoter Region for MicL-S Small RNA Prevents the Ability of DnaA(L366K) to Decrease Lpp(C21G) Expression and Only Partially Restore Growth

The *E. coli* transcription factor σ^E has been recognized as an envelope stress-responsive sigma factor that senses an abnormality of the outer membrane integrity. It has been previously reported that abundance of Lpp results in the σ^E activity and that MicL and Lpp comprise a new regulatory loop that opposes membrane stress (Guo et al., 2014). Interestingly, when cells were grown without IPTG, we only observed expression *via* the *lpp* promoter of Lpp (lanes 2–3, **Figures 2E,F**) but not Lpp(C21G) (lanes 2–3, **Figures 2A,D**). This observation indicated the possible role of σ^E-dependent MicL-S-mediated envelope protective mechanisms in preventing expression of Lpp(C21G) from the *lpp* promoter in cells harboring the empty vector (lanes 2–3, **Figure 2A**) or expressing DnaA(L366K) (lanes 2–3, **Figure 2C**). However, when cells were treated with IPTG to induce Lpp(C21G), the ectopic expression of DnaA(L366K) was required to decrease Lpp(C21G) levels (lane 5, **Figure 2C**) and allow growth (**Figure 1C**), as opposed to the empty pSC vector (lane 5, **Figures 1A**, **2A**). The ability of DnaA(L366K) to reduce levels of Lpp(C21G) led us to question whether the ectopic expression of DnaA(L366K) translationally inhibits Lpp(C21G) *via* early activation of the σ^E-dependent MicL-S-mediated envelope protective loop. To test this, we chose an *E. coli* strain lacking the chromosome region for *cutC* (*cutC-null,*

Supplementary Table 1); *cutC* encodes for a copper homeostasis protein. The region for *cutC* comprises an σ^E binding site for the transcription of MicL small RNA (Guo et al., 2014; Nicoloff et al., 2017). Therefore, the lack of the *cutC* disrupts the σ^E-dependent envelope protective loop by preventing transcription of the MicL-S sRNA. Studies suggest that lack of *cutC* leads to increased levels of endogenous Lpp and affects membrane integrity (Guo et al., 2014). So, for this study, we deleted the gene-encoding region for *lpp* in cells already lacking *cutC* (Δ*cutC*) to test the effect of Lpp(C21G) on cell growth.

In *cutC-null* cells, we observed an approximate 250-fold reduction ($p < 0.02$) in cell viability (**Figures 3A,B**) by the

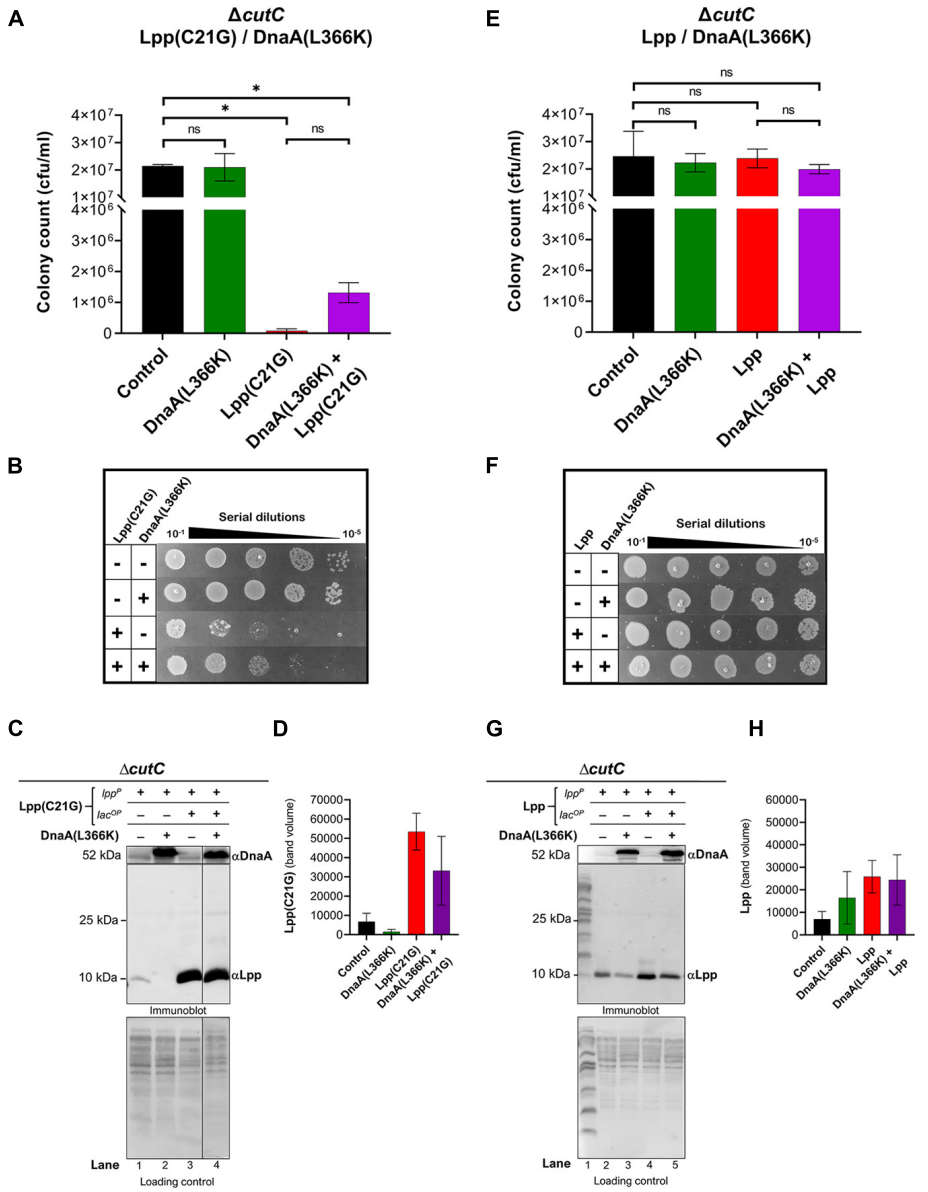


FIGURE 3 | Overexpression DnaA(L366K) restores limited growth but fails *to* reduce levels of Lpp(C21G) when the promoter for MicL-S small RNA is absent. **(A)** Viability of cutC-null cells expressing DnaA(L366K) and Lpp(C21G). No inducer control (black), in the presence of 0.2% Arabinose (green) for DnaA(L366K), 50 μM IPTG (red) for Lpp(C21G), and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp(C21G) and DnaA(L366K) expression in M9+CAA+Glu media. **(B)** Serial dilutions of cutC-null cells, capable of expressing Lpp(C21G) and DnaA(L366K), on agar plates. **(C,D)** Immunoblotting for the detection of DnaA(L366K) and Lpp(C21G) and quantitative analysis of Lpp(C21G) levels. **(E)** Viability of cutC-null cells expressing DnaA(L366K) and Lpp. No inducer control (black), in the presence of 0.2% Arabinose (green) for DnaA(L366K), 50 μM IPTG (red) for Lpp, and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp and DnaA(L366K) expression in M9+CAA+Glu media. **(F)** Serial dilutions of cutC-null cells, capable of expressing plasmid-derived Lpp and DnaA(L366K), on agar plates. **(G,H)** Immunoblotting for the detection of DnaA(L366K) and Lpp and quantitative analysis of Lpp levels. Lpp(C21G) and Lpp are expressed via Lpp (lpp^P) and Lac ($lacUV5^{OP}$) promoters. All loading controls are ponceau-S staining for total protein normalization presented in grayscale. Viability is expressed in cfu/ml and presented on a linear scale. Data are means ± SEM of at least three independent experiments. $^*p < 0.05$, ns $p > 0.05$ in one-way ANOVA with Dunnett's multiple comparison test.

limited (50 μM IPTG) expression of Lpp(C21G) as compared to that for cells not induced for Lpp(C21G) expression. In line with this observation, we co-expressed in *cutC-null* cells Lpp(C21G) along with the DnaA(L366K). DnaA(L366K) overproduction resulted in an approximate 16-fold increase in cell viability, even when Lpp(C21G) expression was induced (**Figure 3A**). Although, in comparison with *cutC*$^{+}$ cells (**Figures 1C,D**), *cutC-null* cells co-expressing Lpp(C21G) and DnaA(L366K) showed limited capacity to restore cell viability (**Figures 3A,B**).

The immunoblotting analysis further showed no significant decrease in levels of Lpp(C21G) in *cutC-null* cells in the absence or presence of DnaA(L366K) (lanes 3–4, **Figures 3C,D**), as opposed to cells with *cutC*$^{+}$ cells (lanes 4–5, **Figures 2C,D**). Of note, we also observed a limited expression of Lpp(C21G) (lane 1, **Figure 3C**) from the *lpp* promoter in cells with DnaA(L366K). Interestingly, with the wild-type Lpp control, we witnessed no change in viability (**Figures 3E,F**). Moreover, there were no significant differences in Lpp expression in the presence or absence of DnaA(L366K) overexpression (lanes 4–5, **Figures 3G,H**), indicating that DnaA(L366K) selectively decreases the expression of Lpp(C21G), but not Lpp, in a σ^{E}-dependent MicL-S-mediated manner.

Δ*fis* Cells Show the Ability to Grow Even When the Lpp(C21G) Expression Is Induced

The capacity of DnaA(L366K) to still rescue growth to some extent in the absence of MicL-S sRNA led us to then question whether the lack of acidic phospholipid-mediated accumulation of pro-lipoprotein and the ectopic expression of Lpp(C21G) poisons the DnaA-dependent, *oriC*-mediated DNA replication initiation process. Considering that DNA replication occurs in the vicinity of inner cellular membranes, where the outer membrane lipoprotein accumulates due to faulty maturation, we wondered whether the accumulated, unprocessed pro-Lpp intermediate creates stress on the chromatin structure, causing growth arrest linked to DNA replication. To characterize the adverse effect of Lpp(C21G) on the initiation, we employed

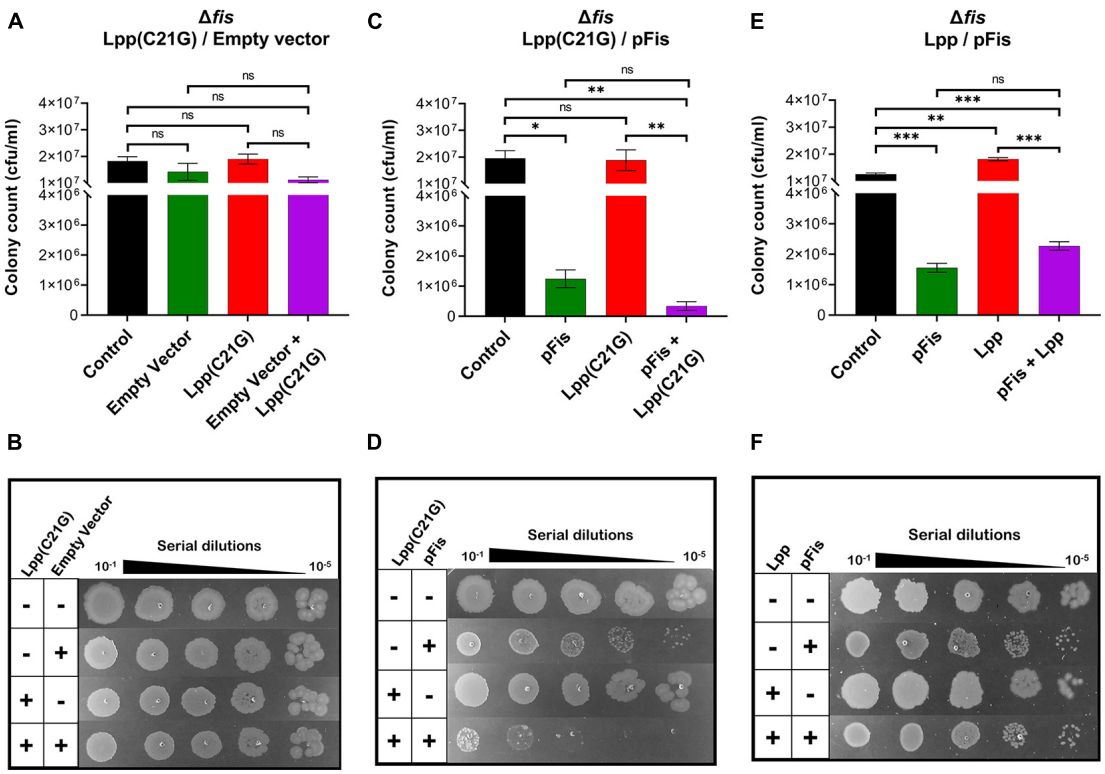


FIGURE 4 | Cells lacking fis can grow even when the inducer for Lpp(C21G) is present. **(A)** Viability of *fis-null* cells transformed with empty vector and plasmids expressing Lpp(C21G). No inducer control (black), in the presence of 0.2% Arabinose (green) for empty vector, 50 μM IPTG (red) for Lpp(C21G), and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp(C21G) and empty vector expression in M9+CAA+Glu media. **(B)** Serial dilutions of *fis-null* cells, capable of expressing plasmid-derived Lpp(C21G) and empty vector, on agar plates. **(C)** Viability of *fis-null* cells expressing Fis (pFis) and Lpp(C21G). No inducer control (black); in the presence of 0.2% Arabinose (green) for Fis, 50 μM IPTG (red) for Lpp(C21G), and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp(C21G) and Fis expression in M9+CAA+Glu media. **(D)** Serial dilutions of *fis-null* cells, capable of expressing plasmid-derived Lpp(C21G) and Fis on agar plates. **(E)** Viability of *fis-null* cells expressing Fis (pFis) and Lpp. No inducer control (black), in the presence of 0.2% Arabinose (green) for Fis, 50 μM IPTG (red) for Lpp, and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp and Fis expression in M9+CAA+Glu media. **(F)** Serial dilutions of *fis-null* cells, capable of expressing plasmid-derived Lpp and Fis, on agar plates. Viability is expressed in cfu/ml and presented on a linear scale. Data are means ± SEM of at least three independent experiments. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, ns $p > 0.05$ in one-way ANOVA with Dunnett's multiple comparison test.

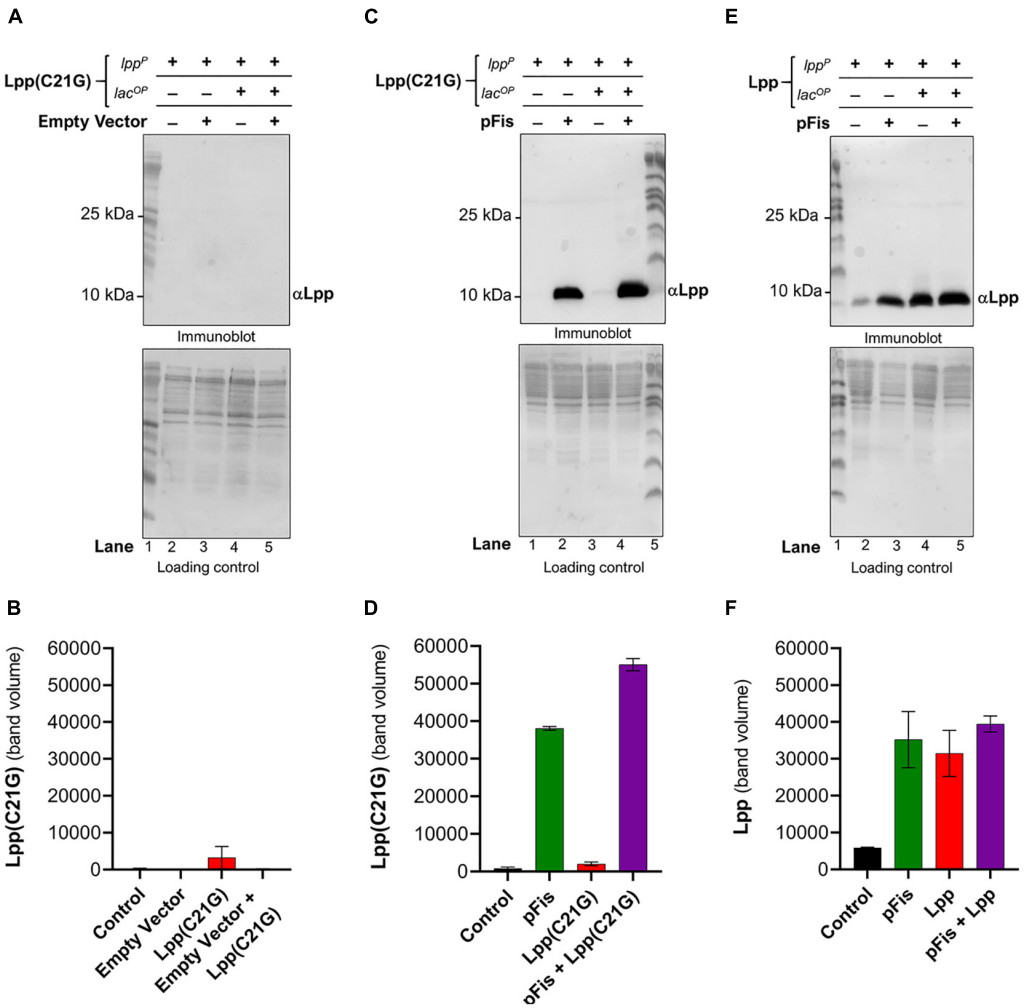


FIGURE 5 | Cells lacking fis reduces levels of Lpp(C21G) even when the inducer is present. **(A,B)** Immunoblotting of *fis-null* cells carrying an empty vector and a plasmid expressing Lpp(C21G) to detect Lpp(C21G) and quantitative analysis of Lpp(C21G) levels. **(C,D)** Immunoblotting of *fis-null* cells carrying plasmid expressing Fis (pFis) and Lpp(C21G) to detect Lpp(C21G) and quantitative analysis of Lpp(C21G) levels. **(E,F)** Immunoblotting of *fis-null* cells carrying plasmid expressing Fis (pFis) and Lpp to detect Lpp and quantitative analysis of Lpp levels. Lpp(C21G) and Lpp are expressed *via* Lpp (lpp^P) and Lac ($lacUV5^{OP}$) promoters. All loading controls are ponceau-S staining for total protein normalization presented in grayscale. Data are means ± SEM of at least three independent experiments.

a loss-of-function study targeting NAPs that participate in DnaA-dependent initiation of DNA replication at *oriC* (Riber et al., 2016). *E. coli* strains lacking gene-encoding regions for HU-α, HU-β, Ihf-A, Ihf-B, Fis, and SeqA (**Supplementary Table 1**) were used to see which of those proteins might be adversely affected by the accumulation of the Lpp intermediate by testing which NAP deletions allowed growth upon the overproduction of Lpp(C21G). Cells lacking HU-α, HU-β, Ihf-A, Ihf-B, and SeqA (**Supplementary Figures 3A–J**) were still adversely affected by the overproduction of Lpp(C21G), thus suggesting that the Lpp(C21G) does not directly affect these NAPs. In contrast, the cells lacking Fis were able to retain approximately 84 and 64% viability even when Lpp(C21G) expression was induced with IPTG at 50 and 1,000 μM, respectively. It was noted, however, that colonies were large, opaque, and irregularly lobate (**Supplementary Figures 3K,L**).

A two-plasmid system introduced Fis into *fis*-null cells expressing Lpp(C21G). As was expected, no adverse effects of Lpp(C21G) on cell viability (**Figures 4A–D**) were observed in the *fis-null* cells. However, when Fis was exogenously expressed in cells also expressing Lpp(C21G), a 57.4-fold ($p < 0.006$) loss of cell viability (**Figures 4C,D**) was observed. We also noticed that the ectopic expression of Fis alone led to a 15.7-fold ($p < 0.01$) reduction in cell viability (**Figure 4C**). However, the combined expression of Lpp(C21G) and Fis [pFis+Lpp(C21G), **Figure 4C**] had a more severe effect on viability than the expression of Fis alone (pFis, **Figure 4C**: approximately fourfold reduction). As a control, we tested the effect of Fis in cells expressing wild-type Lpp. Cell viability (**Figures 4E,F**) remained unaffected following the overexpression of Lpp alone. Like cells expressing Lpp(C21G), the overexpression of Fis alone affected viability in *fis-null* cells

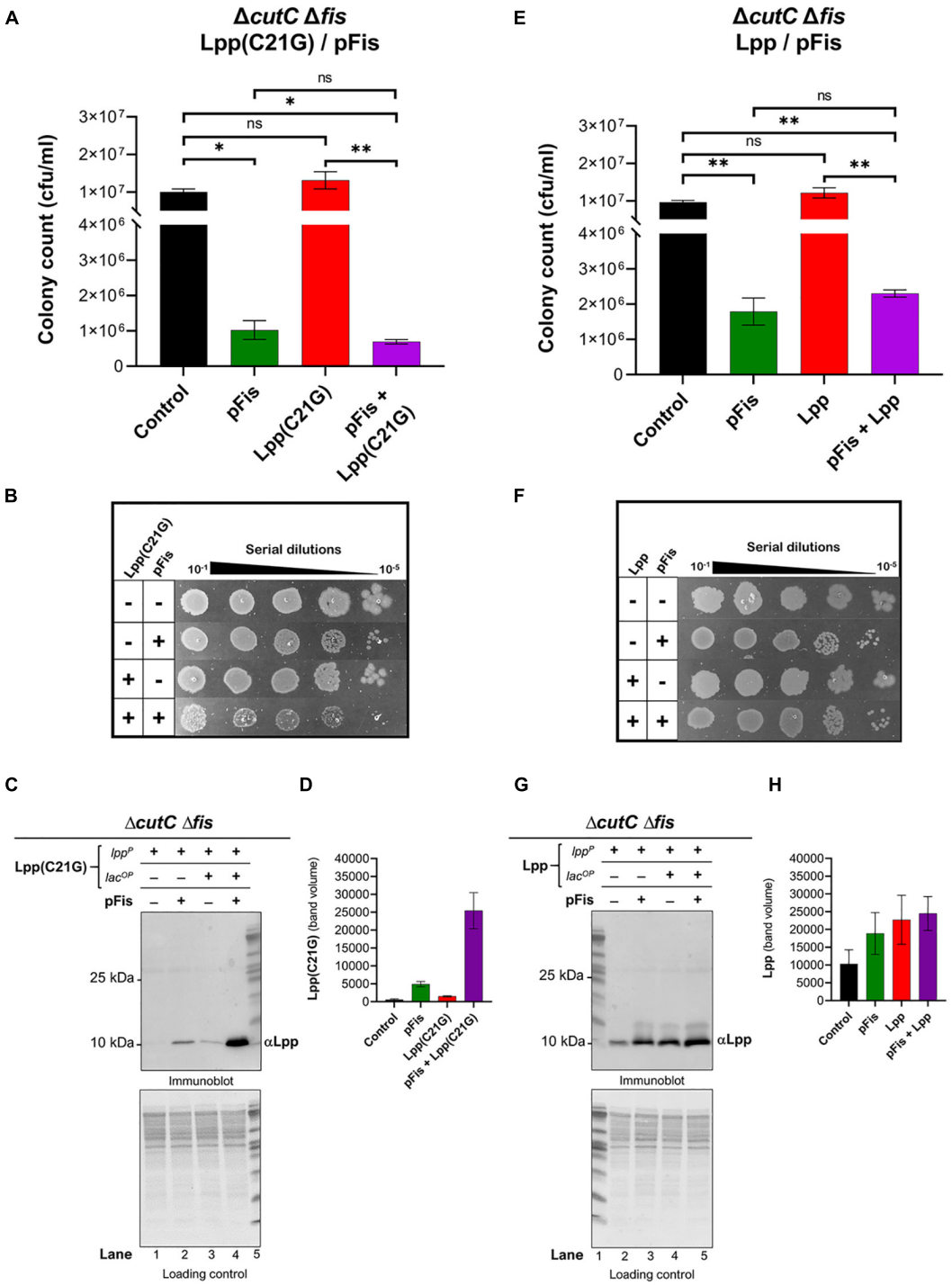


FIGURE 6 | *fis-null* cells lacking promoter for MicL-S small RNA can grow by expressing only low levels Lpp(C21G). **(A)** Viability of *cutC-null fis-null* cells expressing Fis (pFis) and Lpp(C21G). No inducer control (black); in the presence of 0.2% Arabinose (green) for Fis, 50 μM IPTG (red) for Lpp(C21G), and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp(C21G) and Fis expression in M9+CAA+Glu media. **(B)** Serial dilutions of *cutC-null fis-null* cells, capable of expressing plasmid-derived Lpp(C21G) and Fis, on agar plates. **(C,D)** Immunoblotting for the detection of Lpp(C21G) and quantitative analysis of Lpp(C21G) levels. **(E)** Viability of *cutC-null fis-null* cells expressing Fis (pFis) and Lpp. No inducer control (black), in the presence of 0.2% Arabinose (green) for Fis, 50 μM IPTG (red) for Lpp, and 50 μM IPTG + 0.2% Arabinose (purple) for both Lpp and Fis expression in M9+CAA+Glu media. **(F)** Serial dilutions of *cutC-null fis-null* cells, capable of expressing plasmid-derived Lpp and Fis, on agar plates. **(G,H)** Immunoblotting for the detection of Lpp and quantitative analysis of Lpp levels. Lpp(C21G) and Lpp are expressed *via* Lpp (lpp^{P}) and Lac ($lacUV5^{OP}$) promoters. All loading controls are ponceau-S staining for total protein normalization presented in grayscale. Viability is expressed in cfu/ml and presented on a linear scale. Data are means ± SEM of at least three independent experiments. *$p < 0.05$, **$p < 0.01$, ns $p > 0.05$ in one-way ANOVA with Dunnett's multiple comparison test.

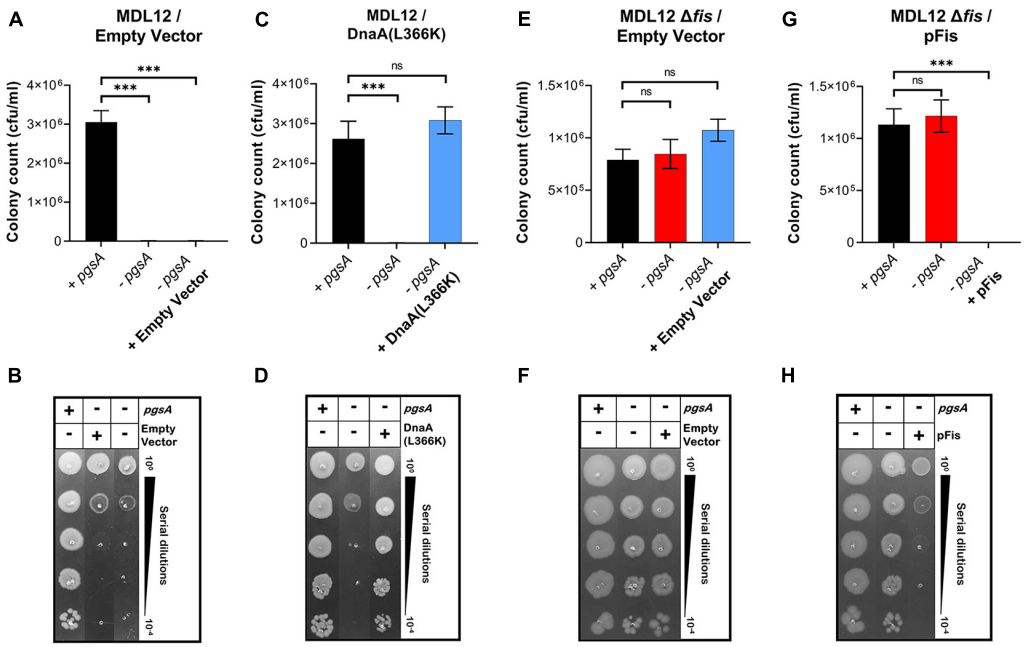


FIGURE 7 | Lack of fis in cells deficient in phosphatidylglycerol (PG) synthesis permits growth. **(A,B)** Viability of MDL12 cells carrying empty vector with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red), and 0.2% Arabinose control (blue) in M9+CAA+Glu media. **(C,D)** Viability of MDL12 cells carrying plasmid for DnaA(L366K) with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red), 0.2% Arabinose for DnaA(L366K) (blue) in M9+CAA+Glu media. **(E,F)** Viability of MDL12 *fis-null* cells carrying empty vector with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red), and 0.2% Arabinose control (blue) in M9+CAA+Glu media. **(G,H)** Viability of MDL12 *fis-null* cells carrying plasmid for Fis (pFis) with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red), and 0.2% Arabinose for Fis (blue) in M9+CAA+Glu media. Viability is expressed in cfu/ml and presented on a linear scale. Data are means ± SEM of at least three independent experiments. ***$p < 0.001$, ns $p > 0.05$ in one-way ANOVA with Dunnett's multiple comparison test.

carrying a plasmid for Lpp. The co-expression of Lpp and Fis (pFis+Lpp, **Figure 4E**) resulted in increased viability when compared with Fis expression alone (pFis, **Figure 4E**), as opposed to the co-expression of Lpp(C21G) and Fis [pFis+Lpp(C21G), **Figure 4C**].

Cells Lacking Fis Have Reduced Levels of Lpp(C21G) but Not Lpp

Seeing that the DnaA(L366K) reduces levels of Lpp(C21G) but not the wild-type Lpp, as mentioned above, we wanted to test whether *fis-null* cells restore growth to Lpp(C21G)-expressing cells, in a manner similar to the ectopic expression of DnaA(L366K), by reducing Lpp(C21G) expression but not Lpp. Surprisingly, in the cells carrying vector control, we did not detect any Lpp(C21G) when expressed from either the *lpp* promoter (lanes 2–3, **Figures 5A,B**) and only limited expression of Lpp(C21G) was seen from the combined *lpp* and *lacUV5* promoters (lanes 4–5, **Figures 5A,B**). The expression of Lpp(C21G) was evident only in the presence of Fis (lanes 2 and 4, **Figures 5C,D**). In contrast, for cells that carry plasmids for Fis and wild-type Lpp, we detected Lpp from both *lpp* (lanes 2–3, **Figures 5E,F**) and combined *lpp* and *lacUV5* promoters (lanes 4–5, **Figures 5E,F**), as opposed to Lpp(C21G) (lane 3, **Figures 5C,D**). We also noted that in cells lacking *fis*, ectopic expression of Lpp through the *lpp* promoter (lanes 2–3, **Figure 5E**) was similar to that caused by DnaA(L366K) (lanes 2–3, **Figure 2E**).

fis-null Cells Lacking the Promoter Region for MicL-S Small RNA Only Partially Express Lpp(C21G) but Still Are Able to Grow

We examined the role of MicL-S small RNA in the reduction of Lpp(C21G) levels in *fis-null* cells. In cells lacking Fis and the promoter region for MicL-S small RNA (Δ*cutC*Δ*fis*, **Supplementary Table 1**), induction of Lpp(C21G) (**Figures 6A,B**) and Lpp (**Figures 6E,F**) alone did not adversely affect cell viability, in a manner similar to *cutC*$^+$ *fis-null* cells (**Figures 4C–F**). Interestingly, immunoblotting analysis detected an approximate 1.8-fold higher expression of Lpp(C21G) (lane 3, **Figures 6C,D**) when expressed alone, as opposed to that in *cutC*$^+$ *fis-null* cells (lane 3, **Figures 5C,D**). However, the levels of Lpp(C21G) in *cutC-null, fis-null* cells (lane 3, **Figures 6C,D**) were still 16-fold less in comparison with the co-expression of Lpp(C21G) and Fis (lane 4, **Figures 6C,D**).

The ectopic expression of Fis alone still caused growth inhibition (**Figures 6A,E**). At the same time, co-expression of Fis and Lpp(C21G) caused a 14.5-fold reduction ($p = 0.02$, **Figure 6A**) in the cell viability, as opposed to a 4.2-fold decrease ($p = 0.009$, **Figure 6E**) when Fis and wild-type Lpp were co-expressed. Intriguingly, expression of Lpp was unaffected

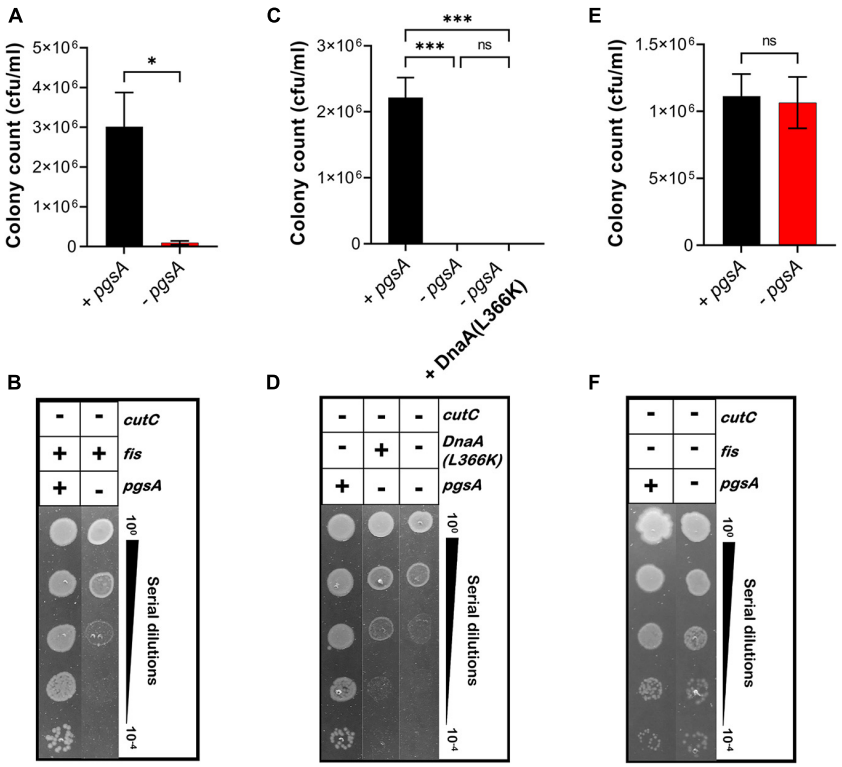


FIGURE 8 | Growth of *fis-null* cells lacking promoter for MicL-S small RNA does not depend on cellular levels of acidic phospholipids. **(A,B)** Viability of MDL12 *cutC-null* cells treated with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red) in M9+CAA+Glu medium. **(C,D)** Viability of MDL12 *cutC-null fis-null* cells treated with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red) in M9+CAA+Glu medium. **(E,F)** Viability of MDL12 *cutC-null* cells carrying plasmid for DnaA(L366K) with 1 mM IPTG for the expression of *pgsA* (black), no inducer control (red), and 0.2% Arabinose control (blue) in M9+CAA+Glu medium. Viability is expressed in cfu/ml and presented on a linear scale. Data are means ± SEM of at least three independent experiments. $*p < 0.05$, $***p < 0.001$, ns $p > 0.05$ in Student's *t*-test for analysis in plots A and C, and one-way ANOVA with Dunnett's multiple comparison test.

by the lack of *fis* (lanes 3–5, **Figures 5E,F**) or *fis* and *cutC* (lanes 3–5, **Figures 6G,H**), as was seen for cells expressing DnaA(L366K) (lanes 3–5, **Figures 2E,F**), showing a mutation-specific decrease in the gene expression of the Lpp(C21G) but not Lpp.

fis-null Cells Unable to Synthesize PG and CL Are Viable

Previously, it was seen that the ectopic expression of DnaA(L366K) allows cell growth even when acidic phospholipids are absent. We now questioned whether deletion of *fis* permitted growth in lpp^+ cells lacking normal levels of acidic phospholipids. *E. coli* MDL12 cells, which have chromosomal *pgsA* under control of the *lac* promoter, undergo growth arrest in the absence of inducer (IPTG) due to diminished levels of anionic phospholipids (Fingland et al., 2012; **Figures 7A,B**). MDL12 *fis-null* cells (**Supplementary Table 1**) were able to grow even when the inducer IPTG was absent (**Figures 7E,F**). To confirm that this outcome was due to the loss of Fis, we exogenously expressed Fis in MDL12 *fis-null* cells. We noted that the ectopic expression of Fis indeed resulted in severely diminished viability (**Figures 7G,H**) when acidic phospholipid biosynthesis was not induced.

Furthermore, we examined whether the growth-arrested phenotype of *E. coli* MDL12 cells lacking *cutC* gene could be suppressed by either overexpression of DnaA(L366K) or deletion of chromosomal *fis*. Our result suggested that the growth of cells lacking the chromosomal *cutC* allele is dependent on cellular levels of acidic phospholipids (**Figures 8A,B**). Whereas cells carrying the intact *pgsA* allowed growth when *cutC* is not present, the absence of *pgsA* allele adversely affected growth of cells lacking *cutC* (**Figures 8A,B**). Moreover, overexpression of DnaA(L366K) cannot suppress the growth-arrested phenotype when *cutC*-null cells are unable to synthesize acidic phospholipids (**Figures 8C,D**). In contrast, removal of the chromosomal *fis* allele allowed growth in a *pgsA* independent manner (**Figures 8E,F**). These results agree with our findings that either the cells overexpressing mutant Lpp protein or the lack of cellular levels of acidic phospholipids resulted in the accumulation of cytotoxic form of Lpp(C21G), causing a growth-arrested phenotype.

fis-null Cells Unable to Synthesize PG Have an Altered Lipidomic Profile

We assessed the lipid profile of MDL12 cells with and without Fis to confirm that the absence of Fis did not somehow bestow

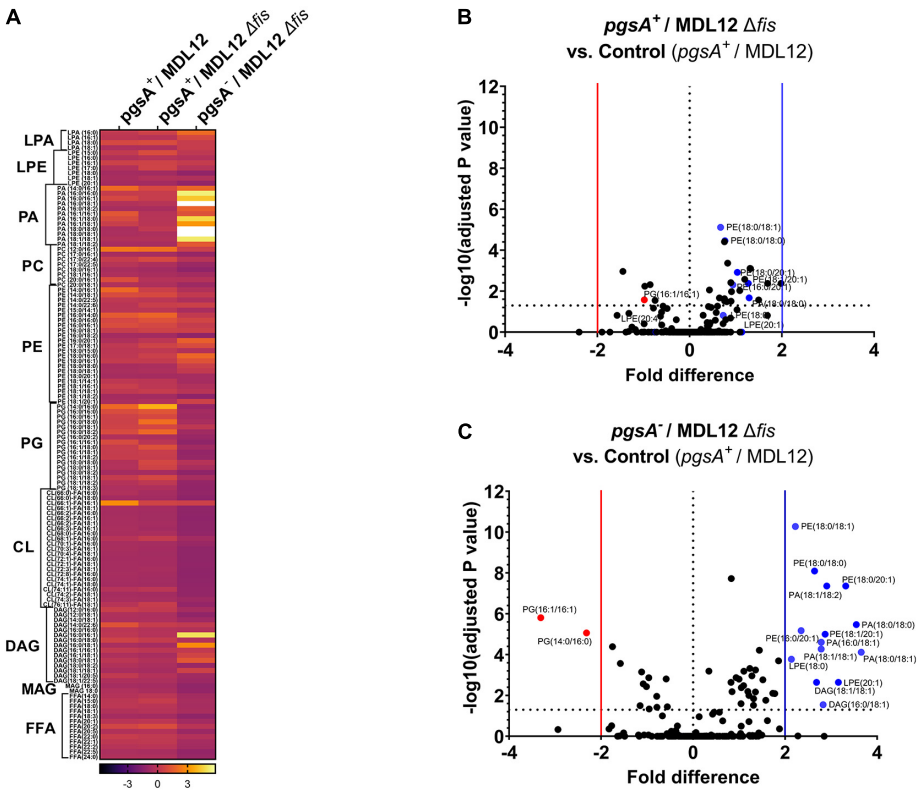


FIGURE 9 | Lipidomic analysis of MDL12 cells in the presence or absence of Fis. **(A)** Lipidomic analysis represented in heatmap comparing MDL12 cells with 1 mM IPTG for the expression of *pgsA* (*pgsA*$^+$/MDL12), MDL12 *fis-null* cells with 1 mM IPTG for the expression of *pgsA* (*pgsA*$^+$/MDL12Δ*fis*), and MDL12 *fis-null* cells without expression of *pgsA* (*pgsA*$^-$/MDL12Δ*fis*). Each row indicates the ratio values of individual lipid species for at least three biological replicates. ANOVA was performed to compare statistical differences between each group. Volcano plots comparing distribution of 191 lipids species among. **(B)** MDL12 *fis-null* cells expressing *pgsA* (*pgsA*$^+$/MDL12Δ*fis*) and control [MDL12 cells expressing of *pgsA* (*pgsA*$^+$/MDL12)]. **(C)** MDL12 *fis-null* cells *without* expression of *pgsA* (*pgsA*$^-$/MDL12Δ*fis*) and control [MDL12 cells expressing of *pgsA* (*pgsA*$^+$/MDL12)]. Horizontal dotted line indicates statistical significance with $p = 0.05$ for at least three biological replicates. A vertical red line separates red dots indicating data points with statistically significant fold decrease. A vertical blue line separates blue dots, indicating data points with statistically significant fold increase.

a normal lipid composition. A targeted LC-MS approach to estimate total lipid content detected no evident change in phospholipid content in *pgsA*-expressing cells with or without endogenous levels of Fis (**Figure 9A**). However, in cells lacking *fis* and *pgsA*, we observed that out of 191 lipid metabolites, 60 phospholipid species were differentially produced as compared to *fis*$^+$ and *pgsA*$^+$ cells with levels of PG significantly reduced in the absence of active *pgsA* and *fis* (**Figures 9B,C**). The depletion in the levels of PG was accompanied by significantly higher levels of phosphatidic acid (PA), phosphatidylethanolamine (PE), diacylglycerol (DAG), and lysophosphatidylethanolamine (LPE) in comparison with the *fis*$^+$ and *pgsA*$^+$ cells.

DISCUSSION

Cellular membranes have long been hypothesized to be involved in chromosomal replication, including accumulating evidence indicating that membranes have an influence on DnaA protein function. *E. coli pgsA*-null cells unable to synthesize PG and CL contain decreased chromosomal DNA content due to reduced frequency of initiation from chromosomal origin and thus undergo growth-arrested phenotype (Fingland et al., 2012). In addition, cells deficient in PG and CL are unable to process outer membrane lipoprotein intermediates, with the intermediates accumulating at the inner membrane. Alternatively, defective lipoprotein sorting in bacterial cells containing an intact *pgsA* allele can be achieved by overproducing mutant Lpp protein in which the Cys-21 residue is substituted with Gly (**Figures 10A,B**). Although two not obviously related processes, initiation of replication and the biogenesis of lipoproteins are associated with the cellular composition of bacterial membranes, and any linkage among them is unknown, yet. We hypothesized the possibility that unprocessed pro-Lpp intermediate might poison the DnaA–*oriC*-mediated process of replication of initiation.

A specific point mutant of chromosomal replication initiator protein, DnaA(L366K), when working in concert with wild-type DnaA, can initiate DNA synthesis both *in vivo* (Zheng et al., 2001) and *in vitro* (Saxena et al., 2011) and suppress the arrested growth of *pgsA* null cells. However, any mechanistic insights into how that occurs remain unknown. Secondly, the inability of cells to synthesize acidic membranes results in accumulation

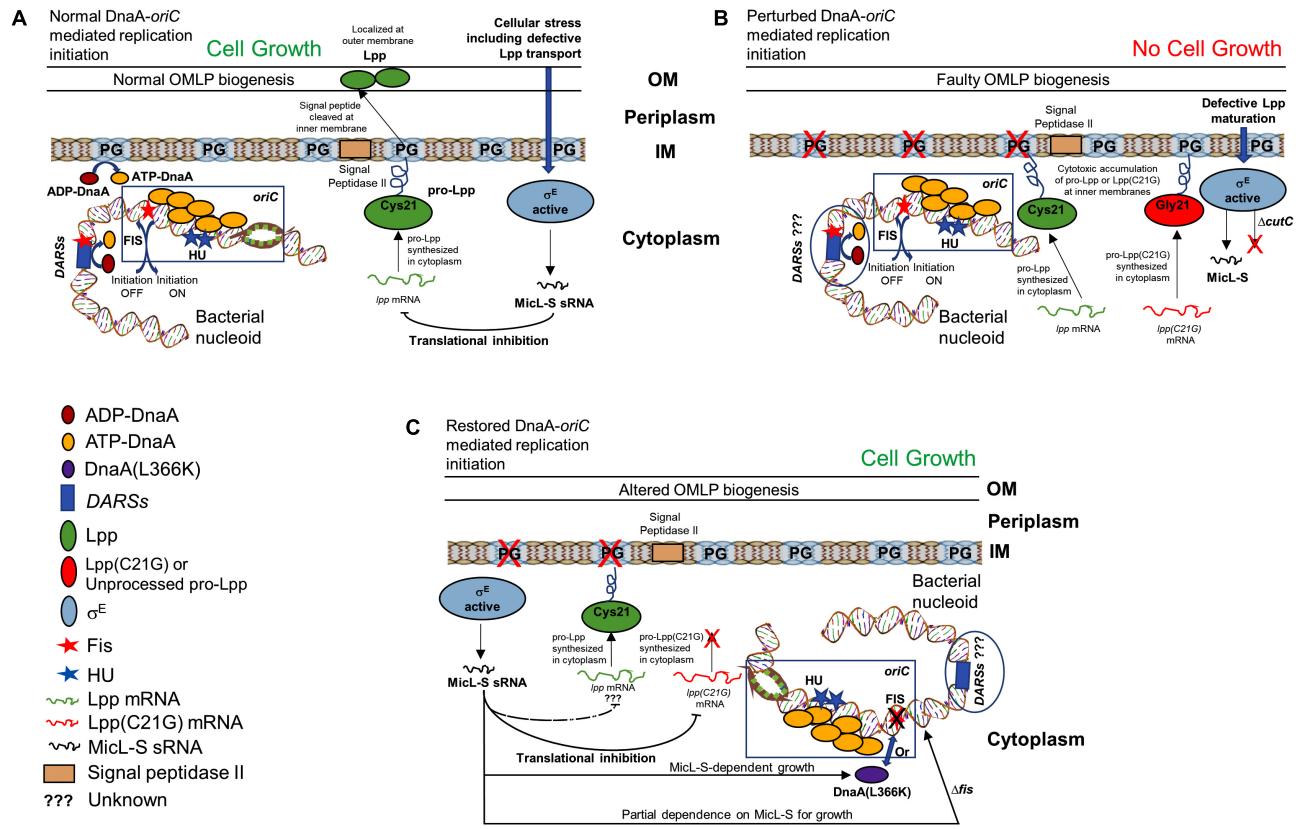


FIGURE 10 | Schematic representation of how membrane stress caused by unprocessed outer membrane lipoprotein intermediate pro-Lpp affects DnaA and Fis-dependent growth. **(A)** Cells carrying normal cellular levels of inner membrane acidic phospholipids, PG, and CL support biogenesis of outer membrane lipoprotein (Lpp) and initiate DNA replication at the chromosomal origins. ATP-DnaA binds at specific *oriC* sites, and in the presence of HU protein, unwinds the duplex DNA, leading to the formation of a functional replisome. Acidic membranes and the specific chromosomal site DARSs play a role in the rejuvenation of replication-defective ADP-DnaA to replication-proficient ATP-DnaA. Binding of Fis protein at *oriC* and DARSs helps maintain timely inhibition of replication initiation. Stress to the cellular envelope stimulates the activation of an alternative sigma factor (such as sigma factor E, σ^E), which regulates the production of Lpp protein through MicL-S sRNA in response to the defective Lpp localization. **(B)** Cells unable to synthesize PG or that accumulate cytotoxic mutant Lpp(C21G) on the inner membrane are negatively affected for growth. **(C)** Cells expressing Lpp(C21G) or lacking PG can grow in a σ^E-dependent MicL-S sRNA manner, when overexpressing DnaA(L366K). In addition, restoration of growth can occur in cells lacking *fis*, independently of MicL-S sRNA.

of unprocessed pro-lpp protein. That leads to the question of whether accumulating, unprocessed pro-Lpp poisons DnaA-dependent replication. Therefore, a valid question to ask is whether DnaA(L366K) can allow growth in cells carrying mutant Lpp(C21G). We found that the bacterial cells expressing both DnaA(L366K) and Lpp(C21G) (**Figures 1C,D**) can survive, but only with a marked decrease in the expression of the latter (lane 5, **Figures 2C,D**). There is evidence that the cellular abundance of Lpp is regulated by RNA polymerase sigma factor σ^E-dependent, small RNA, MicL-S (Guo et al., 2014). This led us to further test if cells overexpressing mutant Lpp(C21G) but lacking the σ^E-dependent MicL-S-mediated regulatory loops are capable of growth. If so, it would seem that DnaA(L366K) contains the potential to override cellular toxicity independent of MicL-S. Our results indicate that overexpression of DnaA(L366K) can bestow partial growth when Lpp(C21G) is present in growth-arresting amounts even when the σ^E-dependent envelope-protective mechanism is inactive (**Figure 3A**). Earlier work indicates that DnaA overproduction increases polA (encodes DNA polymerase I) expression in stationary-phase cultures (Quinones et al., 1997). The stimulation effect was independent of rpoS, which encodes the sigma factor for stationary-phase-inducible genes, which also include σ^E involved in the transcription of MicL.

From the subsequent loss-of-function study to determine whether Lpp(C21G) adversely affects the orisome, we found that cells lacking Fis were able to grow even when the inducer for Lpp(C21G) expression is present. However, the combined expression of Fis and Lpp(C21G) (**Figures 4C,D**) was highly inhibitory for growth in contrast to co-expression of Fis and Lpp (**Figures 4E,F**). These observations confirmed that Fis, at least when overproduced, is adversely affected by the inner membrane perturbed by accumulating Lpp intermediates. Strikingly, the near-complete depletion of Lpp(C21G) (lane 4, **Figure 5A**, and lane 3, **Figure 5C**) but not Lpp (lane 4, **Figure 5E**) in *fis-null* cells, similar to the ectopic expression of DnaA(L366K) (lane 4, **Figure 2C**, and lane 4, **Figure 2E**), underscores a concurrence in the restoration of growth with a decrease in expression of Lpp(C21G). DnaA(L366K), however, shows a strong dependence

on σ^E-mediated MicL-S-dependent envelope protective loop to decrease Lpp(C21G) expression (lane 4, **Figure 3C**), while *fis-null* cells exhibit only a partial dependence on MicL-S-mediated translational inhibition of Lpp(C21G) (lane 3, **Figure 6C**) and thus indicates the pleiotropic role of Fis.

Cells lacking Fis grow even when normal levels of PG are absent (**Figures 7E–H**), confirming that the perturbation of inner membranes affects growth in a Fis-dependent manner. Our targeted LC-MS lipidomic approach to determine changes in total phospholipid content in the presence or absence of Fis indicates a significant fold increase in PA and other phospholipid species such as PE, DAG, and LPE when both *pgsA* and *fis* are absent (**Figures 9A–C**). This observation agrees with a previous study involving the ectopic expression of DnaA(L366K) to rescue cell growth in PG-deficient cells, accumulating higher levels of PA (Zheng et al., 2001). However, the elevated levels of PA do not fully substitute for the normal combined levels of PG and CL (Heacock and Dowhan, 1987; **Figures 7A,B**). Intervention, such as (1) high levels of DnaA(L366K) (Zheng et al., 2001; **Figures 7C,D**) or (2) lack of Fis (**Figures 7E–H**), may serve to reorganize the orisome. In conclusion, we now suggest that the poisoning of orisome by unprocessed immature pro-lipoprotein present at the inner membranes causes growth inhibition when cells are lacking normal levels of acidic phospholipids (**Figure 10**). This report suggests an intricate network between the physiological state and the composition of bacterial membranes and the orisome for the proper initiation of the chromosomal DNA replication (**Figure 10C**).

The role of auxiliary replication initiation factor Fis is well-established as a global transcription factor undertaking several functions (Cho et al., 2008). In addition to acting as a global transcription factor, Fis, like other nucleoid-associated proteins, remains associated with bacterial nucleoid and serves as a DNA bending–binding protein that helps to maintain the proper confirmation of *oriC* DNA (Gille et al., 1991). The binding of Fis protein to *oriC* DNA prevents untimely initiation of DNA replication in the cells, therefore serving as a negative regulator of DnaA-mediated replication at *oriC* (Wold et al., 1996). At the time of initiation, Fis needs to be removed and replaced with DnaA protein to form replication-proficient DnaA–*oriC* complexes (Margulies and Kaguni, 1998). Certain synthetic *oriC* sequences carrying mutations in DnaA recognition sites as well as Fis or IHF binding sites cause asynchronous initiations (Weigel et al., 2001). In addition to binding *oriC*, Fis binding to *DARSs* may regulate timely initiation of replication (Kasho et al., 2014). It remains to be determined whether the lack of *fis* might alter any ongoing primary or secondary mechanisms required to maintain cellular levels of replication-proficient ATP-DnaA, and help bacterial cells to escape membrane-mediated cytotoxicity. Considering the important roles that Fis plays in both nucleoid structure and gene regulation, further studies are also needed to see how these roles are impacted in cells either stressed from insufficient acidic phospholipid levels or accumulating Lpp intermediates. These studies include how the cellular level of Fis protein controls the expression of the σ^E sigma factor, the processing of the small RNA such as MicL-S, and the levels and function of other important proteins such as different cellular proteases, which are known to upregulate when cells are stressed.

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/s.

AUTHOR CONTRIBUTIONS

RS and DP conceived the idea along with EC. DP, RS, and DX performed experiments. MS provided the assistance. DP, RS, and EC analyzed the results and wrote the manuscript. DP, SB, and AC performed and analyzed lipidomic experiments. All authors have reviewed the results and approved the final version of the manuscript.

FUNDING

The authors would like to acknowledge the Metabolomics Shared Resource at Georgetown University, which is partially supported by the NIH/NCI/CCSG grant P30-CA051008. The financial support to conduct this research was provided by the Office of the Dean of Research, Georgetown University Medical Center.

ACKNOWLEDGMENTS

We thank Dr. Thomas Silhavy for sharing plasmid pC2 and anti-Lpp antiserum, Tsutomu Katayama for sharing *Escherichia coli* KP7364 strain, Kyle Divito for careful reading of the manuscript, and Hiroshi Nakai for early discussion in bacterial growth and quorum sensing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.677812/full#supplementary-material

REFERENCES

Aranovich, A., Braier-Marcovitz, S., Ansbacher, E., Granek, R., Parola, A. H., and Fishov, I. (2015). N-terminal-mediated oligomerization of DnaA drives the occupancy-dependent rejuvenation of the protein on the membrane. *Biosci. Rep.* 35:e00250.

Aranovich, A., Gdalevsky, G. Y., Cohen-Luria, R., Fishov, I., and Parola, A. H. (2006). Membrane-catalyzed nucleotide exchange on DnaA. Effect of surface

molecular crowding. *J. Biol. Chem.* 281, 12526–12534. doi: 10.1074/jbc.m510266200

Aranovich, A., Parola, A. H., and Fishov, I. (2007). The reactivation of DnaA(L366K) requires less acidic phospholipids supporting their role in the initiation of chromosome replication in *Escherichia coli*. *FEBS Lett.* 581, 4439–4442. doi: 10.1016/j.febslet.2007.08.019

Atlung, T., Løbner-Olesen, A., and Hansen, F. G. (1987). Overproduction of DnaA protein stimulates initiation of chromosome and minichromosome replication in *Escherichia coli*. *Mol. Gen. Genet. MGG* 206, 51–59. doi: 10.1007/bf00326535

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006–2008.

Baker, T. A., and Bell, S. P. (1998). Polymerases and the replisome: machines within machines. *Cell* 92, 295–305. doi: 10.1016/s0092-8674(00)80923-x

Boeneman, K., Fossum, S., Yang, Y., Fingland, N., Skarstad, K., and Crooke, E. (2009). *Escherichia coli* DnaA forms helical structures along the longitudinal cell axis distinct from MreB filaments. *Mol. Microbiol.* 72, 645–657. doi: 10.1111/j.1365-2958.2009.06674.x

Bramhill, D., and Kornberg, A. (1988). A model for initiation at origins of DNA replication. *Cell* 54, 915–918. doi: 10.1016/0092-8674(88)90102-x

Braun, V. (1975). Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim. Biophys. Acta (BBA) Rev. Biomembr.* 415, 335–377. doi: 10.1016/0304-4157(75)90013-1

Castuma, C. E., Crooke, E., and Kornberg, A. (1993). Fluid membranes with acidic domains activate DnaA, the initiator protein of replication in *Escherichia coli*. *J. Biol. Chem.* 268, 24665–24668. doi: 10.1016/s0021-9258(19)74517-7

Chaba, R., Grigorova, I. L., Flynn, J. M., Baker, T. A., and Gross, C. A. (2007). Design principles of the proteolytic cascade governing the σE-mediated envelope stress response in *Escherichia coli*: keys to graded, buffered, and rapid signal transduction. *Genes Dev.* 21, 124–136. doi: 10.1101/gad.1496707

Chang, A. C., and Cohen, S. N. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* 134, 1141–1156. doi: 10.1128/jb.134.3.1141-1156.1978

Cho, B. K., Knight, E. M., Barrett, C. L., and Palsson, B. O. (2008). Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Res.* 18, 900–910. doi: 10.1101/gr.070276.107

Cronan, J. E. (2003). Bacterial membrane lipids: where do we stand? *Annu. Rev. Microbiol.* 57, 203–224. doi: 10.1146/annurev.micro.57.030502.090851

Dame, R. T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol. Microbiol.* 56, 858–870. doi: 10.1111/j.1365-2958.2005.04598.x

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

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

Dorman, C. J., and Deighan, P. (2003). Regulation of gene expression by histone-like proteins in bacteria. *Curr. Opin. Genet. Dev.* 13, 179–184. doi: 10.1016/s0959-437x(03)00025-x

Dorman, C. J., Hinton, J. C., and Free, A. (1999). Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol.* 7, 124–128. doi: 10.1016/s0966-842x(99)01455-9

Fingland, N., Flåtten, I., Downey, C. D., Fossum-Raunehaug, S., Skarstad, K., and Crooke, E. (2012). Depletion of acidic phospholipids influences chromosomal replication in *Escherichia coli*. *Microbiol. Open* 1, 450–466. doi: 10.1002/mbo3.46

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

Flatten, I., and Skarstad, K. (2013). The Fis protein has a stimulating role in initiation of replication in *Escherichia coli* in vivo. *PLoS One* 8:e83562. doi: 10.1371/journal.pone.0083562

Fujimitsu, K., Senriuchi, T., and Katayama, T. (2009). Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* 23, 1221–1233. doi: 10.1101/gad.1775809

Garner, J., and Crooke, E. (1996). Membrane regulation of the chromosomal replication activity of *E. coli* DnaA requires a discrete site on the protein. *EMBO J.* 15, 2313–2321. doi: 10.1002/j.1460-2075.1996.tb00585.x

Garner, J., Durrer, P., Kitchen, J., Brunner, J., and Crooke, E. (1998). Membrane-mediated release of nucleotide from an initiator of chromosomal replication, *Escherichia coli* DnaA, occurs with insertion of a distinct region of the protein into the lipid bilayer. *J. Biol. Chem.* 273, 5167–5173. doi: 10.1074/jbc.273.9.5167

Gille, H., Egan, J. B., Roth, A., and Messer, W. (1991). The FIS protein binds and bends the origin of chromosomal DNA replication, oriC, of *Escherichia coli*. *Nucleic Acids Res.* 19, 4167–4172. doi: 10.1093/nar/19.15.4167

Guo, M. S., Updegrove, T. B., Gogol, E. B., Shabalina, S. A., Gross, C. A., and Storz, G. (2014). MicL, a new σE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev.* 28, 1620–1634. doi: 10.1101/gad.243485.114

Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. O. N. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* 177, 4121–4130. doi: 10.1128/jb.177.14.4121-4130.1995

Heacock, P. N., and Dowhan, W. (1987). Construction of a lethal mutation in the synthesis of the major acidic phospholipids of *Escherichia coli*. *J. Biol. Chem.* 262, 13044–13049. doi: 10.1016/s0021-9258(18)45164-2

Inouye, M. (1979). "Lipoprotein of the outer membrane of *Escherichia coli*," in *Biomembranes*, ed. L. A. Manson (Boston, MA: Springer), 141–208. doi: 10.1007/978-1-4615-6564-2_4

Inouye, S., Franceschini, T., Sato, M., Itakura, K., and Inouye, M. (1983). Prolipoprotein signal peptidase of *Escherichia coli* requires a cysteine residue at the cleavage site. *EMBO J.* 2, 87–91. doi: 10.1002/j.1460-2075.1983.tb01386.x

Jensen, S. I., Lennen, R. M., Herrgård, M. J., and Nielsen, A. T. (2015). Seven gene deletions in seven days: fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Sci. Rep.* 5:17874.

Johansen, J., Rasmussen, A. A., Overgaard, M., and Valentin-Hansen, P. (2006). Conserved small non-coding RNAs that belong to the σE regulon: role in down-regulation of outer membrane proteins. *J. Mol. Biol.* 364, 1–8. doi: 10.1016/j.jmb.2006.09.004

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

Khrapunov, S., Brenowitz, M., Rice, P. A., and Catalano, C. E. (2006). Binding then bending: a mechanism for wrapping DNA. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19217–19218. doi: 10.1073/pnas.0609223103

Kikuchi, S., Shibuya, I., and Matsumoto, K. (2000). Viability of an *Escherichia coli* pgsANull mutant lacking detectable phosphatidylglycerol and cardiolipin. *J. Bacteriol.* 182, 371–376. doi: 10.1128/jb.182.2.371-376.2000

Klein, G., and Raina, S. (2019). Regulated assembly of LPS, its structural alterations and cellular response to LPS defects. *Int. J. Mol. Sci.* 20:356. doi: 10.3390/ijms20020356

Li, Z., Kitchen, J. L., Boeneman, K., Anand, P., and Crooke, E. (2005). Restoration of growth to acidic phospholipid-deficient cells by DnaA (L366K) is independent of its capacity for nucleotide binding and exchange and requires DnaA. *J. Biol. Chem.* 280, 9796–9801. doi: 10.1074/jbc.m413923200

Lima, S., Guo, M. S., Chaba, R., Gross, C. A., and Sauer, R. T. (2013). Dual molecular signals mediate the bacterial response to outer-membrane stress. *Science* 340, 837–841. doi: 10.1126/science.1235358

Luijsterburg, M. S., Noom, M. C., Wuite, G. J., 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

Mao, G., Zhao, Y., Kang, X., Li, Z., Zhang, Y., Wang, X., et al. (2016). Crystal structure of *E. coli* lipoprotein diacylglyceryl transferase. *Nat. Commun.* 7:10198.

Margulies, C., and Kaguni, J. M. (1998). The FIS protein fails to block the binding of DnaA protein to oriC, the *Escherichia coli* chromosomal origin. *Nucleic Acids Res.* 26, 5170–5175. doi: 10.1093/nar/26.22.5170

Marszalek, J., and Kaguni, J. M. (1994). DnaA protein directs the binding of DnaB protein in initiation of DNA replication in *Escherichia coli*. *J. Biol. Chem.* 269, 4883–4890. doi: 10.1016/s0021-9258(17)37627-5

Matsuyama, S., Tajima, T., and Tokuda, H. (1995). A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoproteins destined for the outer membrane. *EMBO J.* 14, 3365–3372. doi: 10.1002/j.1460-2075.1995.tb07342.x

Miyazaki, C., Kuroda, M., Ohta, A., and Shibuya, I. (1985). Genetic manipulation of membrane phospholipid composition in *Escherichia coli*: pgsA mutants defective in phosphatidylglycerol synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 82, 7530–7534. doi: 10.1073/pnas.82.22.7530

Nakamura, K., and Inouye, M. (1982). Construction of versatile expression cloning vehicles using the lipoprotein gene of *Escherichia coli*. *EMBO J.* 1, 771–775. doi: 10.1002/j.1460-2075.1982.tb01244.x

Nicoloff, H., Gopalkrishnan, S., and Ades, S. E. (2017). Appropriate regulation of the σE-dependent envelope stress response is necessary to maintain cell envelope integrity and stationary-phase survival in *Escherichia coli*. *J. Bacteriol.* 199, 1–17.

Paetzel, M., Karla, A., Strynadka, N. C., and Dalbey, R. E. (2002). Signal peptidases. *Chem. Rev.* 102, 4549–4580.

Pierucci, O. L. G. A., Rickert, M., and Helmstetter, C. E. (1989). DnaA protein overproduction abolishes cell cycle specificity of DNA replication from oriC in *Escherichia coli*. *J. Bacteriol.* 171, 3760–3766. doi: 10.1128/jb.171.7.3760-3766.1989

Quinones, A., Wandt, G., Kleinstäuber, S., and Messer, W. (1997). DnaA protein stimulates polA gene expression in *Escherichia coli*. *Mol. Microbiol.* 23, 1193–1202. doi: 10.1046/j.1365-2958.1997.2961658.x

Raivio, T. L., and Silhavy, T. J. (2001). Periplasmic stress and ECF sigma factors. *Annu. Rev. Microbiol.* 55, 591–624. doi: 10.1146/annurev.micro.55.1.591

Reisch, C. R., and Prather, K. L. (2015). The no-SCAR (scarless Cas9 assisted recombineering) system for genome editing in *Escherichia coli*. *Sci. Rep.* 5:15096.

Riber, L., Frimodt-Møller, J., Charbon, G., and Løbner-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

Rimsky, S. (2004). Structure of the histone-like protein H-NS and its role in regulation and genome superstructure. *Curr. Opin. Microbiol.* 7, 109–114. doi: 10.1016/j.mib.2004.02.001

Rose, R. E. (1988). The nucleotide sequence of pACYC184. *Nucleic Acids Res.* 16:355. doi: 10.1093/nar/16.1.355

Sankaran, K., and Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J. Biol. Chem.* 269, 19701–19706. doi: 10.1016/s0021-9258(17)32077-x

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

Saxena, R., Rozgaja, T., Grimwade, J., and Crooke, E. (2011). Remodeling of nucleoprotein complexes is independent of the nucleotide state of a mutant AAA+ protein. *J. Biol. Chem.* 286, 33770–33777. doi: 10.1074/jbc.m111.223495

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

Sekimizu, K., and Kornberg, A. (1988). Cardiolipin activation of dnaA protein, the initiation protein of replication in *Escherichia coli*. *J. Biol. Chem.* 263, 7131–7135. doi: 10.1016/s0021-9258(18)68615-6

Suzuki, M., Hara, H., and Matsumoto, K. (2002). Envelope disorder of *Escherichia coli* cells lacking phosphatidylglycerol. *J. Bacteriol.* 184, 5418–5425. doi: 10.1128/jb.184.19.5418-5425.2002

Udekwu, K. I., and Wagner, E. G. H. (2007). Sigma E controls biogenesis of the antisense RNA MicA. *Nucleic Acids Res.* 35, 1279–1288. doi: 10.1093/nar/gkl1154

Varshavsky, A. J., Nedospasov, S. A., Bakayev, V. V., Bakayeva, T. G., and Georgiev, G. P. (1977). Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein. *Nucleic Acids Res.* 4, 2725–2745. doi: 10.1093/nar/4.8.2725

Weigel, C., Messer, W., Preiss, S., Welzeck, M., and Morigen, B. E. (2001). The sequence requirements for a functional *Escherichia coli* replication origin are different for the chromosome and a minichromosome. *Mol. Microbiol.* 40, 498–507. doi: 10.1046/j.1365-2958.2001.02409.x

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

Xia, W., and Dowhan, W. (1995). In vivo evidence for the involvement of anionic phospholipids in initiation of DNA replication in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 783–787. doi: 10.1073/pnas.92.3.783

Yung, B. Y., and Kornberg, A. (1988). Membrane attachment activates dnaA protein, the initiation protein of chromosome replication in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 85, 7202–7205. doi: 10.1073/pnas.85.19.7202

Zheng, W., Li, Z., Skarstad, K., and Crooke, E. (2001). Mutations in DnaA protein suppress the growth arrest of acidic phospholipid-deficient *Escherichia coli* cells. *EMBO J.* 20, 1164–1172. doi: 10.1093/emboj/20.5.1164

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 August 2021
doi: 10.3389/fmicb.2021.697712

Whole-Genome Analysis Reveals That the Nucleoid Protein IHF Predominantly Binds to the Replication Origin *oriC* Specifically at the Time of Initiation

Kazutoshi Kasho[1], Taku Oshima[2]*, Onuma Chumsakul[3], Kensuke Nakamura[4], Kazuki Fukamachi[2] and Tsutomu Katayama[1]*

[1] Department of Molecular Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan, [2] Department of Biotechnology, Toyama Prefectural University, Toyama, Japan, [3] Graduate School of Biological Sciences, Nara Institute of Science and Technology, Nara, Japan, [4] Department of Life Science and Informatics, Maebashi Institute of Technology, Maebashi, Japan

OPEN ACCESS

Edited by:
Morigen Morigen, Inner Mongolia University, China

Reviewed by:
Gregory Marczynski, McGill University, Canada
Mitsuo Ogura, Tokai University, Japan

***Correspondence:**
Kazutoshi Kasho
kazutoshi.kasho@phar.kyushu-u.ac.jp
Taku Oshima
taku@pu-toyama.ac.jp

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *20 April 2021*
Accepted: *26 July 2021*
Published: *12 August 2021*

Citation:
Kasho K, Oshima T, Chumsakul O, Nakamura K, Fukamachi K and Katayama T (2021) Whole-Genome Analysis Reveals That the Nucleoid Protein IHF Predominantly Binds to the Replication Origin oriC Specifically at the Time of Initiation. Front. Microbiol. 12:697712. doi: 10.3389/fmicb.2021.697712

The structure and function of bacterial chromosomes are dynamically regulated by a wide variety of nucleoid-associated proteins (NAPs) and DNA superstructures, such as DNA supercoiling. In *Escherichia coli*, integration host factor (IHF), a NAP, binds to specific transcription promoters and regulatory DNA elements of DNA replication such as the replication origin *oriC*: binding to these elements depends on the cell cycle but underlying mechanisms are unknown. In this study, we combined GeF-seq (genome footprinting with high-throughput sequencing) with synchronization of the *E. coli* cell cycle to determine the genome-wide, cell cycle-dependent binding of IHF with base-pair resolution. The GeF-seq results in this study were qualified enough to analyze genomic IHF binding sites (e.g., *oriC* and the transcriptional promoters of *ilvG* and *osmY*) except some of the known sites. Unexpectedly, we found that before replication initiation, *oriC* was a predominant site for stable IHF binding, whereas all other loci exhibited reduced IHF binding. To reveal the specific mechanism of stable *oriC*–IHF binding, we inserted a truncated *oriC* sequence in the *terC* (replication terminus) locus of the genome. Before replication initiation, stable IHF binding was detected even at this additional *oriC* site, dependent on the specific DnaA-binding sequence DnaA box R1 within the site. DnaA oligomers formed on *oriC* might protect the *oriC*–IHF complex from IHF dissociation. After replication initiation, IHF rapidly dissociated from *oriC*, and IHF binding to other sites was sustained or stimulated. In addition, we identified a novel locus associated with cell cycle-dependent IHF binding. These findings provide mechanistic insight into IHF binding and dissociation in the genome.

Keywords: *Escherichia coli*, cell cycle, IHF, GeF-seq, *oriC*

INTRODUCTION

Within bacterial cells, chromosomal DNA forms a dynamic and highly condensed structure called the nucleoid. In *Escherichia coli*, the nucleoid is organized by a wide variety of nucleoid-associated proteins (NAPs), RNA, and DNA supercoiling. Major bacterial NAPs such as integration host factor (IHF), heat unstable (HU), and factor for inversion stimulation (Fis), along with DNA supercoiling, regulate various cellular events such as DNA replication, transcription, recombination, and nucleoid condensation (Dillon and Dorman, 2010; Seah et al., 2014). HU, one of the most abundant NAPs, binds to AT-rich DNA without sequence specificity (Ali Azam et al., 1999; Dillon and Dorman, 2010). IHF, a hetero-dimeric protein that consists of α and β subunits (encoded by *ihfA* and *ihfB*, respectively), binds to DNA in a sequence-specific manner and causes sharp (>120°) bending (Rice et al., 1996; Aeling et al., 2006). Fis, which is expressed specifically in log-phase cells (Ali Azam et al., 1999; Dillon and Dorman, 2010), also binds DNA in a sequence-specific manner, but bends DNA more moderately (∼60°) (Stella et al., 2010). The chromosomal DNA is not randomly condensed but instead dynamically forms specific structures during DNA replication and segregation (Toro and Shapiro, 2010); however, the regulatory mechanism underlying these structural changes remains unclear.

Initiation of chromosomal DNA replication is rigidly controlled to ensure that it occurs only once during the cell cycle (Katayama et al., 2010, 2017; Skarstad and Katayama, 2013; Riber et al., 2016). The initiator protein DnaA and IHF play crucial roles in initiating replication at the chromosomal origin, *oriC* (**Figures 1A–C**). DnaA forms complexes with ATP or ADP, and ATP–DnaA is active in initiation (Shimizu et al., 2016; Sakiyama et al., 2017). *oriC* contains an AT-rich duplex unwinding element (DUE), a single IHF-binding sequence (IBS), and a DnaA oligomerization region (DOR) that contains two subregions with oppositely oriented clusters of DnaA-binding sites (DnaA boxes; **Figure 1B**; Ozaki et al., 2012; Noguchi et al., 2015). The high-affinity DnaA boxes R1 and R4 at the outer edges of the DOR are oriented in opposite directions, and cooperative binding of ATP–DnaA molecules to lower affinity sites in the Left and Right DORs results in the formation of two DnaA subcomplexes (**Figure 1C**). IHF specifically binds to the 13-mer IBS and bends DNA sharply (Shimizu et al., 2016; Sakiyama et al., 2017). At initiation, binding of IHF and ATP–DnaA molecules to *oriC* induces a conformational change and local unwinding of the *oriC* DUE (**Figure 1C**), followed by loading of the DNA replication machinery.

The level of ATP–DnaA in the cell is tightly regulated, peaking at the stage of replication initiation (Kurokawa et al., 1999). During replication, DnaA-bound ATP is hydrolyzed by a complex containing Hda and the DNA-loaded clamp subunit of DNA polymerase III holoenzyme, yielding initiation-inactive ADP–DnaA (Kato and Katayama, 2001; Katayama et al., 2010). This replication-coupled negative feedback system is termed regulatory inactivation of DnaA (RIDA). In addition, the specific chromosomal locus *datA*, which contains four DnaA boxes and an IBS, is required to prevent untimely initiations (**Figures 1A,B**; Nozaki et al., 2009). Recently, we showed that ATP–DnaA molecules form specific complexes on IHF-bound *datA*, which stimulates hydrolysis of DnaA-bound ATP (Kasho and Katayama, 2013; Kasho et al., 2017). *datA*–IHF binding specifically occurs at the post-initiation stage of the cell cycle. This system for timely inactivation of DnaA, termed DDAH (*datA*-dependent DnaA–ATP hydrolysis), plays a supplemental role to RIDA in timely yielding of ADP–DnaA.

In contrast to *datA*, the DnaA-binding chromosomal loci called *DARS*s (DnaA-reactivating sequences) increase the level of ATP–DnaA by promoting nucleotide exchange of ADP–DnaA (**Figures 1A,B**; Fujimitsu et al., 2009). The *E. coli* chromosome contains at least two *DARS*s, *DARS1*, and *DARS2*, which are required for timely initiation of replication during the cell cycle. *DARS1* and *DARS2* share three highly conserved "Core" DnaA boxes that are necessary for nucleotide exchange (Fujimitsu et al., 2009; Sugiyama et al., 2019). In contrast to *DARS1*, *DARS2* requires two activator proteins, IHF and Fis (Kasho et al., 2014). *DARS2*–IHF binding is temporally regulated to occur at the pre-initiation stage of the cell cycle. Thus, IHF bindings to *oriC*, *datA*, and *DARS2* are regulated such that they occur at different times during the cell cycle; however, the regulatory mechanism and the genome-wide dynamics of timely IHF binding during the cell cycle remain to be elucidated.

In this study, we utilized GeF-seq, a unique method for identifying protein-binding sites with base-pair resolution (Chumsakul et al., 2013), to identify genome-wide distribution of IHF-binding sites. Based on results obtained by combining GeF-qPCR, GeF-seq and cell cycle synchronization, we identified novel cell cycle-coordinated IHF dynamics: at the replication initiation stage, IHF specifically binds to *oriC* and dissociates from other genomic IHF-binding loci, whereas after replication initiation, IHF is dissociated from *oriC*, and many IHF molecules stably bind to other binding loci. We analyzed IHF-binding consensus sequences at each cell cycle stage and suggest that IHF can temporarily bind to a secondary IBS on *oriC* at the stage of replication initiation. Further mechanistic analysis of *oriC* revealed that the presence of DnaA box R1, but not the chromosomal location of *oriC*, was required for stable *oriC*–IHF binding at the initiation stage. In addition, we comprehensively analyzed genomic IBS and found novel binding loci in the *ttcA* gene that are likely cell cycle-specific. Based on these findings, we propose a model of the specific mechanism involved in stable IHF binding at *oriC* and hypothesize that the modes of genomic IHF binding drastically change during the cell cycle, potentially having a global effect on the dynamics of nucleoid structures and functions.

MATERIALS AND METHODS

Bacterial Strains and Cultures

For GeF-seq, *E. coli* SH022 (*dnaC2 ihfA-cHis12*) cells were used (**Table 1**; Kasho et al., 2014; Inoue et al., 2016). To introduce *oriC*ΔDUE at the *terC* locus, a pBR322-based plasmid

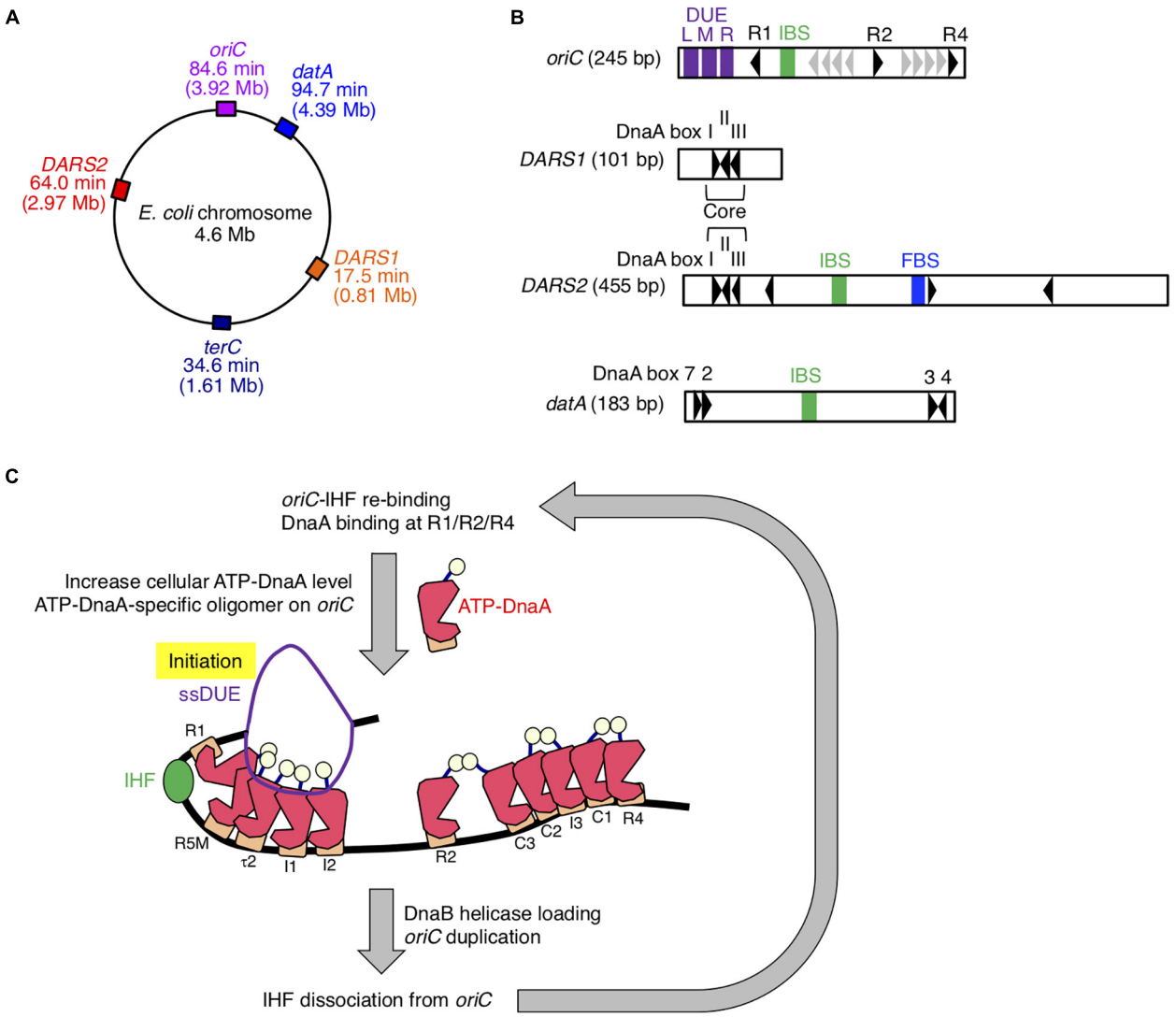


FIGURE 1 | Role of IHF binding at *oriC*. **(A)** The locations of *oriC*, *datA*, *DARS1/2*, and *TerC* loci are indicated on the chromosome. **(B)** Schematic representation of the structures of *oriC*, *datA*, and *DARS1/2* are described. Black arrowheads represent DnaA-binding sites (DnaA boxes) that match the 9-mer consensus sequence completely or with only a single mismatch. Gray arrowheads in *oriC* indicate low-affinity ATP–DnaA-binding sites R5, I1–3, τ2, and C1–3. For simplicity, a non-essential τ1 site partially overlapping with IHF-interacting region including IHF-binding sequence (IBS; green bars) is omitted (Sakiyama et al., 2017). DnaA boxes 2, 3, 4, and 7 in *datA* and "Core" DnaA boxes I–III in *DARS1/2* are essential for their activities. IBS, Fis-binding sequence (FBS; blue bar), and AT-rich repeats L, M, and R (DUE; purple bars) are also indicated. **(C)** Role and timing of IHF binding at *oriC*. IHF binds to a unique IBS and bends DNA during the pre-initiation period. At the initiation stage, ATP–DnaA becomes more abundant and forms specific oligomers on *oriC*. The left DnaA pentamer with bound IHF unwinds the DUE to initiate replication. After replication initiation, IHF quickly dissociates from *oriC*. DnaA is shown as pink polygon (domain III) with light yellow ball (domain I) and orange rectangle (domain IV). IHF is also shown as a green ball.

with *oriC*ΔDUE and *kan* (kanamycin resistant) gene flanking *frt* sequences (pKX136) was constructed. The DNA fragment bearing the *oriC*ΔDUE-*frt-kan* with the chromosomal sequence of the *terC* -proximal locus (intergenic region between *pntA* and *ydgH*; Ter-2 locus in Inoue et al., 2016) was amplified using primers TERori1-U and TERori-L (**Table 2**). Site-directed recombination was performed using SH022 cells, and the *frt-kan* region was removed using pCP20 to yield KX237 (*dnaC2 ihfA-cHis12* TER-*oriC*ΔDUE) and KX238 [*dnaC2 ihfA-cHis12* TER-*oriC*Δ(DUE-R1)] (**Table 1**; Datsenko and Wanner, 2000). To introduce *oriC*Δ(DUE-R1) at the *terC* locus, *oriC*Δ(DUE-R1)-*frt-kan* in pKX136 was amplified using primers TERori2-U and TERori-L, and similarly inserted into the SH022 genome.

Cell cultivation and cell cycle synchronization were performed according to a previously described method with minor modifications (Kasho and Katayama, 2013; Kasho et al., 2014; Inoue et al., 2016). To synchronize the *E. coli* cell cycle, cells were grown in supplemented M9 medium at 30°C, the permissive temperature for *dnaC2*, until the A_{660} of the culture reached 0.03, followed by further incubation at 38°C, the restrictive

TABLE 1 | list of *E. coli* strains.

Strains	Genotypes	References
SH022	MG1655 *ihfA-cHis12 dnaC2 zjj18::cat*	Kasho et al., 2014
KYA018	MG1655 *dnaC2 zjj18::cat*	Kasho and Katayama, 2013
KX237	SH022 *zdg7::oriC*ΔDUE	This study
KX238	SH022 *zdg7::oriC*Δ(DUE-R1)	This study

TABLE 2 | List of oligonucleotides.

Names	Sequences
ORI_1	CTGTGAATGATCGGTGATC
KWoriCRev	GTGGATAACTCTGTCAGGAAGCTTG
RTYLCC-L	GGCGTGGTAAAGGGTATCG
RTYLCC-R	TCTGCGGGGTGATGGTAAAG
ilvG-U	TCCTCGGTTATGTTTTTAAGGTC
ilvG-L	TGCACTTGGACGAGGAAAG
rhlB-U	TACGTCACGACCCGCCAG
rhlB-L	CATCCGAAGGTTGTAGAAGC
glnH-U	AATGGTGCATCTTCAGGGTATTG
glnH-L	CACATATATGAAAAAATCGTGCCAG
osmY-U	ATCACAATTTTGAAACCGCTC
osmY-L	CTGTCAATTTCCCTTCCTTATTAGC
TERori1-U	GATAAAGACTGATAATTGTCTTCGACGGTCGGGT AAAACGAGACAATCGCACTGCCCTGTGGATAAC
TERori2-U	GATAAAGACTGATAATTGTCTTCGACGGTCGGGT AAAACGAGACACAAGGATCCGGCTTTTAAGATCAAC
TERori-L	TGTATAAGTTAATTTAATGTTAAGTAGTGATTCGTG CCGGGGCGACCATATGAATATCCTCCTTAGTTCC
RT-TERoriC-U	CTCGCAAAATATTAACGATTCAGCCG
RT-TERoriC-L	TGTCTCGTTTTACCCGACCG
RT-NoriC-U2	GATCTGTTCTATTGTGATCTCTTATTAGGATCG
RT-NoriC-L2	CACAGTTAATGATCCTTTCCAGGTTG

temperature, for 90 min. Cells were immediately cooled to 30°C by addition of ice-cold medium and then incubated for an additional 10 or 20 min. Cell samples were withdrawn at the indicated time points, collected, and crosslinked with 3% (final) formaldehyde for 5 min.

In situ DNase I Digestion, His-Tag Affinity Purification of IHF–DNA Complexes, and Sequencing

To hydrolyze the cell wall without osmotic burst, cells were treated with 1 mg/ml egg white lysozyme in 2 ml isotonic PeriPrep buffer [200 mM Tris–HCl (pH 8.0), 50%(v/v) sucrose] in the presence of 1 mM phenylmethylsulfonyl fluoride (PMSF). After incubation for 15 min at 37°C with mixing, cells were collected by centrifugation at 7,000 rpm for 5 min at 4°C, and then resuspended in 550 μl king2 buffer [100 mM Tris–HCl (pH 7.5), 200 mM NaCl, 1% (v/v) Triton X-100, 0.1% (w/v) Na-deoxycholate, 0.2% (w/v) Brij 58, and 20% (v/v) glycerol]. *In situ* DNase I treatment was performed by adding 50 μl MgCa buffer (100 mM $MgCl_2$ and 50 mM $CaCl_2$), 100 μg RNase A, and 20 units of DNase I (New England Biolabs, Ipswich, Massachusetts, United States) and incubating at 37°C for 15 min. Reactions were stopped by adding 3 ml UT buffer [50 mM HEPES-KOH (pH 7.6), 250 mM NaCl, 0.5% (v/v) Triton X-100, 5 mM imidazole, 5 mM β-mercaptoethanol, 9 M urea, and 1 mM PMSF]. The resultant suspensions were sonicated for 2 min (4 s "on"/10 s "off," 30 times, output 2), and cell debris was removed by centrifugation at 12,000 rpm for 5 min at 4°C. A portion (200 μl) of the resultant supernatant was used to check DNA size by 2% agarose gel electrophoresis. The rest (3.5 ml) was mixed with 100 μl Dynabeads His-tag Isolation and Pulldown (Life Technologies, Carlsbad, California, United States), followed by incubation at 4°C overnight with a gentle rotation. Beads and bound materials were washed seven times with UT buffer, resuspended in 200 μl elution buffer [100 mM Tris–HCl (pH 7.5), 500 mM imidazole, 1% (w/v) SDS, and 10 mM dithiothreitol]. Proteins were degraded by Proteinase K treatment at 42°C for 2 h, followed by further incubation at 65°C for 6 h for de-crosslinking. After removal of proteins by phenol–chloroform–isoamyl alcohol extraction, DNA was recovered by ethanol precipitation in the presence of glycogen and resuspended in 10 μl nuclease-free water.

The DNA library for next-generation sequencing (NGS) was produced using the NEB Next DNA Sample Prep Reagent kit (New England Biolabs) according to the manufacturer's instructions for "Preparing Samples for Sequencing Genomic DNA" (Illumina). The DNA fragments were then purified using a WIZARD SV Gel and PCR Clean-Up System (Promega), and amplified by 15 cycles of PCR. The sequence of the library was then determined by BioAnalyzer (Agilent Technologies).

Short-read sequencing was performed by BGI (Shenzhen, Guangdong, China) with the paired-end procedure (100 bp × 2) on an Illumina HiSeq 2000 instrument (Illumina, San Diego, California, United States). The fastq files of forward and reverse short-read sequencing for each DNA library were concatenated for read mapping, and IHF-binding regions were detected using the pmapsr program (see following section).

Determination of Highest IHF-Binding Regions and Determination of IHF-Binding Motifs

The regions protected by IHF in the *E. coli* genome, which would be sandwiched between the edges of DNase I digestion corresponding to the 5′ and 3′ ends of the forward and reverse short reads, were identified precisely using the pmapsr program (Chumsakul et al., 2013). The DNase I-digested short fragments were estimated to be 70–110 bp long; because we performed 100 bp Illumina sequencing, the reads frequently included the primer sequences added for DNA library construction (see before section), causing severe mismatches for mapping of the reads to the *E. coli* K-12 MG1655 genome (reference sequence) and decreasing the number of mapped reads. To reduce the number of unmapped reads, we initially mapped reads using the mpsmap program permitting 35 bp mismatch in order to be able to map the reads, including the primer sequences (Chumsakul et al., 2013). We then determined the boundaries of the primer sequences and the homologous sequences to the reference genome in the forward and reverse short reads, which

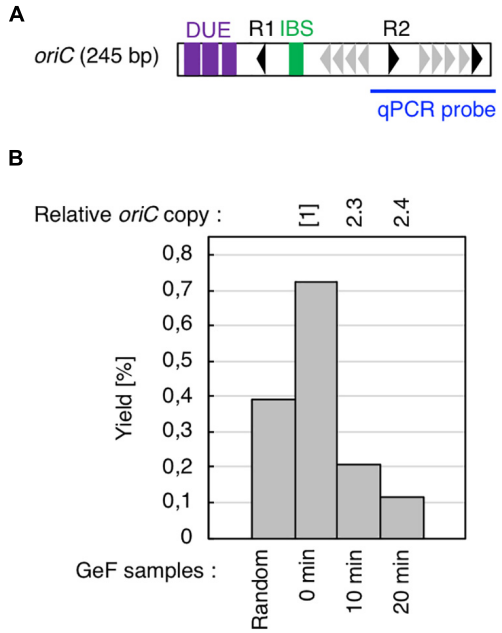


FIGURE 2 | Cell cycle-specific IHF binding at *oriC* in GeF-qPCR. **(A)** Amplified region of *oriC* in qPCR experiments. Symbols used for *oriC* structure are shown as in **Figure 1B**. **(B)** IHF-GeF-qPCR of *oriC*. SH022 (*dnaC2 ihfA-cHis12*) cells growing at 30°C were transferred to 38°C and incubated for 90 min. The cells were then transferred to 30°C (Time 0) and further incubated for 10 or 20 min at 30°C. The relative *oriC* levels before and after Ni-affinity purification were determined using real-time qPCR, and yield was calculated (expressed as %). In addition, relative copies of the chromosomal *oriC* and *TerC* loci in the Input samples were quantified; the relative ratios of *oriC*/*ter* are expressed relative to the ratio at 0 min (determined as 1).

represent the 5′ or 3′ ends of DNase I-digested short fragments. We identified the 70–110 bp regions sandwiched between the 5′ or 3′ ends with high read depths as candidate IHF-binding regions (see details described below).

Detection of highest IHF-binding regions was performed according to the instruction for the pmapsr program for GeF-seq analysis with minor modifications (the original version used in this study; [1], the new version with some minor bug fix; [2]). We did not use the options -pbo, -pbt, or -pbs, which have been adapted to analyze randomly replicating cells and have a genome dosage bias from the replication origin to the termination site in exponentially growing cells. In this study, we used *dnaC2*-based cells to synchronize the replication cycle; consequently, such bias would not be present. We used the following options: -primer to detect primer sequence, -ewf 70 to set the minimal length of the binding regions to 70 bp, and -ewt 110 to set the maximal length of the binding regions to 110 bp; thus, we could detect the 5′ or 3′ ends of 70–110 bp regions protected from DNase I digestion, and -tp with appropriate values to set the threshold values, to select highest IHF-binding regions using the pmapsr program, which have the highest average read depths in IHF-binding regions. We set the threshold value at 10,000 reads for the 0 min dataset, and then adjusted the threshold values of other datasets to reflect the differences of mapped read numbers to the *E. coli* genome among datasets, i.e., 10,000 reads for the 0 min dataset (number of mapped reads, 15,289,849), 19,606 reads for the 10 min dataset (number of mapped reads, 29,977,846), 11,697 reads for the 20 min dataset (number of mapped reads, 17,885,459) and 12,784 reads for the Random dataset (number of mapped reads, 19,546,737). Those high threshold values allowed us to select comparable highest IHF-binding regions in different samples. Finally, we checked IHF-binding peaks and highest IHF-binding regions by visual inspection of IHF-binding profiles visualized on the Integrative Genomics Viewer (IGV; [3]) (Robinson et al., 2017). We removed probable artifacts of highest IHF-binding regions, i.e., contamination of *rRNA* operons including *rrnC* (**Figure 3A**), which occur frequently in chromatin immunoprecipitation (Waldminghaus and Skarstad, 2010).

To visualize the IHF-binding profiles and highest IHF-binding regions estimated by pmapsr on IGV, which was used for visual inspection of false IHF-binding regions (described above) and to prepare Figures (i.e., **Figure 3**), we independently mapped the short reads onto the reference genome using Bowtie 2 to prepare sorted BAM files (Langmead and Salzberg, 2012), the format read by IGV. Bowtie 2 was used with the default settings using Illumina short reads with the primer sequence removed by the cutadapt program (Martin, 2011). Although some of highest IHF-binding regions overlapped and were consequently included in one IHF-binding peak, we used all of highest IHF-binding peaks to estimate the IHF-binding consensus sequence in each dataset. The consensus sequences in the highest IBS were estimated using the MEME suite with default settings (Bailey et al., 2009).

Quantitative PCR

Quantitative PCR (qPCR) experiments were performed as previously described (Kasho and Katayama, 2013; Kasho et al., 2014; Inoue et al., 2016). The levels of *oriC* and *ylcC* were quantified by real-time qPCR using SYBR Premix Ex Taq II (Perfect Real Time; Takara Bio) and primers ORI_1 and KWoriCRev for *oriC*; RTYLCC-L and RTYLCC-R for *ylcC* in SH022 or KYA018 (**Figure 2B**); ilvG-U and ilvG-L for ilvG, rhlB-U and rhlB-L for rhlB, glnH-U and glnH-L for glnH, osmY-U and osmY-L for osmY in KYA018 (**Supplementary Figure 7**); RTNoriC-U2 and RTNoriC-L2 for native *oriC*; RTTERoriC-L and RTTERoriC-L for TER-*oriC*; and RTYLCC-L and RTYLCC-R for *ylcC* in KX237 or KX238 (**Figures 5B,C**).

Chromatin Affinity Precipitation

Chromatin affinity precipitation (ChAP) experiments were performed according to a previously described method (Kasho et al., 2014; Inoue et al., 2016).

[1] https://github.com/NGS-maps/gefseq01

[2] https://github.com/NGS-maps/gefseq011

[3] https://software.broadinstitute.org/software/igv/

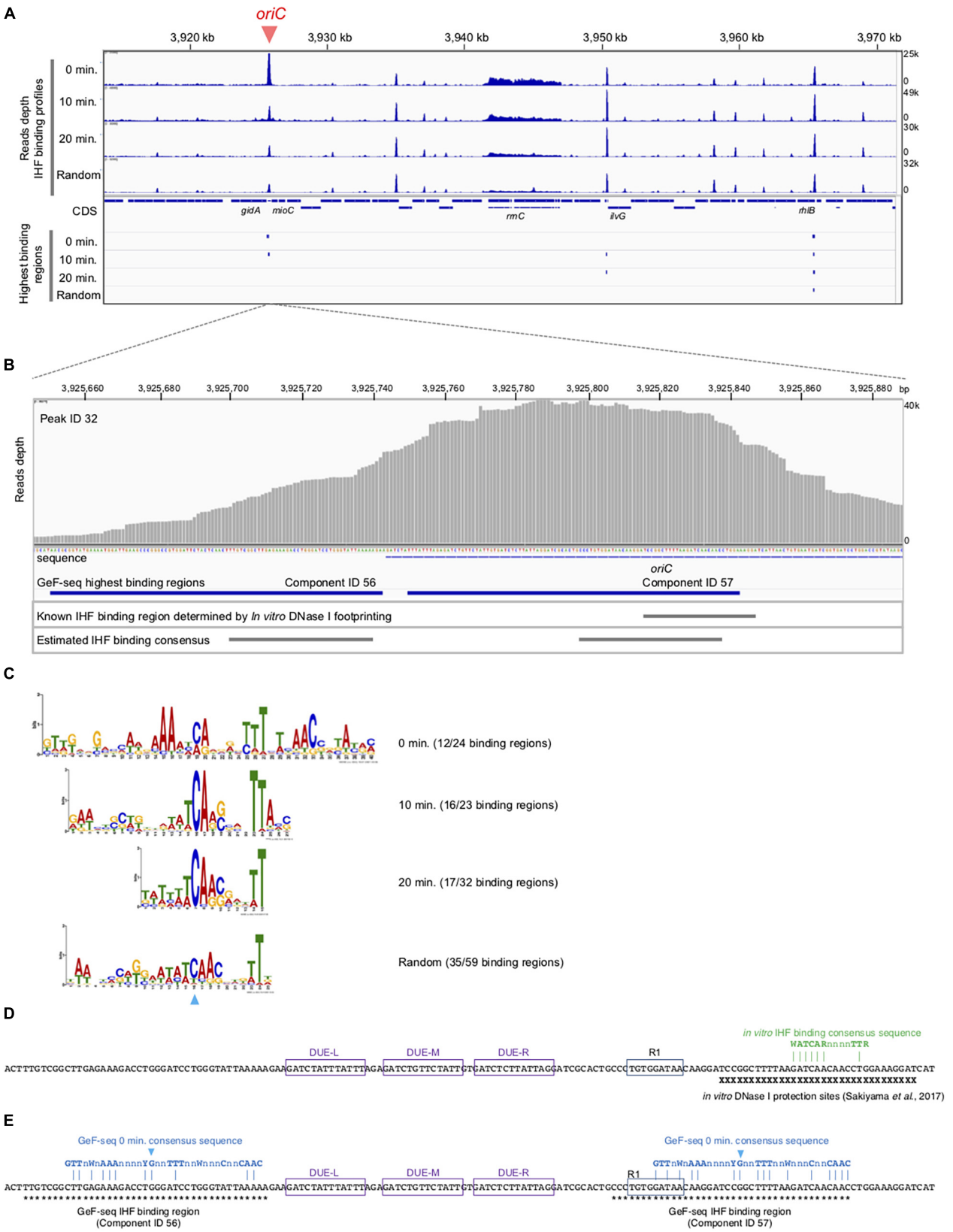


FIGURE 3 | Cell cycle-specific IHF binding at *oriC* in GeF-seq. **(A)** The IHF-binding profile in the region including 3,920–3,970 kb part in the *E. coli* K-12 genome. *(Continued)*

FIGURE 3 | Continued
Upper panels indicate the IHF-binding profiles in the 0 min, 10 min, 20 min, and Random samples. Read depth indicates the strength of IHF binding with the *E. coli* genome. To show each GeF-seq result, we altered the vertical scale in each dataset to allow direct comparison of peak height in each dataset shown in this figure. The vertical scale was expanded 1.96-fold (for the 10 min dataset), 1.2-fold (20 min), or 1.28-fold (Random) relative to the 0 min dataset, indicated at the right of the panels, because the mapped read number differed among samples: 15,289,849, 29,977,846, 17,885,459, and 19,546,737 in 0 min, 10 min, 20 min, and Random datasets, respectively. Therefore, this adjustment of the vertical scale makes it possible to directly compare the height of the IHF binding peaks visualized in figures without concerns about false differences due to differences in mapped read numbers. The threshold values are indicated as blue broken lines: 10,000, 19,606, 11,697, and 12,784 in 0 min, 10 min, 20 min, and Random datasets, respectively. Lower panels indicate highest IHF binding regions (Component IDs 56, 57, 58, and 59 in **Supplementary Table 1**). As shown here, the IBS in lower IHF-binding peaks has not been identified, although the lower-binding peaks may also have the IHF-binding consensus sequences and specifically interact with IHF. **(B)** Extension of **Figure 3A** at the *oriC* locus. Upper panel indicates read depth. At the bottom of this figure, the DNA sequences of the region and the location of the *oriC* locus are indicated. The lower panels indicate the locations of highest IHF-binding regions in the 0 min dataset (Components ID 56, 57 in **Supplementary Table 1**), the IBS determined by *in vitro* DNase I footprinting (Sakiyama et al., 2017), and the estimated IHF-binding consensus sequences based on the GeF-seq result for the 0 min dataset (highest IHF-binding regions). **(C)** Logos indicating consensus sequences estimated from highest IHF-binding regions in each dataset. On the right side of the Logos, datasets and the numbers of highest IHF-binding regions used to compute the Logos are indicated. The numbers indicate the proportion of binding regions including the consensus sequence (removing overlapping) among highest IHF-binding regions used for this analysis in each dataset (for instance, in the 0 min dataset, 12 highest IHF-binding regions include the consensus sequence indicated by the Logo, whereas a total of 24 highest-binding regions were detected. The sky-blue arrowhead indicates the position of the "C" residue at position 19 (0 min), which is highly conserved in all IBSs predicted in this study. **(D)** Known IBSs determined by *in vitro* DNase I footprinting. X indicates sequences in the region protected by IHF from DNase I digestion (Sakiyama et al., 2017). Three DUE elements (L, M, and R) and DnaA box R1 are indicated by purple or blue boxes. IHF binding consensus sequence previously determined by *in vitro* experiments is shown by green characters (Swinger and Rice, 2004). Dam-dependent methylation sequence GATC are labeled as yellow background. **(E)** The locations of the 0 min consensus sequences (blue characters) in highest IHF binding regions, ID 56 and 57. As shown in *panel C*, position for the conserve "C" residue is also indicated with the same sky-blue arrowhead. Asterisks indicate sequences in the indicated regions.

RESULTS

Specific IHF Binding to the *oriC* Locus Before Replication Initiation

Previous studies have identified or predicted thousands of IBS on the *E. coli* chromosome (Grainger et al., 2006; Prieto et al., 2012), and we recently revealed that IHF binding is dynamically regulated at two intergenic sites, *datA* and *DARS2* (Kasho and Katayama, 2013; Kasho et al., 2014), in addition to *oriC*. These observations suggest that the specific time of IHF binding is crucial for regulation of DNA replication and progression of the cell cycle. In this study, to identify the genome-wide regulation of IHF binding, we applied the GeF-seq method to temperature-sensitive *dnaC2* cells, in which it is possible to synchronize the replication cycle. DnaC is the helicase loader; in the *dnaC2* mutant, replication initiation at *oriC* is specifically inhibited at high temperatures (38–42°C) and thus we can synchronize the cell cycle just before replication initiation (0 min). By decreasing the temperature to low temperature (30°C), DnaC is immediately activated to concordantly initiate replication within 5 min in the cells. Previous studies have identified cell cycle-dependent expression of the genes such as *dnaA*, *mioC*, and *gidA* using *dnaC2* cell-based synchronization, and thus this well-established method should be the most suitable for the purpose in this study. By combining DNase I-dependent DNA cleavage, ChAP (a modified ChIP), and NGS, the GeF-seq method identifies protein-bound sites throughout the genome at base-pair resolution (Chumsakul et al., 2013).

Before mapping by GeF-seq, we confirmed by GeF-qPCR that IHF binding was regulated at the *oriC* locus in the *dnaC2* mutant, i.e., IHF stably bound to the left part of *oriC* before initiation (0 min; **Figure 2**) and dissociated after initiation (10 or 20 min; **Figure 2**), consistent with previous studies (Ryan et al., 2002; Kasho and Katayama, 2013; Kasho et al., 2014).

Next, using the same samples as in **Figure 2**, we performed genome-wide mapping of IHF binding at specific cell cycle stages using GeF-seq (**Figure 3**). Unlike standard ChIP-seq, GeF-seq includes the DNase I digestion in the DNA fragmentation process instead of sonication in ChIP-seq. As a result, protein-bound DNA fragments purified in the GeF-seq procedure are shorter than those in standard ChIP-seq and the binding peaks detected by GeF-seq become sharper than those detected by standard ChIP-seq (Chumsakul et al., 2013). In addition, we have the unique program for the GeF-seq analysis, pmapsr which was developed to accurately detect protein-binding regions on genomic DNA using GeF-seq datasets (**Supplementary Table 1**; Chumsakul et al., 2013). Pmapsr is the program to detect the regions sandwiched by the 5′ and 3′ ends of reads sequenced by Illumina sequencing, which are possible to represent the "edge" of the regions protected from the DNase I digestion by protein binding. Previously we succeeded to make the map of the genome wide footprinting of AbrB, the *Bacillus subtilis* global transcriptional regulator, by GeF-seq and the pmapsr analysis with similar resolution of *in vitro* DNase I footprinting and detected weak binding consensus sequence (Chumsakul et al., 2013). By this method, the presence of multiple binding regions (components) is possible to be estimated in one large peak. We, therefore, discriminate the "binding peak (peak)" and "binding region (component)."

GeF-seq results of the synchronized *dnaC2* cells showed a large peak at *oriC* in the 0 min data set, in which two possible binding regions (components) were estimated by pmapsr (**Figure 3A**). Those binding regions might indicate the higher-order complex formation by multiple DNA binding proteins and DNA (for instance, DnaA-IHF complexes and higher-order structures made by binding of multiple IHF molecules). In addition, our GeF-seq results included the representative IHF binding sites determined in previous *in vitro* DNase I footprinting experiments such as those at transcriptional promoters of *ilvG*

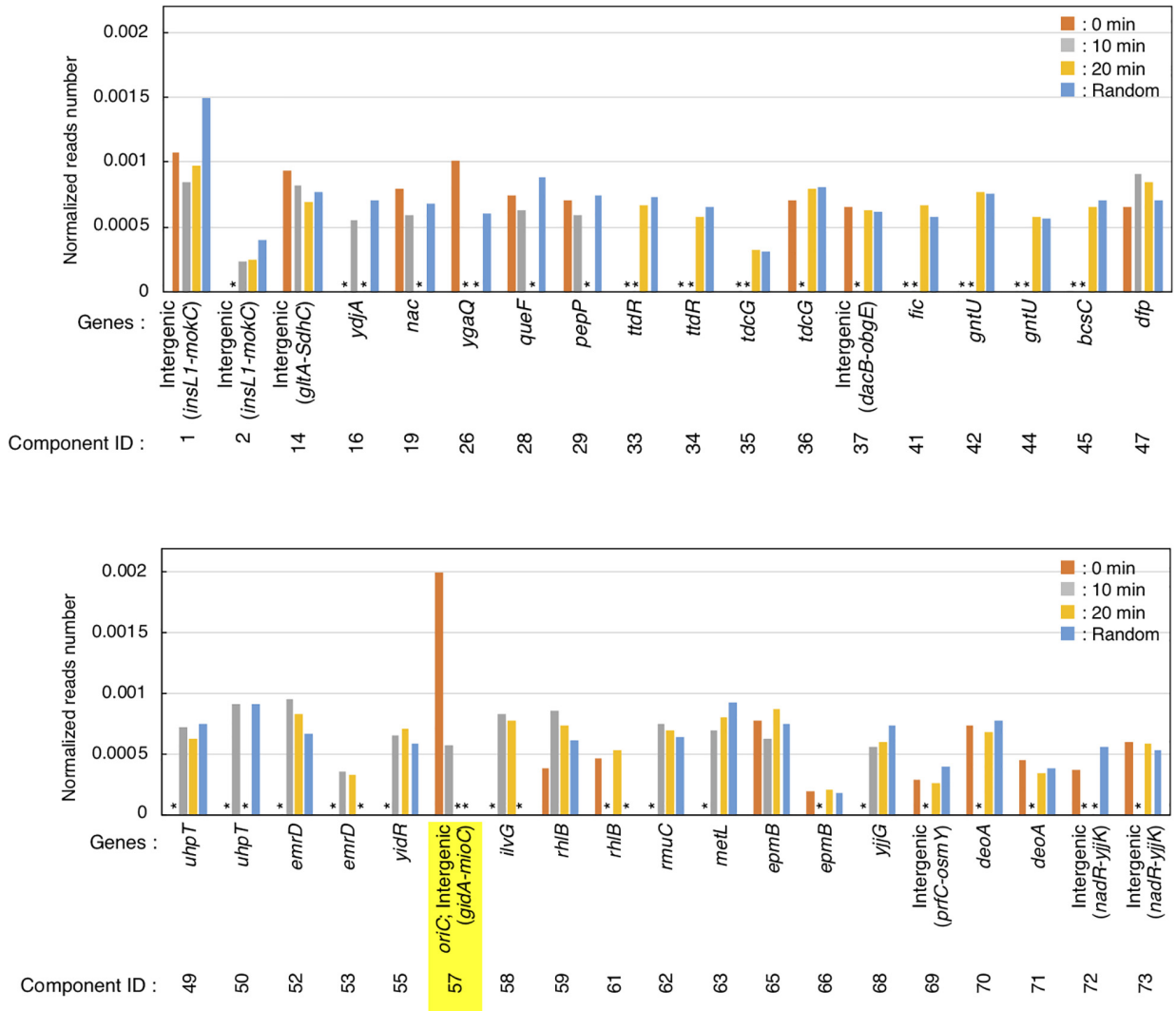


FIGURE 4 | Preferential IHF binding at *oriC* at replication pre-initiation stage. Highest IHF-binding peaks at each timepoint are listed as bar charts. Y-axis indicates the normalized reads number which means the relative read numbers of each component divided by the total mapped reads numbers of the reference genome at each timepoint. Stars (*) indicate "lower than the threshold values," which means that the component was not identified as highest IHF binding peaks at each timepoints.

or *glnH* (**Supplementary Figures 1**, **3**; Tsui and Freundlich, 1988; Claverie-Martin and Magasanik, 1991), which supports that our GeF-seq experiments were qualified enough to determine the genomic IHF binding regions. As shown in **Figure 3**, the IHF-binding peak at *oriC* was prominent relative to all other IHF-binding peaks in 0 min sample (**Figure 3A**). To determine how specifically IHF binds to *oriC* before replication initiation, we selected highest IHF-binding peaks and binding regions, and compared the read numbers of IHF binding among those regions. Highest IHF-binding peaks (Peak ID 1-42) and binding regions (Component ID 1-73) were selected by the pmapsr program using high threshold values (see section "Materials and Methods"). Peak ID 32 (Component ID 57, genomic positions 3,925,749 to 3,925,842), highest IHF-binding peak in the 0 min dataset, overlapped with the *oriC* region (**Figures 3A**, **4** and **Supplementary Table 1**), which had the highest read depth in all of highest IHF-binding peaks selected in the 0 min dataset. The average read depth of the *oriC* locus was 30,601.5, whereas that of Peak ID 1 (Component ID 1, the genomic positions of 16,520 to 16,606), the second highest IHF-binding peak (**Figure 4**, **Supplementary Figure 6**, and **Supplementary Table 1**), was 16,341.3. By contrast, the same *oriC* locus in the 10 min, 20 min, and Random datasets did not have the highest read depth relative to other IHF-binding regions (**Figures 3A**, **4** and **Supplementary Table 1**), consistent with the data shown in **Figure 2B** and previous studies that IHF binding at *oriC* locus is increased specifically at pre-initiation period (Kasho and Katayama, 2013; Kasho et al., 2014). This observation clearly indicates that IHF binds to the *oriC* locus most preferentially before replication initiation; this was a highly distinctive property of IHF binding at the *oriC* locus, at least among highest binding regions (see next section).

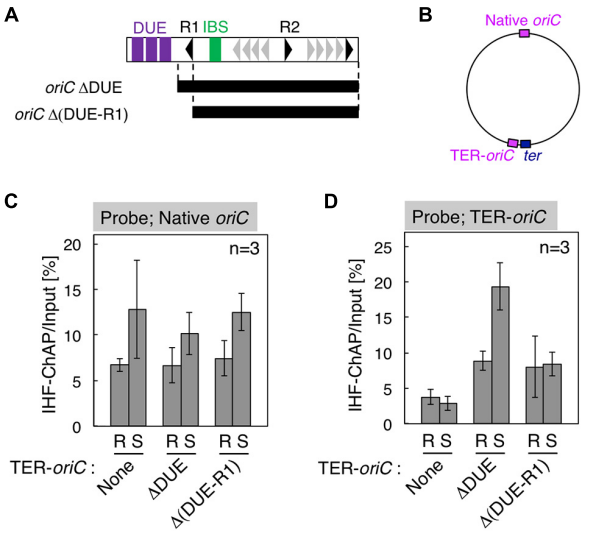


FIGURE 5 | Requirement for DnaA box R1 for *oriC*–IHF binding at the replication initiation period. **(A)** Deletion derivatives of *oriC*. Bars indicate the regions of *oriC*ΔDUE and *oriC*Δ(DUE-R1), which were inserted at the position indicated as TER-*oriC* in the *terC* locus [shown in panel **(B)**] of SH022 (*dnaC2*) strain, resulting in strains KX237 and KX238. Symbols used for the *oriC* structure are shown as in **Figure 1B**. **(C,D)** IHF-ChAP-qPCR with TER-*oriC*-inserted cells. SH022 (*dnaC2*), KX237 (*dnaC2* TER-*oriC*ΔDUE), and KX238 [*dnaC2* TER-*oriC*Δ(DUE-R1)] cells grown at 30°C (R; Random) were transferred to 38°C and incubated for 90 min (S; Synchronized). The relative levels of native *oriC* **(C)** and TER-*oriC* **(D)** before and after Ni-affinity purification were determined using real-time qPCR, and yield was calculated (expressed as %). Error bars represent standard deviation from three independent experiments.

In addition, before initiation, the IHF-binding peak was only prominent at the *oriC* locus, and in contrast, the signals of all other IHF-binding peaks including the *rrnC* locus, one of the rDNA operons, were relatively modest compared with the IHF binding signals at *oriC* (**Figures 2B**, **3A**, 0 min sample). In contrast, after initiation, the IHF-binding peak at *oriC* was relatively modest or discreet, and the peaks at other loci including that at *ilvG* (Peak ID 33; Component ID 58) or the *rhlB* loci (Peak ID 34; Component ID 59 and 60) were mostly comparable or even larger compared with the IHF binding signals at *oriC* (**Figure 3A,** 10 and 20 min samples; **Supplementary Figures 2**, **3**), consistent with the decreased *oriC* copies in qPCR analysis (**Figure 2B**). Notably, the representative IHF binding sites at transcriptional promoter of *osmY* determined in previous *in vitro* DNase I footprinting experiments was included in highest IHF binding regions (**Supplementary Figure 4** and **Supplementary Table 1**; Colland et al., 2000), which supports that our methodology was suitable for identifying IHF binding regions. In addition, IHF binding at representative non-*oriC* loci such as *ilvG*, *rhlB*, *glnH*, and *osmY* regions was demonstrated in ChIP-qPCR experiments using IHF antibody (**Supplementary Figure 7**). As previously shown (Kasho and Katayama, 2013; Kasho et al., 2014), o*riC*-IHF binding was increased only at pre-initiation period (**Supplementary Figure 7A**). In contrast, under the same conditions, IHF binding at those non-*oriC* loci was not largely changed or even decreased in the initiation periods (**Supplementary Figures 7B–E**). These results are basically consistent with the GeF-seq data (see also Discussion).

Prediction of Specific IHF-Binding Consensus Sequences Before Replication Initiation

To determine how IHF preferentially binds to *oriC* before initiation, we investigated the consensus sequences of highest IHF-binding peaks in each dataset using the MEME suite (Bailey et al., 2009). As shown in **Figure 3C**, the consensus sequences in all datasets (0 min: CAnnnnTTT at position 19–27, 10 min: WWTCARSnnnTTA at position 13–25, 20 min: WWCARSnnnTT at position 5–15, and Random: WAWCAACnnnTT at position 13–24, where S is G or C) are similar and contain essential DNA elements with the known consensus sequence WATCARnnnnTTR (W is A or T; n is any nucleotides; R is A or G) (Swinger and Rice, 2004). The 0 min consensus sequence is more enriched in "T/A" at positions 11–16 and 25–29 beside the conserved "CA" at position 19–20 (**Figure 3C**). In addition, unique "GTTG" and "AAC" elements at positions 1–4 and 31–33, respectively, locate beside these "T/A" elements (**Figure 3C**). The highest IHF binding region at *oriC* (Component ID 57) is one of the loci with the best match to the 0 min consensus (**Supplementary Figure 5**), supporting the idea that these specific DNA elements in the 0 min consensus sequence are relevant to the preferential *oriC*–IHF binding before initiation.

Whole Genome Analysis Predicts the Preferential IHF-Binding at *oriC* Locus Before Replication Initiation

As shown in **Figure 3B**, we identified two IHF-protected regions (Component ID 56/57) in the prominent IHF-binding peak at *oriC* (Peak ID 32; **Figure 4**, **Supplementary Figure 6**, and **Supplementary Table 1**). The right protected region (Component ID 57) at *oriC* partly overlaps with a known IHF-binding consensus sequence (**Figure 3E**), which was previously identified as an *oriC* IBS by *in vitro* DNase I footprinting analysis (**Figure 3D**; Sakiyama et al., 2017). This study provides the first direct evidence for *in vivo* IHF binding to the *oriC* IBS in *E. coli* cells. The second protected region (Component ID 56) has weak IHF binding signal. Although no known IBS had been identified in this region, we identified a sequence that matches the new 0 min consensus sequence in Component ID 56 (**Figures 3C,E**). Alternative possibility is that the DnaA–IHF complex bound to the wide area of the *oriC* region, and the protected region would be expanded outside *oriC*. However, the existence of a secondary IBS with unstable IHF binding cannot be ruled out. Taken together, this preferential *oriC*–IHF binding before initiation requires the additional mechanism to the 0 min consensus sequences; i.e., weak IHF binding to the secondary site and the presence of two IBSs at the *oriC* locus may support stable *oriC*–IHF binding at a specific stage of the cell cycle (see below and Discussion).

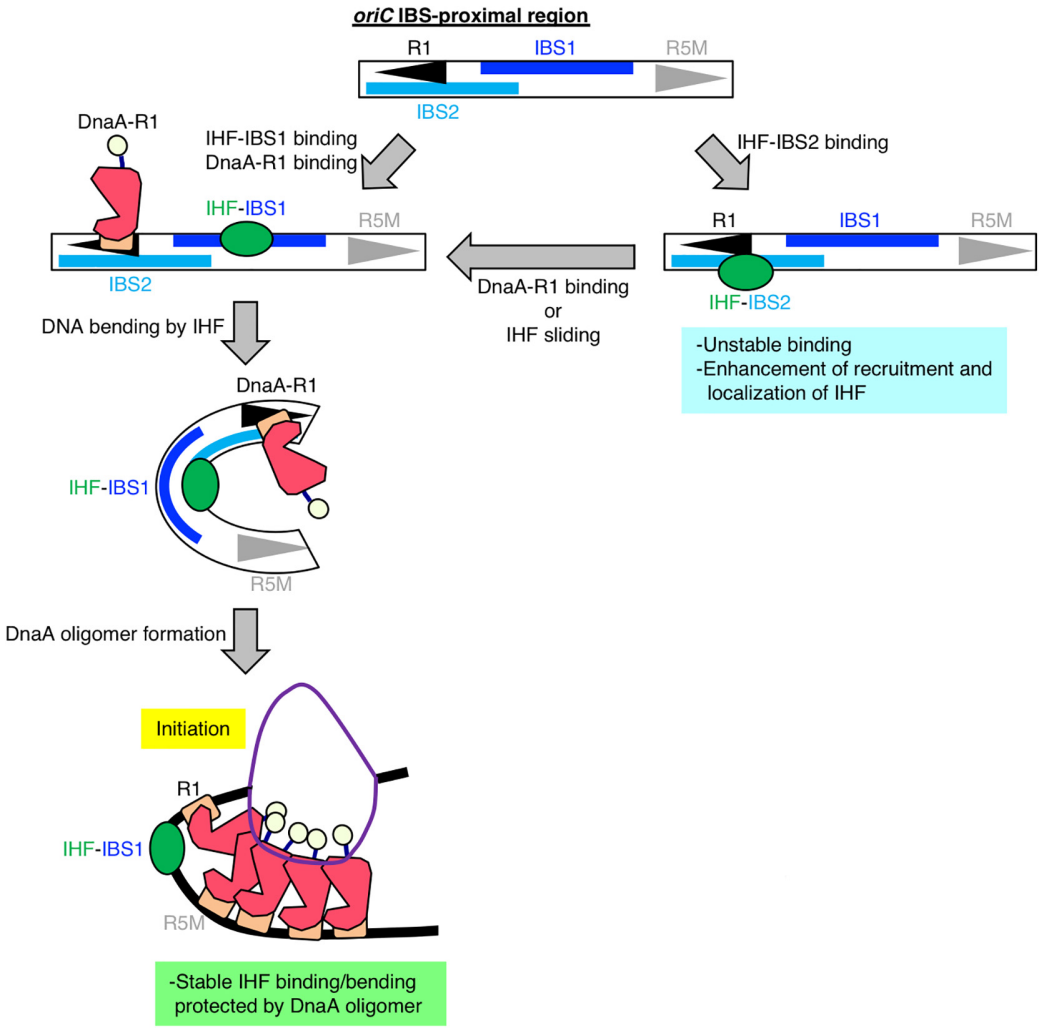


FIGURE 6 | Model of DnaA box R1-mediated stabilization of IHF binding at *oriC*. Our findings suggest that, specifically at the stage of replication initiation, a secondary IBS (termed IBS2 in this study; sky-blue bar) overlaps with DnaA box R1, providing a new possible role for DnaA box R1 (black arrowhead). First, IHF binds to either a known IBS (IBS1; blue bar) or IBS2. When IHF binds to IBS2, DnaA cannot bind to R1; thus we hypothesize that even though *oriC* IBS2–IHF binding is relatively unstable, it plays a supportive role in stabilizing overall IHF binding at *oriC* by preventing free diffusion of IHF dissociated from *oriC* and promoting re-binding to *oriC*. When IHF binds to IBS1 with higher affinity and bends DNA, ATP–DnaA can form higher-order oligomers to initiate replication from *oriC*. In the initiation complex, IHF binding and bending at IBS1 is further stabilized by an ATP–DnaA oligomer, which is formed by interaction between ATP–DnaA molecules bound at R1 and another DnaA box R5 (gray arrowhead).

We did not detect substantial specific signals at the *datA* and *DARS2* loci, although our previous study using ChIP-qPCR indicated that those should appear at 10 and 20 min after initiation. This is probably because of a difference in experimental conditions between ChIP and GeF, e.g., DNase I treatment or NGS sample preparation (see details in Discussion).

A Role for *oriC* DnaA Box R1, but Not Chromosomal Position of *oriC*, in Stable IHF Binding

To examine the regulatory mechanism by which IHF binding at the *oriC* locus is stabilized before initiation, we focus on two specific structural features of *oriC*: (1) overall structure of 1 Mb chromosomal region containing *oriC*, termed Ori-macrodomain (Niki et al., 2000; Valens et al., 2016), and (2) the local structure of *oriC* complexed with ATP–DnaA. First, to examine the requirement of the Ori-macrodomain structure, we constructed genome-edited cells carrying an insertion of DUE-deleted *oriC* sequence at *terC*-proximal intergenic region between *pntA* and *ydgH* genes (TER-*oriC*ΔDUE; **Figures 5A,B**), and analyzed IHF-binding patterns before initiation. These cells have intact *oriC*s at the native position (Native *oriC*), and as expected, insertion of TER-*oriC*ΔDUE had a minimal effect on stabilization of IHF binding at the native *oriC* locus before initiation (**Figure 5C**). Even at the *terC* locus, the signal of IHF binding was dependent on insertion of TER-*oriC*ΔDUE, and it increased before initiation (**Figure 5D**). These observations suggest that

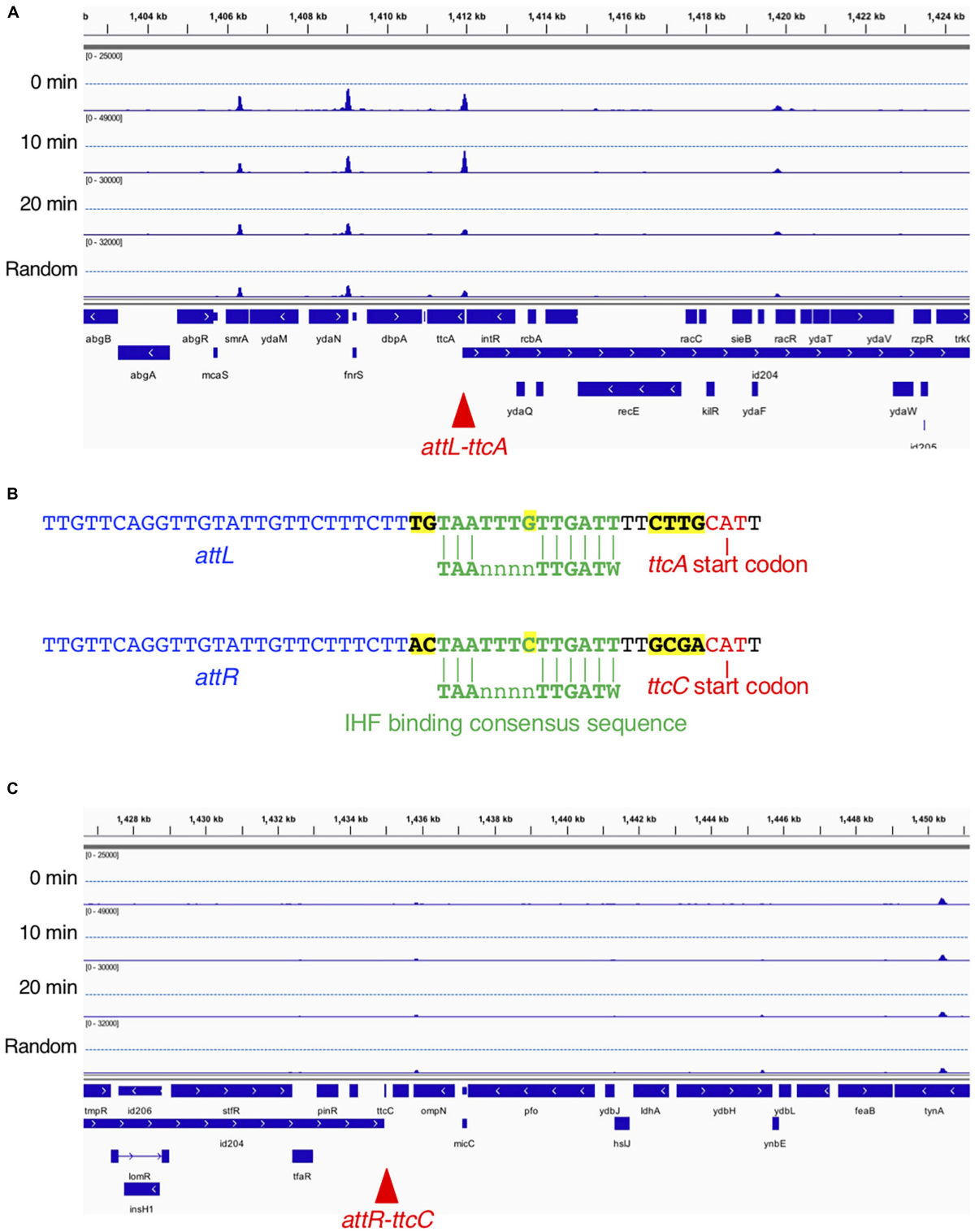


FIGURE 7 | Time-resolved IHF binding at the *ttcA*-Rac prophage locus. **(A)** IHF-binding profiles at the *ttcA* upstream region in the 0 min, 10 min, 20 min, and Random samples. Red arrowhead represents the locus with IHF binding at specific timing (10 min) after replication initiation, which corresponds to Rac prophage attachment site *attL*. **(B)** Sequence comparison of Rac prophage attachment sites and IHF-binding consensus sequences in *ttcA* (top) or *ttcC* genes (bottom), and characters with yellow highlight is sequences different in *ttcA* or *ttcC* genes. The start codons of the *ttcA* or *ttcC* genes are shown by red characters. IHF-binding consensus (TAAnnnnTTGATW, where W means T or A) is shown by green characters. Rac prophage attachment sites *attL* or *attR* are shown by blue characters. **(C)** IHF-binding profiles at the *ttcC* upstream region in the 0 min, 10 min, 20 min, and Random samples. Red arrowhead represents the locus with Rac prophage attachment site *attR*.

timely stabilization of *oriC*–IHF binding is independent of Ori-macrodomain structure. Interestingly, IHF binding signal at TER-*oriC*ΔDUE might be higher than that at Native *oriC*, which could suggest two possibilities that DNA structure at Ter-macrodomain could be preferable for *oriC*-IHF binding, or that *oriC* DUE could play an inhibitory role for IHF binding by unidentified mechanism.

Second, to assess the requirement of ATP–DnaA oligomers on *oriC* for stabilizing IHF binding before initiation, we further constructed genome-edited cells carrying an insertion of an *oriC* derivative lacking a DUE-R1 region at the *TerC* locus [TER-*oriC*Δ(DUE-R1); **Figure 5A**]. Deletion of DnaA box R1 should drastically change the structure of ATP–DnaA oligomers on *oriC*. As with TER-*oriC*ΔDUE, insertion of TER-*oriC*Δ(DUE-R1) had little effect on the IHF-binding pattern at the native *oriC* locus (**Figure 5C**); at the *terC* locus, the signal of IHF binding was observed in a manner dependent on insertion of TER-*oriC*Δ(DUE-R1; **Figure 5D**). Notably, in contrast to the TER-*oriC*ΔDUE, the signal of IHF binding at TER-*oriC*Δ(DUE-R1) was not increased even before initiation (**Figure 5D**). These results indicate that DnaA box R1 is required for stabilization of IHF binding at *oriC* at the stage of replication initiation. The ATP–DnaA molecule bound at DnaA box R1 is predicted to interact with the one bound at DnaA box R5 *via* DNA bending at the primary IBS (IBS1; **Figure 6**) induced by IHF binding (**Figures 1B,C**). The ATP–DnaA oligomer formed on *oriC* might stabilize DNA bending by IHF and thus form a rigid *oriC*–IHF complex. Another possibility is that the secondary IBS (IBS2; **Figure 6**), which partly overlaps with DnaA box R1, might play a supporting role in stabilizing IHF binding (see details in Discussion).

An Unique Locus Binds IHF at a Specific Time

From the genome-wide mapping of IBS by GeF-seq, we identified a novel cell cycle-dependent IHF binding in the *ttcA* gene region (**Figure 7**). This locus has the unique feature that IHF binds temporarily 10 min after initiation but dissociates 20 min after initiation and before initiation, as observed for all other IBSs except for *oriC* (**Figure 7A**). The IHF-binding peak corresponded to the IHF-binding consensus near the start codon of the *ttcA* gene, which encodes a tRNA-thioltransferase (**Figure 7B**; Liu et al., 2015). In addition, this locus is adjacent to the Rac prophage excision site *attL* (**Figure 7B**). A region including the other Rac prophage excision site *attR*, located in the *ttcC* gene, has a similar IHF-binding consensus but does not have IHF-binding peaks (**Figures 7B,C**). The IHF-binding region in the *ttcA* gene could play a major regulatory role in Rac prophage excision; however, this hypothesis requires further experimental testing.

DISCUSSION

In this study, we sought to characterize the cell cycle-dependent regulation of the genomic IHF-binding pattern using GeF-seq at base-pair resolution. The results provided evidence of unique and specific regulation of strong *oriC*–IHF binding at the replication pre-initiation stage. This suggests that at that stage, IHF preferentially binds to *oriC* rather than other sites with affinity for IHF. The mechanisms responsible for this preference remain unknown (see below). Previous studies have attempted genome-wide analysis of IHF binding by antibody-based immunoprecipitation, but no previous study detected IHF binding at the *oriC* locus (Grainger et al., 2006; Prieto et al., 2012). This is consistent with our observations that the prominent binding peak at *oriC* was specific for the initiation stage; in other stages, as well as in random cultures, the IHF-binding peaks at *oriC* were smaller than other evident binding peaks. In addition, formation of bulky *oriC*–DnaA-IHF super-complexes may inhibit the interaction between IHF and the antibodies used in previous studies. By contrast, our method, which is based on His-tag affinity purification, improves the yield of the IHF complex with bulky initiation complexes (Kasho et al., 2014; Inoue et al., 2016), and successfully detected *oriC*–IHF binding in a comprehensive analysis of genomic IHF binding for the first time.

However, our analyses also have limitations. Some of the known IHF-binding loci, such as *datA* and *DARS2*, were not substantially detected, which could be explained in three ways. First, compared with IHF binding to *oriC*, that to *datA* and *DARS2* could be more dynamic, with a rapid binding/dissociation equilibrium to ensure timely interaction during the cell cycle, making it difficult to detect the binding using the present methods. Second, in regard to cell sample preparation, our GeF-seq experiments were performed using cells at early exponential phase to determine the effect of the cell cycle on actively growing and replicating cells, whereas previous studies were performed using cells at mid/late exponential or stationary phases (Grainger et al., 2006; Prieto et al., 2012). Expression of the IHF protein is 3–6-fold higher in stationary phase than in exponential phase (Ali Azam et al., 1999), which might destabilize IHF binding to *datA* and *DARS2* in early exponential phase. Third, for the DNA samples used for NGS, protein-bound DNA was treated with DNase I to precisely determine the protein-binding site; however, previous studies suggested that DNase I may have some sequence specificities that would prevent detection of some chromosomal loci (Herrera and Chaires, 1994; Koohy et al., 2013), e.g., input read depth at the *DARS2* locus was very low relative to the surrounding regions. In addition, results obtained by ChIP-qPCR experiments were overall consistent with the GeF-seq results (**Supplementary Figure 7**), in certain cases relative IHF binding levels at 10 or 20 min after replication initiation were moderately different between GeF-seq and ChIP-qPCR data (**Figure 3A** and **Supplementary Figures 7B,C**). This might be caused from changes in the local copy number of genomic DNA during replication. In GeF-seq data, we cannot normalize IHF signals according to this change. Also, the IHF binding at *glnH* locus was lower than the threshold in this experiment, whereas it was detected by *in vitro* DNase I footprinting (Claverie-Martin and Magasanik, 1991). This difference might be because we set the high threshold value to identify the highest IHF binding regions which was to enable the identification of the timepoint-specific IHF binding in GeF-seq results. Therefore, the relatively lower binding peaks may also represent specific

IHF binding regions. Consistently, in the previous GeF-seq experiment, many of the lower binding peaks were specific and had the consensus sequence of the *B. subtilis* transcriptional regular AbrB (Chumsakul et al., 2013). Another possible reason for the difference in the IHF binding profiles between this and previous experiments is that in experimental conditions.

Our results provide a new model for the specific regulation of local *oriC*–IHF binding at the replication initiation stage. We successfully reconstituted cell cycle-dependent IHF binding/dissociation by introducing *oriC*ΔDUE at the *terC* locus and performing *oriC* mutation analysis. The results suggested that *oriC* DnaA box R1, but not *oriC* location, is essential for stable *oriC*–IHF binding (**Figure 5C**). In initiation complexes, ATP–DnaA forms specific tight and bulky oligomers, which could stabilize the bending of DNA bound to IHF at the IBS-proximal region (**Figure 1C**), implying that ATP–DnaA oligomers might prevent IHF from freely dissociating from the *oriC* region during DNA bending promoted by IHF binding (the process is discussed below). In addition, it should be noted that in the CRISPR–Cas system of *E. coli*, IHF binding and the resultant DNA bending promotes DNA binding by the Cas1–Cas2 integrase complex, and Cas1 directly interacts with IHF (Wright et al., 2017). Similarly, IHF could directly interact with DnaA in the initiation complex. Consistent with this, HU, a structural homolog of IHF, interacts directly with DnaA (Chodavarapu et al., 2008). In addition, the possibility of a DnaA–IHF interaction could explain the appearance of an expanded protection region including DnaA box R1 (Component ID 57 in **Figures 3B,E**). Another possibility is that specific subcellular localization of IHF by liquid–liquid phase separation (LLPS) could occur before and at the initiation stage, followed by dramatic changes in IHF dynamics after initiation. Given that intrinsically disordered regions (IDRs) of proteins are suggested to stimulate LLPS (Borcherds et al., 2020), this hypothesis is consistent with predictions for IDRs of IHF in the MobiDB database: only α-subunits of proteobacterial IHF (such as *Salmonella typhimurium*, *Myxococcus xanthus*, *Vibrio cholerae*, etc.), but not β-subunits, have IDRs (Potenza et al., 2015). For example, amino acids 49–73 of *E. coli* IHF-α, which forms β-sheets and interacts directly with DNA (Rice et al., 1996), are predicted to be an IDR; however, further analysis is required to prove these possibilities.

Our comprehensive sequence analysis of the IHF-binding consensus suggested that before replication initiation, IHF has a specific consensus sequence that consists of conserved elements "CA" at positions 19–20 and "TT" at positions 26–27, as well as unique elements "GTTG" and "AAC" at positions 1–4 and 31–33, respectively (**Figure 3C**). Previous studies have tried to address the relationship between the IBS and its binding affinity; however, the requirement of the surrounding AT-rich elements in the IHF-binding consensus remains unclear (Aeling et al., 2006). This study raises the possibility that at a specific cell cycle stage, other NAPs, supercoiling state, or transcriptional profile might change the higher-order genomic structure and strengthen the requirement for surrounding AT-rich elements of the IHF-binding consensus. IHF binding to this new IHF consensus sequence might regulate expression of cell cycle-dependent genes before initiation. DnaA is known to regulate cell cycle-dependent genes such as *nrdAB* and *mioC* (Gon et al., 2006; Hansen et al., 2007). Also, SeqA protein, which specifically recognizes hemi-methylated GATC sequences after initiation, represses transcription of *dnaA* and *gidA* genes after initiation (Bogan and Helmstetter, 1997). Thus in addition to DnaA and SeqA, IHF could be a novel regulator for cell cycle-dependent gene expression.

A previous kinetic study suggested that DNA binding and bending introduced by IHF occur in a stepwise manner, and that IHF–DNA binding is a rapid process, whereas IHF-induced DNA bending is much slower and therefore rate-limiting (Sugimura and Crothers, 2006). This kinetic model and our sequence analysis of IHF-binding consensus at *oriC* suggested that *oriC* has a secondary IBS (IBS2) at a region overlapping with R1, which could also explain the mechanism by which *oriC*–IHF binding is stabilized at the replication initiation stage: first, IHF could bind to either the primary IBS (IBS1) or IBS2, and then if DnaA box R1 is occupied by DnaA, IHF can no longer bind to IBS2; in this case, however, IHF binding to IBS1 would be stabilized by *oriC*-ATP–DnaA complexes as described (**Figure 6**). A similar mode of discrimination of NAPs–DNA binding was proposed to occur in the regulation of *E. coli ftnA* transcription by H-NS and Fur, i.e., H-NS dimers cooperatively bind to the *ftnA* promoter to repress transcription, and Fur expression induces switching from H-NS to Fur to activate *ftnA* transcription (Nandal et al., 2010). Also, as with other DNA binding/bending proteins such as human mitochondrial transcription factor A (TFAM) (Farge et al., 2012), IHF could be mobile and slide along DNA from IBS2 to IBS1. Alternatively, IBS2 could be a reservoir of IHF, i.e., if an IHF molecule bound at the primary site IBS1 is accidentally dissociated, it could re-bind to IBS2, preventing free diffusion and thereby stabilizing overall *oriC*–IHF binding. These features of IHF-binding dynamics and IBSs in *oriC* imply that *oriC* structure is designated in a sophisticated manner to ensure IHF binding for regulation of replication initiation.

IHF dissociates from *oriC* within 5–10 min after initiation (**Figures 2**, **3**; Kasho and Katayama, 2013; Kasho et al., 2014), although the mechanisms remain unclear. The passage of the replication machinery, which is loaded onto the unwound *oriC* region, or unknown factors could contribute to IHF dissociation. In addition, *oriC* contains eleven copies of Dam-dependent methylation sequence GATC and SeqA-dependent DnaA dissociation at post-initiation period can be considered as an analogous mechanism (Nievera et al., 2006; Katayama et al., 2010). Notably, GATC sequence is also present in IHF binding regions identified in this study (**Figures 3D,E**), suggesting that SeqA sequestrates DnaA and IHF from *oriC* to inhibit over-initiations. Also, after initiation, ATP–DnaA is converted to ADP–DnaA, causing ADP–DnaA to become predominant in cells (Kurokawa et al., 1999). ADP–DnaA molecules form unstable oligomers on *oriC*, suggesting that dissociation of DnaA molecules from *oriC* might impede stable interaction of IHF with *oriC in vivo*.

In addition, we identified a novel IHF-binding region at the Rac prophage excision site *attL* (**Figure 7**). In the CRISPR–Cas

system, IHF interacts directly with Cas1 integrase and promotes the interaction of Cas1–Cas2 complexes with DNA, thereby suppressing off-target integration by Cas1–Cas2 (Wright et al., 2017). Also, IHF is a stimulatory factor for both integration and excision of bacteriophage lambda (Segall and Nash, 1996; MacWilliams et al., 1997). In *E. coli*, Rac prophage plays a role in cellular stress responses and uses a lambda phage-like mechanism for integration and excision by IntR integrase (Liu et al., 2015). Interestingly, the function of Rac prophage is stimulated under specific environmental conditions such as high nutrition or growth in early exponential phase (Liu et al., 2015), suggesting that similar significant role of IHF might be present in Rac prophage excision, and that IHF binding could stimulate the excision at Rac prophage excision site *attL* at a specific cell cycle stage, which remains to be determined. The significance of these cell cycle-specific interactions remains unknown.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. Genome sequencing data were deposited in the DDBJ Sequence Read Archive (DRA) under BioProject number PRJDB11576 and accession numbers DRA012298 (0 min pulldown), DRA012299 (10 min pulldown), DRA012300 (20 min pulldown), DRA012301 (Random pulldown), DRA012302 (0 min Input), DRA012303 (10 min Input), DRA012304 (20 min Input), and DRA012305 (Random Input).

AUTHOR CONTRIBUTIONS

KK and TO performed the experiments. TO, OC, KN, and KF analyzed the IHF binding profile from GeF-seq data. KK, TO, and TK wrote the manuscript. All authors conceived the experiments and analyzed the data.

FUNDING

This study was supported by Grant-in-aid for Scientific Research, JSPS KAKENHI Grant numbers JP21K19233, JP20H03212, JP17H03656, and JP15K18479, and was partly supported by Institute for Fermentation, Osaka (G-2020-2-103 to TO).

ACKNOWLEDGMENTS

We would like to thank Sjoerd Wanrooij for understanding in preparation of this manuscript when KK was a postdocotoral researcher in his laboratory of Umeå University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.697712/full#supplementary-material

REFERENCES

Aeling, K. A., Opel, M. L., Steffen, N. R., Tretyachenko-Ladokhina, V., Hatfield, G. W., Lathrop, R. H., et al. (2006). Indirect recognition in sequence-specific DNA binding by *Escherichia coli* integration host factor: the role of DNA deformation energy. *J. Biol. Chem.* 281, 39236–39248. doi: 10.1074/jbc.M606363200

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. doi: 10.1128/JB.181.20.6361-6370.1999

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Bogan, J. A., and Helmstetter, C. E. (1997). DNA sequestration and transcription in the *oriC* region of *Escherichia coli*. *Mol. Microbiol.* 26, 889–896. doi: 10.1046/j.1365-2958.1997.6221989.x

Borcherds, W., Bremer, A., Borgia, M. B., and Mittag, T. (2020). How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation? *Curr. Opin. Struct. Biol.* 67, 41–50. doi: 10.1016/j.sbi.2020.09.004

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

Chumsakul, O., Nakamura, K., Kurata, T., Sakamoto, T., Hobman, J. L., Ogasawara, N., et al. (2013). High-resolution mapping of in vivo genomic transcription factor binding sites using in situ DNase I footprinting and ChIP-seq. *DNA Res.* 20, 325–337. doi: 10.1093/dnares/dst013

Claverie-Martin, F., and Magasanik, B. (1991). Role of integration host factor in the regulation of the *glnHp2* promoter of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 88, 1631–1635. doi: 10.1073/pnas.88.5.1631

Colland, F., Barth, M., Hengge-Aronis, R., and Kolb, A. (2000). Sigma factor selectivity of *Escherichia coli* RNA polymerase: role for CRP, IHF and Lrp transcription factors. *EMBO J.* 19, 3028–3037. doi: 10.1093/emboj/19.12.3028

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

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

Farge, G., Laurens, N., Broekmans, O. D., Van Den Wildenberg, S. M. J. L., Dekker, L. C. M., Gaspari, M., et al. (2012). Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. *Nat. Commun.* 3:1013. doi: 10.1038/ncomms2001

Fujimitsu, K., Senriuchi, T., and Katayama, T. (2009). Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* 23, 1221–1233. doi: 10.1101/gad.1775809

Gon, S., Camara, J. E., Klungsøyr, H. K., Crooke, E., Skarstad, K., and Beckwith, J. (2006). A novel regulatory mechanism couples deoxyribonucleotide synthesis and DNA replication in *Escherichia coli*. *EMBO J.* 25, 1137–1147. doi: 10.1038/sj.emboj.7600990

Grainger, D. C., Hurd, D., Goldberg, M. D., and Busby, S. J. W. (2006). Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res.* 34, 4642–4652. doi: 10.1093/nar/gkl542

Hansen, F. G., Christensen, B. B., and Atlung, T. (2007). Sequence characteristics required for cooperative binding and efficient in vivo titration of the replication initiator protein DnaA in *E. coli*. *J. Mol. Biol.* 367, 942–952. doi: 10.1016/j.jmb.2007.01.056

Herrera, J. E., and Chaires, J. B. (1994). Characterization of preferred deoxyribonuclease I cleavage sites. *J. Mol. Biol.* 236, 405–411. doi: 10.1006/jmbi.1994.1152

Inoue, Y., Tanaka, H., Kasho, K., Fujimitsu, K., Oshima, T., and Katayama, T. (2016). Chromosomal location of the DnaA-reactivating sequence *DARS2* is important to regulate timely initiation of DNA replication in *Escherichia coli*. *Genes Cells* 21, 1015–1023. doi: 10.1111/gtc.12395

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

Kasho, K., Tanaka, H., Sakai, R., and Katayama, T. (2017). Cooperative DnaA binding to the negatively supercoiled *datA* locus stimulates DnaA-ATP hydrolysis. *J. Biol. Chem.* 292, 1251–1266. doi: 10.1074/jbc.M116.762815

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

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

Koohy, H., Down, T. A., and Hubbard, T. J. (2013). Chromatin accessibility data sets show bias due to sequence specificity of the DNase I enzyme. *PLoS One* 8:e69853. doi: 10.1371/journal.pone.0069853

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

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Liu, X., Li, Y., Guo, Y., Zeng, Z., Li, B., Wood, T. K., et al. (2015). Physiological function of Rac Prophage during biofilm formation and regulation of Rac excision in *Escherichia coli* K-12. *Sci. Rep.* 5:16074. doi: 10.1038/srep16074

MacWilliams, M., Gumport, R. I., and Gardner, J. F. (1997). Mutational analysis of protein binding sites involved in formation of the Bacteriophage λ *attL* complex. *J. Bacteriol.* 179, 1059–1067. doi: 10.1128/jb.179.4.1059-1067.1997

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Nandal, A., Huggins, C. C. O., Woodhall, M. R., McHugh, J., Rodríguez-Quiñones, F., Quail, M. A., et al. (2010). Induction of the ferritin gene (*ftnA*) of *Escherichia coli* by Fe 2+-Fur is mediated by reversal of H-NS silencing and is RyhB independent. *Mol. Microbiol.* 75, 637–657. doi: 10.1111/j.1365-2958.2009.06977.x

Nievera, C., Torgue, J. J.-C., Grimwade, J. E., and Leonard, A. C. (2006). SeqA blocking of DnaA-*oriC* interactions ensures staged assembly of the *E. coli* pre-RC. *Mol. Cell* 24, 581–592. doi: 10.1016/j.molcel.2006.09.016

Niki, H., Yamaichi, Y., and Hiraga, S. (2000). Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14, 212–223. doi: 10.1101/gad.14.2.212

Noguchi, Y., Sakiyama, Y., Kawakami, H., and Katayama, T. (2015). The Arg fingers of key DnaA protomers are oriented inward within the replication origin *oriC* and stimulate DnaA subcomplexes in the initiation complex. *J. Biol. Chem.* 290, 20295–20312. doi: 10.1074/jbc.M115.662601

Nozaki, S., Yamada, Y., and Ogawa, T. (2009). Initiator titration complex formed at *datA* with the aid of IHF regulates replication timing in *Escherichia coli*. *Genes Cells* 14, 329–341. doi: 10.1111/j.1365-2443.2008.01269.x

Ozaki, S., Noguchi, Y., Hayashi, Y., 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

Potenza, E., Di Domenico, T., Walsh, I., and Tosatto, S. C. E. (2015). MobiDB 2.0: an improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.* 43, D315–D320. doi: 10.1093/nar/gku982

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

Riber, L., Frimodt-Møller, J., Charbon, G., and Løbner-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

Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996). Crystal Structure of an IHF-DNA Complex: a protein-induced DNA U-turn. *Cell* 87, 1295–1306. doi: 10.1016/s0092-8674(00)81824-3

Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A., and Mesirov, J. P. (2017). Variant review with the integrative genomics viewer. *Cancer Res.* 77, e31–e34. doi: 10.1158/0008-5472.CAN-17-0337

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

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H., and Katayama, T. (2017). Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in *Escherichia coli*. *Nucleic Acids Res.* 45, 12354–12373. doi: 10.1093/nar/gkx914

Seah, N. E., Warren, D., Tong, W., Laxmikanthan, G., Van Duyne, G. D., and Landy, A. (2014). Nucleoprotein architectures regulating the directionality of viral integration and excision. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12372–12377. doi: 10.1073/pnas.1413019111

Segall, A. M., and Nash, H. A. (1996). Architectural flexibility in lambda site-specific recombination: three alternate conformations channel the *attL* site into three distinct pathways. *Genes Cells* 1, 453–463. doi: 10.1046/j.1365-2443.1996.d01-254.x

Shimizu, M., Noguchi, Y., Sakiyama, Y., Kawakami, H., Katayama, T., and Takada, S. (2016). Near-atomic structural model for bacterial DNA replication initiation complex and its functional insights. *Proc. Natl. Acad. Sci. U.S.A.* 113, E8021–E8030. doi: 10.1073/pnas.1609649113

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Stella, S., Cascio, D., and Johnson, R. C. (2010). The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes Dev.* 24, 814–826. doi: 10.1101/gad.1900610

Sugimura, S., and Crothers, D. M. (2006). Stepwise binding and bending of DNA by *Escherichia coli* integration host factor. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18510–18514. doi: 10.1073/pnas.0608337103

Sugiyama, R., Kasho, K., Miyoshi, K., Ozaki, S., Kagawa, W., Kurumizaka, H., et al. (2019). A novel mode of DnaA-DnaA interaction promotes ADP dissociation for reactivation of replication initiation activity. *Nucleic Acids Res.* 47, 11209–11224. doi: 10.1093/nar/gkz795

Swinger, K., and Rice, P. (2004). IHF and HU: flexible architects of bent DNA. *Curr. Opin. Struct. Biol.* 14, 28–35. doi: 10.1016/j.sbi.2003.12.003

Toro, E., and Shapiro, L. (2010). Bacterial chromosome organization and segregation. *Cold Spring Harb. Perspect. Biol.* 2:a000349. doi: 10.1101/cshperspect.a000349

Tsui, P., and Freundlich, M. (1988). Integration host factor binds specifically to sites in the *ilvGMEDA* operon in *Escherichia coli*. *J. Mol. Biol.* 203, 817–820. doi: 10.1016/0022-2836(88)90212-4

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

Waldminghaus, T., and Skarstad, K. (2010). ChIP on chip: surprising results are often artifacts. *BMC Genomics* 11:414. doi: 10.1186/1471-2164-11-414

Wright, A. V., Liu, J.-J., Knott, G. J., Doxzen, K. W., Nogales, E., and Doudna, J. A. (2017). Structures of the CRISPR genome integration complex. *Science* 357, 1113–1118. doi: 10.1126/science.aao0679

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

REVIEW
published: 20 September 2021
doi: 10.3389/fmicb.2021.732270

Blocking, Bending, and Binding: Regulation of Initiation of Chromosome Replication During the *Escherichia coli* Cell Cycle by Transcriptional Modulators That Interact With Origin DNA

Julia E. Grimwade and Alan C. Leonard*

Microbial Genetics Laboratory, Biological Sciences Program, Department of Biomedical and Chemical Engineering and Sciences, Florida Institute of Technology, Melbourne, FL, United States

Genome duplication is a critical event in the reproduction cycle of every cell. Because all daughter cells must inherit a complete genome, chromosome replication is tightly regulated, with multiple mechanisms focused on controlling when chromosome replication begins during the cell cycle. In bacteria, chromosome duplication starts when nucleoprotein complexes, termed orisomes, unwind replication origin (*oriC*) DNA and recruit proteins needed to build new replication forks. Functional orisomes comprise the conserved initiator protein, DnaA, bound to a set of high and low affinity recognition sites in *oriC*. Orisomes must be assembled each cell cycle. In *Escherichia coli*, the organism in which orisome assembly has been most thoroughly examined, the process starts with DnaA binding to high affinity sites after chromosome duplication is initiated, and orisome assembly is completed immediately before the next initiation event, when DnaA interacts with *oriC*'s lower affinity sites, coincident with origin unwinding. A host of regulators, including several transcriptional modulators, targets low affinity DnaA-*oriC* interactions, exerting their effects by DNA bending, blocking access to recognition sites, and/or facilitating binding of DnaA to both DNA and itself. In this review, we focus on orisome assembly in *E. coli.* We identify three known transcriptional modulators, SeqA, Fis (factor for inversion stimulation), and IHF (integration host factor), that are not essential for initiation, but which interact directly with *E. coli oriC* to regulate orisome assembly and replication initiation timing. These regulators function by blocking sites (SeqA) and bending *oriC* DNA (Fis and IHF) to inhibit or facilitate cooperative low affinity DnaA binding. We also examine how the growth rate regulation of Fis levels might modulate IHF and DnaA binding to *oriC* under a variety of nutritional conditions. Combined, the regulatory mechanisms mediated by transcriptional modulators help ensure that at all growth rates, bacterial chromosome replication begins once, and only once, per cell cycle.

Keywords: replication origin, DnaA, bacterial cell cycle, SeqA, factor for inversion stimulation, integration host factor

OPEN ACCESS

Edited by:
Morigen Morigen,
Inner Mongolia University, China

Reviewed by:
Rahul Saxena,
Georgetown University, United States
Gregory Marczynski,
McGill University, Canada

****Correspondence:***
Julia E. Grimwade
grimwade@fit.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *28 June 2021*
Accepted: *18 August 2021*
Published: *20 September 2021*

Citation:
Grimwade JE and Leonard AC (2021) Blocking, Bending, and Binding: Regulation of Initiation of Chromosome Replication During the Escherichia coli Cell Cycle by Transcriptional Modulators That Interact With Origin DNA. Front. Microbiol. 12:732270. doi: 10.3389/fmicb.2021.732270

INTRODUCTION TO ORISOME ASSEMBLY

Chromosome replication in bacteria ensues from a small region termed the origin of replication (*oriC)*, where the double-stranded DNA is unwound to provide the appropriate configuration for assembly of two new replication forks (Leonard and Méchali, 2013). The timing of replication initiation is tightly coupled to the cell's growth rate, such that, once a cell has reached a critical mass (and critical amount of initiator protein), a new round of chromosome replication is triggered (Cooper and Helmstetter, 1968; Løbner-Olesen et al., 1989). Once an origin has fired, no new rounds of chromosome replication are started until the next cell division cycle (Cooper and Helmstetter, 1968). To achieve this precise "once and only once per cell cycle" initiation, origin activation must be strictly regulated (Skarstad and Katayama, 2013; Katayama, 2017). Regulation is focused on assembly of a DNA unwinding machine (termed the orisome) that must be re-made each cell cycle. In all known bacterial types, orisomes comprise multiple copies of the conserved bacterial initiator protein, DnaA (Leonard and Grimwade, 2015; Bleichert et al., 2017). DnaA is also a transcriptional regulator of multiple genes (Messer and Weigel, 2003; Washington et al., 2017), including the *dnaA* gene itself (Speck et al., 1999; Hansen and Atlung, 2018).

Binding of DnaA to *oriC* is cooperative (Fuller et al., 1984; Messer, 2002), and in *Escherichia coli*, DnaA binding is directed by the arrangement of specific high and low affinity nine mer recognition sites within the origin (Rozgaja et al., 2011). *E. coli oriC* contains 11 of these recognition sites (Rozgaja et al., 2011; **Figure 1A**), but there is tremendous diversity in the numbers and arrangements of DnaA recognition sites among bacterial origins (Zawilak-Pawlik et al., 2005; Gao et al., 2013). A detailed analysis of orisome assembly has been performed in only a few bacterial types, with *E. coli* orisomes being the most extensively studied.

In *E. coli*, three of the eleven recognition sites (R1, R2, and R4) have the high affinity "consensus" sequence 5′-TTATNCACA (Schaper and Messer, 1995). These three sites are occupied throughout the cell cycle (Samitt et al., 1989; Cassler et al., 1995; Nievera et al., 2006) and may interact to constrain the origin (indicated by dotted lines in **Figure 2A**; constraint is also discussed below). The two gap regions between R1 and R2, and R2 and R4 each contain an array of four lower affinity sites, separated by two base pairs, whose sequences differ from consensus at two or more positions (**Figure 1A**; Leonard and Grimwade, 2010; Rozgaja et al., 2011). The lower affinity sites are occupied only at the time of initiation (Nievera et al., 2006). Completion of orisome assembly requires that the high affinity sites recruit and donate DnaA-ATP molecules to the arrays (Schaper and Messer, 1995; Miller et al., 2009; **Figure 2A**). Since DnaA-ADP bound to high affinity sites are capable of donation (Noguchi et al., 2015), it is likely the first interactions utilize the N-terminal domains (Domain 1). However, only the ATP-bound form of DnaA can engage in the cooperative binding needed to "fill the gaps" because only this form is recognized by six of the eight low affinity sites (tau2, I1, I2, I3, C2, and C3; McGarry et al., 2004; Grimwade et al., 2018). Domain III interactions are also required for this "gap filling" stage (Kawakami et al., 2005). Once bound to the arrayed sites, it is likely that the DnaA molecules form two pentameric oligomers in each half of *oriC* (Rozgaja et al., 2011; Shimizu et al., 2016). Although the exact structure of the pentamers is not known, it is likely that they cause the DnaA-*oriC* complex to assume a curved configuration (Shimizu et al., 2016), since DnaA binding is reported to bend DNA about 40 degrees (Schaper and Messer, 1995). It has been suggested that this DnaA-induced bending of *oriC* might create some of the stress required for origin unwinding (Grimwade et al., 2018). Consistent with this idea, immediately after the arrays are filled with DnaA-ATP, the A-T rich DNA unwinding element (DUE) at the left end of *oriC* becomes unwound and available for replicative helicase loading (Kowalski and Eddy, 1989; Bell and Kaguni, 2013; Rozgaja et al., 2011; **Figure 2A**). One of the single strands of DNA in the unwound region is also occupied by oligomeric DnaA-ATP, which may interact with sequences similar to the so called DnaA trio motifs found in other bacterial origins (Ozaki and Katayama, 2012; Richardson et al., 2016). Binding to single-stranded DNA has been proposed to help induce unwinding and to stabilize the unwound strands (Speck and Messer, 2001; Ozaki and Katayama, 2012). Because the ATP-bound form of DnaA is required for both double-stranded and single-stranded *oriC* binding, mechanisms that regulate cellular levels of DnaA-ATP play a key role in precise initiation timing (Skarstad and Katayama, 2013; Riber et al., 2016; Katayama et al., 2017).

Although DnaA-ATP alone is sufficient to build an orisome capable of triggering initiation in *E. coli*, its *oriC* also contains recognition sites for three accessory proteins: SeqA (Slater et al., 1995; Waldminghaus and Skarstad, 2009), integration host factor (IHF; Nash and Robertson, 1981; Polaczek, 1990; Rice, 1997), and factor for inversion stimulation (Fis; Finkel and Johnson, 1992; Filutowicz et al., 1992; **Figure 1A**). Binding of these three proteins to *oriC* regulates DnaA-ATP's occupation of lower affinity recognition sites (Grimwade et al., 2000; Ryan et al., 2004). At other chromosomal loci, SeqA, IHF, and Fis function as transcriptional modulators and may also be involved in organizing the bacterial chromosome (Schmid, 1990; Dorman and Deighan, 2003; Løbner-Olesen et al., 2003; Klungsøyr and Skarstad, 2004; Seshasayee et al., 2011; Dorman, 2014; Wang and Maier, 2015; Hołówka and Zakrzewska-Czerwińska, 2020). Although none of these three proteins is essential for *E. coli* viability, loss of any of them results in altered host growth (Filutowicz and Roll, 1990; Filutowicz et al., 1992; Von Freiesleben et al., 2000; Waldminghaus and Skarstad, 2009). Below, we review these transcriptional modulators and describe how they are utilized during orisome assembly to impart precise regulation of replication initiation.

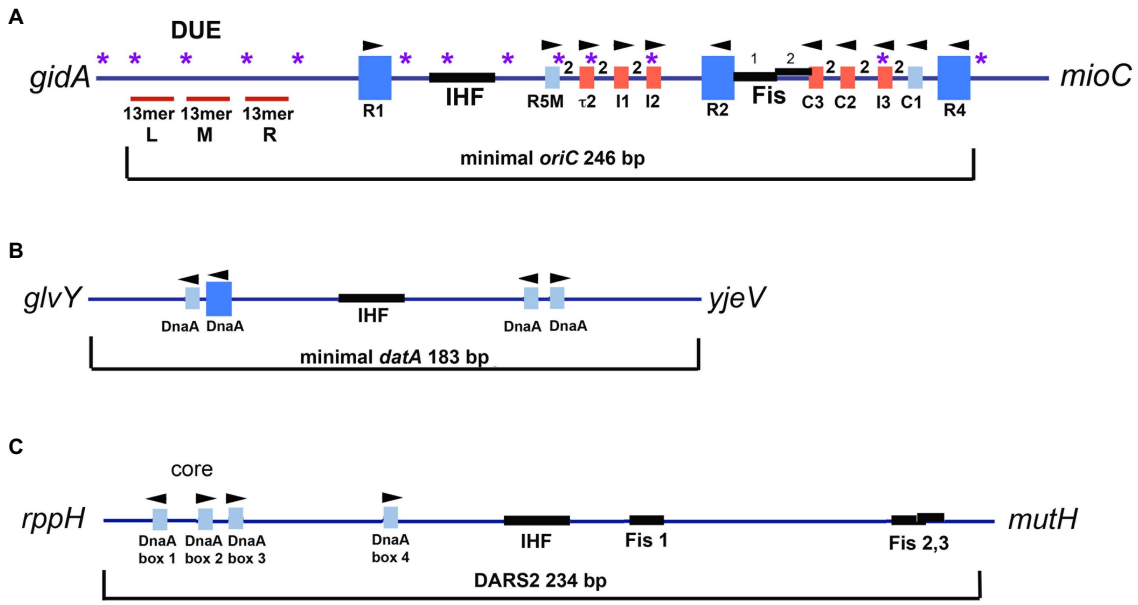


FIGURE 1 | (A) Map of *E. coli oriC*. The large blue rectangles are high affinity DnaA recognition sites. Smaller rectangles are low affinity DnaA recognition sites that deviate from the high affinity consensus by two or more bp and are shown in red for low affinity sites that preferentially bind DnaA-ATP, and blue for low affinity sites that bind both DnaA-ADP and DnaA-ATP. The number 2 between boxes indicates the 2 bp spacing between sites. The positions of the two binding sites for Fis, and the IHF binding region, are shown, as well as the left (L), middle (M), and right (R) 13mer repeats in the DNA unwinding element (DUE). Purple asterisks indicate the positions of 5′-GATC motifs, which, when hemimethylated after *oriC* is replicated, bind SeqA. Arrowheads indicate the orientation of the R285 residue in bound DnaA. The figure is adapted from Leonard and Grimwade (2015). **(B)** Map of *E. coli datA*. High and low affinity DnaA recognition sites and the IHF binding site are marked. Although the affinity of the sites has not been examined, the sites deviate from consensus by at least one bp. **(C)** Map of *E. coli* DARS2 region. DnaA recognition sites and the IHF and Fis sites are marked.

SeqA, A NEGATIVE REGULATOR OF INITIATION THAT BLOCKS LOW AFFINITY DnaA BINDING SITES

Successful formation of unwound orisomes results in recruitment of the replicative helicase, followed by assembly of new replication forks which then copy *oriC* DNA and move bidirectionally toward the terminus region of the chromosome. These processes remove DnaA from *oriC* but do not immediately reduce the amount of DnaA-ATP available to rebind to the origin, which, if it occurred, would result in re-initiation and over-replication. Seminal studies by Messer and Russell and Zinder (Messer et al., 1985; Russell and Zinder, 1987), examining *in vivo* replication of minichromosomes (plasmids initiating replication from a cloned copy of *oriC*) in strains deficient in DNA methylation, identified a role for DNA hemimethylation in the prevention of re-initiation. When DNA is replicated, the older strand carries methyl-adenines in the palindromic sequence 5′-GATC, while the daughter strand is unmethylated. This state of hemimethylation continues until DNA adenine methylase (encoded by the *dam* gene) methylates the daughter strand, restoring the parental methylation pattern (Geier and Modrich, 1979).

The *E. coli oriC* region contains an unusually high density of 5′-GATC motifs (Oka et al., 1984; sites are marked in **Figure 1A**); the motifs are found throughout the DUE, as well as in the IHF binding region and inside or overlapping several low affinity DnaA recognition sites (R5M, tau2, I2, and I3). After passage of the replication fork, these 5′-GATC sequences in *E. coli oriC* remain hemimethylated for approximately 1/3 of the cell cycle (Campbell and Kleckner, 1990) because the SeqA protein "sequesters" the origin from the *dam* methyltransferase (Lu et al., 1994; **Figure 2A**). SeqA is also a transcriptional regulator, rendering some promoter regions, including those for *dnaA* (Campbell and Kleckner, 1990) and *gidA*, the gene adjacent to the *oriC* DUE (Theisen et al., 1993) inactive for part of the cell cycle. There is also evidence that SeqA can be a positive regulator of transcription in bacteriophage lambda (Slomińska et al., 2001).

SeqA binds preferentially to hemimethylated 5′-GATC sequences, with a much lower affinity for the fully methylated version; it does not bind at all to unmethylated 5′-GATCs (Brendler et al., 1995; Slater et al., 1995). The basal SeqA structure is a dimer, and dimers can oligomerize to form filaments (Guarné et al., 2005). SeqA oligomers can interact with large stretches of DNA, resulting in large toroidal complexes that constrain negative supercoils (Klungsøyr and Skarstad, 2004; Skarstad and Katayama, 2013). These complexes localize at replication forks (Hiraga et al., 2000; Molina and Skarstad, 2004) and SeqA assemblies have been proposed to help organize origins and replication forks by pairing sister origins and forks in rapidly growing *E. coli*, so that they eventually segregate properly (Fossum et al., 2007; Morigen et al., 2009).

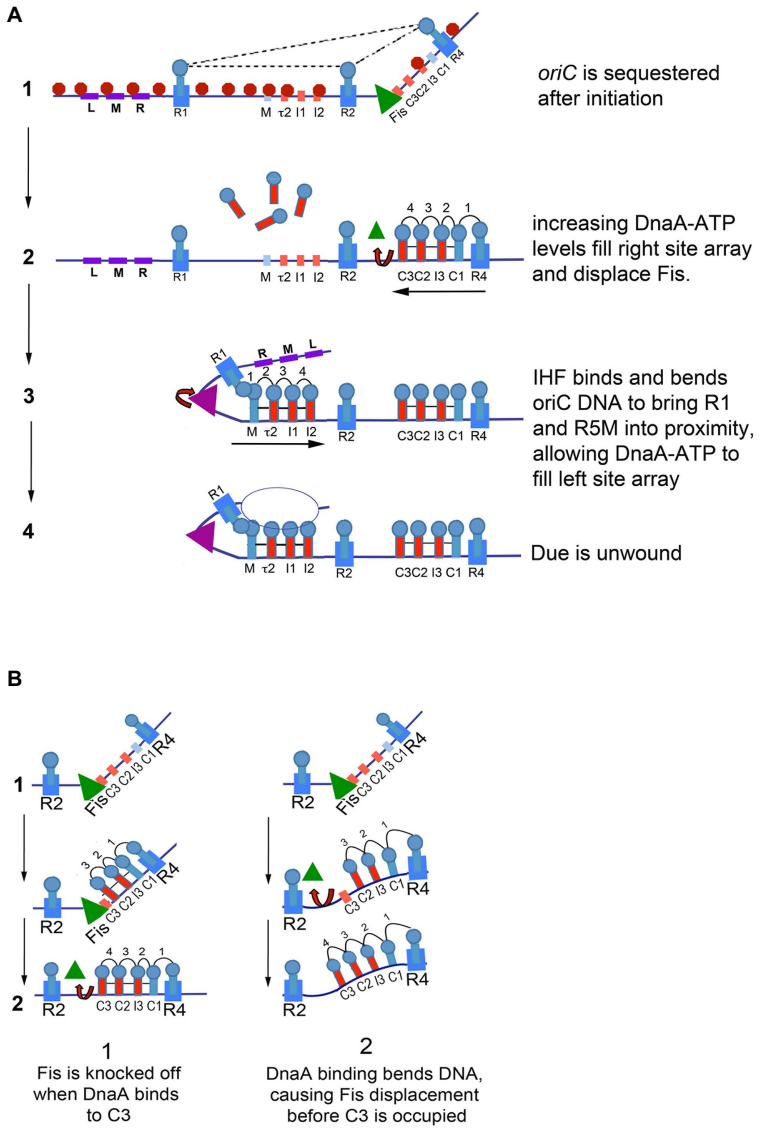


FIGURE 2 | (A) Model of staged *E. coli* orisome assembly. Stage1: After *oriC* is replicated, SeqA (red octagons) binds to hemimethylated 5′-GATC motifs, and DnaA rebinds to high affinity R1, R2, and R4 sites. Fis (green triangle) is also bound. DnaA and Fis binding constrains the origin, possibly through interactions among the bound DnaA molecules (indicated by dotted lines), and prevents IHF from binding to its recognition site. Stage 2: DnaA bound to R4 recruits DnaA for binding to C1. DnaA-ATP then progressively fills the remaining sites in the right array. DnaA displaces Fis, and loss of Fis allows IHF to bind. Stage 3: The IHF-induced bend allows DnaA cooperative binding of DnaA between R1 and R5M. DnaA then progressively fills the remaining sites in the left array. Stage 4: The DUE is unwound and one of the single strands interacts with DnaA-ATP bound to the left array. The figure is adapted from Leonard and Grimwade (2015). **(B)** Possible mechanisms of Fis displacement by DnaA. When sufficient levels of DnaA accumulate, Fis is displaced either by direct competition with DnaA binding to C3 (left panel) or by a conformation change caused by DnaA binding to the right low affinity sites in *oriC* (right panel).

At *oriC*, SeqA prevents orisome re-assembly immediately after an initiation event (Nievera et al., 2006). SeqA occupation of *oriC*'s hemimethylated 5′-GATC sites permits DnaA to rebind higher affinity sites, resetting the origin, but blocks DnaA binding to the lower affinity sites (Nievera et al., 2006; Taghbalout et al., 2000; **Figure 2A**). While the mechanism of SeqA blocking is not fully understood, it seems likely that SeqA, by occupying hemimethylated 5-GATC motifs, physically obstructs DnaA's access to the DUE and to its low affinity recognition sites. SeqA could also inhibit IHF bending, which would reduce cooperative DnaA binding in the entire left arrays of low affinity sites (IHF actions are described in more detail below). Some of the inhibition of initiation caused by SeqA may also result from a reduction in negative superhelicity (Torheim and Skarstad, 1999). Additionally, the inhibition of transcription from the *dnaA* gene by SeqA (Campbell and Kleckner, 1990; Theisen et al., 1993) contributes to the decrease in cellular DnaA-ATP

levels in the cell after initiation (Kurokawa et al., 1999), so that when *oriC* becomes fully remethylated, DnaA-ATP levels are low enough so that the loss of SeqA blocking does not result in immediate re-initiation (Bogan and Helmstetter, 1997).

DYNAMIC BINDING OF THE DNA-BENDING PROTEINS Fis AND IHF TO SPECIFIC SITES IN *oriC* REGULATES ORDERED ORISOME ASSEMBLY

Unlike SeqA, which was identified in a screen for regulators of initiation, both Fis and IHF were originally identified as players in phage recombination and subsequently were discovered to participate in many cellular DNA transactions including replication initiation (described below), transcription, recombination, and transposition, see reviews (Friedman, 1988; Finkel and Johnson, 1992). Both proteins contribute to regulated expression of hundreds of *E. coli* genes, reviewed in (Martínez-Antonio and Collado-Vides, 2003). Depending on the position of recognition sites within promoters, Fis and IHF can be either repressors or activators and there are many genes whose promoters contain recognition sites for both factors (Monteiro et al., 2020). A key feature of both proteins is their ability to bend DNA to an impressive degree; Fis may bend its target by 60 to 90 degrees (Gille et al., 1991) and IHF can produce 180 degree hairpins (Friedman, 1988; Ellenberger and Landy, 1997; Rice, 1997; Travers, 1997; Engelhorn and Geiselmann, 1998). Bending by these proteins also imparts topological effects on supercoiled DNA and plays a role in setting the degree of supercoiling in nucleoids (Schneider et al., 1999; Travers et al., 2001; Opel et al., 2004).

In *E. coli*, regulation of replication initiation by Fis and IHF requires binding of the two proteins to specific recognition sites within *oriC* (Roth et al., 1994). Two closely spaced sites for Fis have been identified between the high affinity R2 site and the array of low affinity sites in the right half of the origin (C3-C1; **Figure 1A**; Gille et al., 1991; Filutowicz et al., 1992; Hengen et al., 1997). One of the Fis sites partially overlaps the C3 site, which may have ramifications with regard to initiation regulation (described below). The binding site for IHF is in the left region of *oriC*, between the high affinity R1 site and the left array of low affinity sites (Polaczek, 1990; **Figure 1A**).

The roles of IHF and Fis in initiation regulation are not as straightforward as that of SeqA. IHF was shown to be a positive regulator in the *in vitro* replication and unwinding of minichromosomes (Hwang and Kornberg, 1992), as well as in the replication of some other plasmid systems (Filutowicz and Appelt, 1988; Biek and Cohen, 1989). In contrast, studies examining the role of Fis in *E. coli* replication initiation report both positive and negative effects (Hiasa and Marians, 1994; Roth et al., 1994; Wold et al., 1996; Margulies and Kaguni, 1998; Flåtten and Skarstad, 2013). Greater insight was provided by examining the binding of Fis and IHF during the cell cycle (Cassler et al., 1995; Ryan et al., 2002).

During the cell cycle, Fis and IHF bind to *oriC* at different times (Cassler et al., 1995). Fis occupies its recognition site most of the cell cycle but is not bound during the time of replication initiation; in contrast, IHF interacts with its *oriC* site only transiently, immediately before the DNA unwinding step and when Fis is not bound (**Figure 2**). These results provided evidence that the two proteins play opposing roles in initiation, with the simplest scenario being that Fis acts as an inhibitor of orisome formation, while IHF is stimulatory.

The positive effect of IHF relies on the bending of *oriC* DNA, which facilitates DnaA cooperative binding (Grimwade et al., 2000). IHF, when bound to its recognition site between R1 and R5M, stabilizes a bend in *oriC* that puts R1 in proximity to R5M, allowing DnaA bound at R1 to recruit additional DnaA and donate it to R5M (**Figure 2**); DnaA-ATP binding can then extend to all the arrayed sites (Leonard and Grimwade, 2010; Rozgaja et al., 2011). Additionally, the IHF-induced bend places the DUE in *E. coli oriC* near the low affinity sites in the left half of the origin, which permits interaction between the array-bound DnaA-ATP and single-stranded DNA in the DUE (Ozaki and Katayama, 2012; Ozaki et al., 2012; Sakiyama et al., 2017; **Figure 2**).

Cellular levels of IHF should be high enough to allow it to bind to *oriC* throughout the cell cycle. However, the finding that IHF binding is restricted to the time of initiation (Cassler et al., 1995) suggested that some factor prevented the interaction of IHF with its *oriC* recognition site. Examination of orisome formation *in vitro* revealed this inhibitory factor was Fis, even though the Fis and IHF recognition sites are separated by nearly 90 bp (Ryan et al., 2004). Fis inhibition begins in the early stages of orisome assembly, when *oriC* is sequestered by SeqA (described above). At this time, the origin is bound to Fis as well as to DnaA at R1, R2, and R4 (**Figure 2A**). Binding of Fis and DnaA topologically constrains *oriC,* and the constraint is sufficient to prohibit binding of IHF at physiological concentrations (Kaur et al., 2014; **Figure 2**). Although the details of the configuration of the constrained complex remain incomplete, there are data that suggest it is maintained by interactions among DnaA molecules bound to R1, R2, and R4 (Kaur et al., 2014). Whatever its nature, the constraint imparted by DnaA and Fis blocks IHF binding until increasing concentrations of DnaA displace Fis shortly before initiation (Ryan et al., 2004; **Figure 2B**). Once Fis leaves *oriC*, IHF binding and bending rapidly promotes DnaA binding and DUE unwinding (described above). Thus, the dynamic binding of Fis and IHF during the cell cycle establishes an initiation "on switch" (Ryan et al., 2004).

Key to the Fis-IHF initiation switch is the interplay between Fis and DnaA-ATP, mediated by the position of the Fis binding site at the end of the right array of low affinity. In the presence of Fis, only the right half of *oriC* can direct cooperative DnaA binding, since the interaction of DnaA at R1 and R5M in *oriC*'s left half is prohibited until IHF binding and bending takes place (Rozgaja et al., 2011; **Figure 2A**). Growth of the right DnaA pentamer competes with Fis binding, and initially Fis wins the contest and delays DnaA-ATP binding (Ryan et al., 2004). However, as cellular DnaA-ATP levels rise, DnaA

displaces Fis while simultaneously completing assembly of the DnaA pentamer associated with the right array of sites (Ryan et al., 2004). It is not entirely clear how DnaA displaces Fis, but there are two possible scenarios (**Figure 2B**). The first is that Fis is simply knockedoff its site when the DnaA oligomer extends to C3. The second possibility is that a DnaA-induced curve in *oriC* DNA (Schaper and Messer, 1995; Shimizu et al., 2016) weakens Fis interactions enough for Fis to be released from *oriC* before DnaA occupation of the right site array is complete.

THE Fis-IHF SWITCH ALSO REGULATES INITIATION SYNCHRONY IN RAPID GROWTH

Members of the *Enterobacteriaceae* (including *E. coli*) and some other bacterial types are capable of rapid growth (doubling their numbers every 20 min) under nutrient-rich conditions, for example, see (Schaechter et al., 1958). However, except in very slow growing cells, the time to replicate the chromosome is constant over a range of growth rates, measured to be around 40–50 min in a variety of *E. coli* strains (Cooper and Helmstetter, 1968; Helmstetter et al., 1992). If the time required to replicate the genome exceeds the generation time, new initiations from *oriC* must take place before previous rounds of chromosome replication are completed (Cooper and Helmstetter, 1968). Thus, bacteria growing rapidly in rich media contain multiple copies of *oriC*, all of which fire synchronously (Skarstad et al., 1986). Synchronous initiation is observed even when the cell contains extra copies of *oriC* carried on minichromosomes (Helmstetter and Leonard, 1987), indicating that the initiation mechanism does not "count" the number *oriC* copies. Since assembly of each orisome requires the same amount of new DnaA-ATP, and mass at initiation is fairly constant regardless of how many origins the cell contains (Donachie, 1968), the attainment of initiation synchrony presents a bit of a puzzle. One model suggests that there is an initiation cascade that resembles a chain reaction, where the DnaA that is displaced when one copy of initiating *oriC* is rapidly picked up and used to complete assembly of another orisome (Løbner-Olesen et al., 1994). However, even a cascade would result in some asynchrony, suggesting that there could be additional regulation to ensure that all origins fire at the same time.

The switch mechanism imparted by the dynamic interactions of Fis and IHF, described above, could provide the regulation that fine tunes an initiation cascade, if the amount of DnaA required to displace Fis during the cell cycle exceeded the amount required to fill the remaining empty DnaA binding sites on other origin copies (**Figure 3A**). In this scenario, the cell would contain more DnaA than needed to complete orisome assembly at any individual *oriC* copy, so loss of Fis would cause both the rapid binding of IHF and provide available DnaA for unfired *oriC* copies (Ryan et al., 2004). Initiation and the resulting orisome disassembly would further increase the amount of available DnaA-ATP, which could cascade to any remaining unfired origins (**Figure 3A**). Evidence for both Fis and IHF acting as modulators of initiation synchrony is shown by the asynchronous initiations observed in fast-growing *E. coli* cells carrying loss-of-function mutations in the genes encoding Fis or IHF (Riber et al., 2016). Further, rapidly growing Fis null cells not only have asynchronous initiations, they also have fewer origins than wild-type cells (i.e., the Fis null cells under-initiate; Flåtten and Skarstad, 2013; Rao et al., 2018), consistent with the idea that, without Fis, there is not enough DnaA-ATP in the cell for every orisome to complete its assembly. In other words, by being an inhibitor of orisome assembly, Fis becomes a positive regulator of initiation synchrony.

GROWTH REGULATION OF Fis PRODUCES DIFFERENT ORISOME ASSEMBLY PATHS FOR FAST GROWTH, SLOW GROWTH, AND STATIONARY PHASE

The regulatory role of Fis in orisome assembly (described above) must be limited to the more rapid growth rates of cells containing multiple copies of *oriC*, because Fis expression is growth rate regulated, with high levels (10–15 thousand molecules per cell) maintained during rapid growth and low to no protein (less than 100 molecules per cell) present during slow growth and stationary phase (Ball et al., 1992; Schneider et al., 1997). Since IHF is not growth rate regulated (Ditto et al., 1994), the absence of Fis in slowly growing cells should result in a lack of Fis inhibition of IHF binding to *oriC*. If so, this should cause orisome assembly to begin with R1 and R5M already in proximity due to bound IHF, and the low affinity site arrays in both left and right regions of *oriC* should start filling with DnaA at approximately the same time (**Figure 3B**). However, the left region of *oriC* has been shown to be more important for unwinding (Ozaki et al., 2012) and it is possible that the right region has little function during slow growth. This idea is consistent with previous findings that some deletions in the right half of *oriC* cause origin inactivation under fast growth, but not slow growth conditions (Stepankiw et al., 2009).

The absence of Fis during slow growth also prevents the cell from accumulating any more DnaA-ATP than is required to assemble an orisome on its single copy of *oriC*, and it is likely that sites in *oriC* become occupied as soon as levels of free DnaA-ATP are high enough to bind to them. Interestingly, recent studies revealed that in slow growth conditions, the tau2 site in the left half of *oriC* may represent a bottleneck in orisome assembly (Rao et al., 2018), determining how much DnaA-ATP is needed. The reason for this is not known, but it may be related to tau2's position as the first low affinity DnaA-ATP site in the array (**Figure 1A**); R5M binds both nucleotide forms of the initiator equally and is likely to be occupied by DnaA-ADP *in vivo* (Grimwade et al., 2007). Logically, slow growth orisomes could start orisome assembly

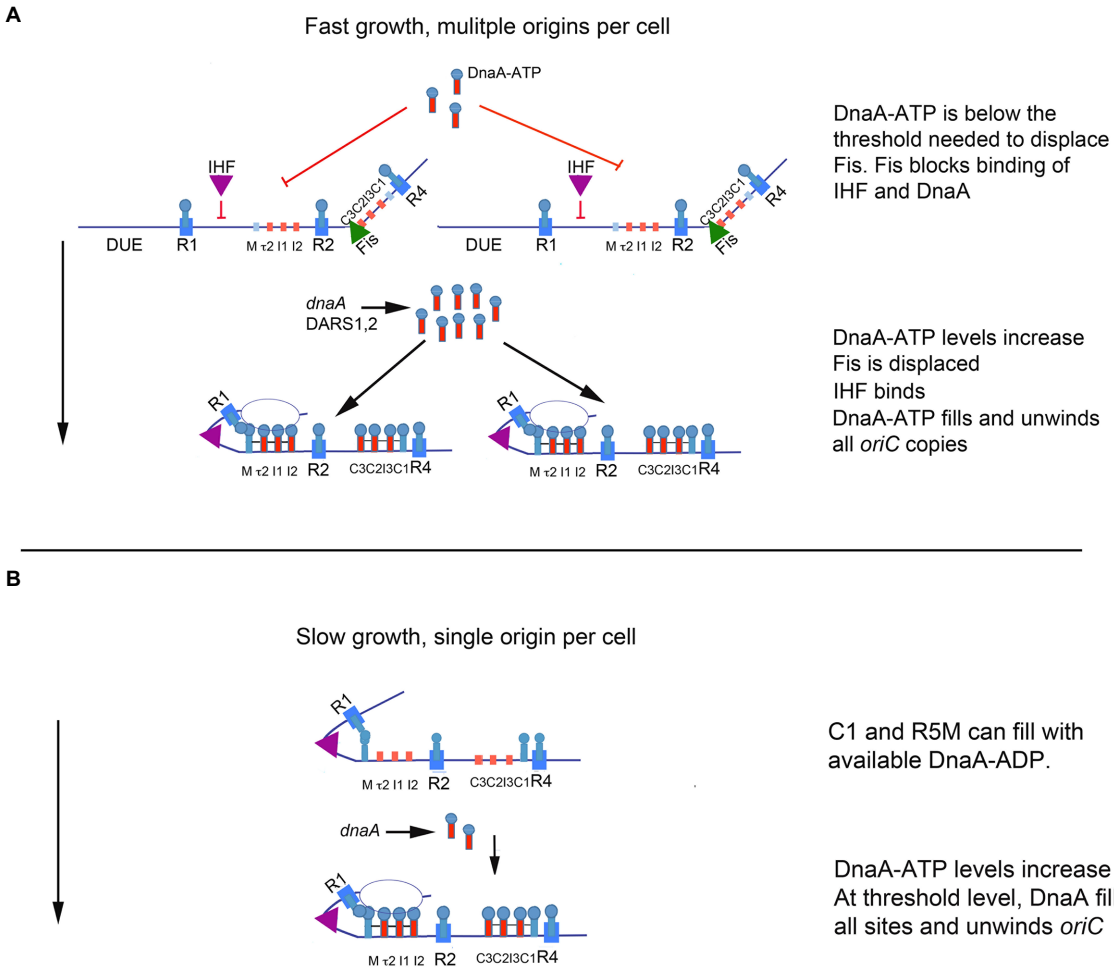


FIGURE 3 | Model of orisome assembly in fast and slow growing cells. **(A)** In rapidly growing cells, DnaA-ATP site occupation is regulated by Fis. Bound Fis prevents free DnaA-ATP from occupying *oriC* and excess DnaA-ATP (provided by new synthesis and DARS2) accumulates. When sufficient DnaA-ATP becomes available, Fis is displaced. Loss of Fis allows IHF to bind to *oriC*, and available DnaA-ATP binds to low affinity sites on all *oriC* copies in the cell. Orisome assembly may be completed by an initiation cascade (see text for details). **(B)** In slowly growing cells, cells contain a single copy of *oriC*, bound to IHF throughout the cell cycle. All low affinity DnaA-ATP sites become occupied as DnaA-ATP becomes available (see text for details). When enough DnaA-ATP has accumulated to fill all sites in the left region of *oriC*, the origin is unwound. The figure is adapted from Rao et al. (2018).

by using available DnaA-ADP for cooperative binding to R5M but would have to pause until levels of DnaA-ATP were high enough to extend binding to tau2.

At intermediate growth rates, cells contain lower but detectable levels of Fis. This creates the interesting, although speculative possibility that depending on growth rate, Fis may occupy one or both of its two recognition sites between R1 and C3, producing different levels of repression over a range of growth conditions (Gille et al., 1991; Hengen et al., 1997, 2003). If this is the case, then a reasonable corollary would be that as Fis repression decreases, the number of DnaA-ATP molecules required to build an unwound orisome would also decrease. Consistent with this idea, there is evidence in the literature that initiation timing is regulated differently in fast and slow growing cells (Wold et al., 1994; Zheng et al., 2020). It is also possible that regulators of DnaA-ATP regeneration (mediated by DARS2 locus, see below) may function differently during fast and slow growth in ways that could affect initiation timing.

The behavior of Fis and IHF also changes at *oriC* as fast-growing bacteria enter stationary phase, stop initiating chromosome replication, and return to the single copy state (Boye and Løbner-Olesen, 1991). Since the levels of IHF increase and the Fis levels become very low, the stationary phase *oriC* is complexed with IHF (Ditto et al., 1994; Cassler et al., 1999). Unless an undiscovered stationary phase repressor of initiation also exists, it must be assumed that these orisomes are incompletely assembled due to insufficient DnaA. This configuration may be advantageous for cells to restart growth when conditions allow, although the effect will be transient, since chromosomal origins that leave stationary phase for nutritionally rich conditions will rebind Fis as cells enter early exponential growth phase,

when Fis levels peak (Ball et al., 1992). It is also possible that IHF binding to *oriC* helps the cell in other ways, perhaps by setting *oriC* in a configuration that is beneficial for long-term non-initiating conditions followed by a stepwise origin reset.

SeqA, Fis, and IHF Modulate Levels and Activity of Other Initiation Regulators

In addition to their participation in orisome assembly, SeqA, Fis, and IHF also indirectly affect initiation timing by interacting with non-*oriC* sites on the *E. coli* chromosome that help regulate intracellular DnaA-ATP levels (Katayama et al., 2010; Leonard and Grimwade, 2015; Riber et al., 2016).

Following each new start of chromosome replication, DnaA-ATP levels in cells decrease (Kurokawa et al., 1999). In part, this is because SeqA sequesters the *dnaA* gene region and prevents new DnaA-ATP synthesis for approximately 20% of the cell cycle (Campbell and Kleckner, 1990). Although the mechanism by which SeqA blocks *dnaA* transcription gene has not been thoroughly explored, there is a high density of 5′-GATC residues in the *dnaA* gene region, similar to that found in *oriC,* which, when hemimethylated, could support a SeqA complex that prevents binding of transcriptional activators. Additionally, a replisome-associated protein (Hda) stimulates the ATP hydrolysis of any DNA-bound DnaA that it encounters (Kato and Katayama, 2001). While this mechanism, termed regulatory inactivation of DnaA (RIDA), is the primary means to inactivate DnaA-ATP, a secondary inactivation mechanism exists for any unbound DnaA-ATP that remains after initiation (for example, from DnaA-ATP oligomers displaced from a newly replicated origin). Secondary inactivation is dependent on a DNA locus termed *datA* (about 47 kbp from *oriC*), whose deletion causes intracellular levels of DnaA-ATP to increase (Ogawa et al., 2002; Kasho and Katayama, 2013). *DatA* contains an essential binding site for IHF and two sets of paired DnaA recognition sites (Kasho and Katayama, 2013; **Figure 1B**). The binding of IHF at *datA*, like *oriC*, is dynamically regulated during the cell cycle by mechanisms that are not yet known, but IHF binding at *datA* is a post-initiation event (detected for 20–30 min after initiation during rapid growth). IHF most likely bends the DNA to bring the two pairs of DnaA sites into proximity (Katayama et al., 2017), to promote the formation of new DnaA multimers on *datA* that are then susceptible to RIDA (Kasho et al., 2017).

A second locus regulated by Fis and IHF, termed DARS2, is located halfway between *oriC* and the chromosomal terminus. DARS2 raises the level of active DnaA *via* a mechanism that removes the nucleotide from bound DnaA-ADP to promote rapid regeneration to DnaA-ATP (Fujimitsu et al., 2009; Sugiyama et al., 2019) (An additional site, termed DARS1, also contributes to increasing DnaA-ATP levels but is not regulated by Fis or IHF). Regenerated DnaA-ATP is used during rapid growth where new synthesis is not sufficient to produce the required threshold level for Fis displacement and synchronous initiation at all origin copies (described above). DARS2 contains four DnaA recognition sites critical for regeneration activity, positioned near an essential IHF recognition site (**Figure 1C**). Three of these sites are closely spaced and form what is termed the DARS "core" (**Figure 1C**; Sugiyama et al., 2019). DARS2 activity is dependent on binding of both IHF and Fis, and similar to their dynamic binding at *oriC*, Fis binds DARS2 throughout the cell cycle, and IHF binds only to activate DARS2 immediately prior to the initiation step (Kasho et al., 2014). It is not yet known whether Fis is responsible for restricting IHF binding, as is the case for the orisome. It is also unclear how Fis and IHF promote the interactions among the three bound DnaA-ADP molecules to effect the disassociation of ADP, although it seems likely that DNA bending is involved to bring DnaA molecules together. However, since Fis is growth rate regulated, it is logical to think that the amount of DnaA-ATP made available by regeneration at DARS2 will also vary in fast and slow growing cells.

Binding of IHF to *oriC*, *datA*, and DARS2 is temporally regulated during cell cycle progression such that *oriC* is bound first, then *datA*, and then DARS2 (Kasho and Katayama, 2013; Kasho et al., 2014). While the mechanism responsible is not yet defined, the movement of the replication fork may cause the displacement of IHF from one location to another as the chromosome is replicated. This idea is supported by studies which show that the chromosomal positions of *datA* and DARS2 relative to *oriC* are important for cell cycle control and bacterial fitness (Frimodt-Møller et al., 2016; Inoue et al., 2016).

CONCLUSION

The dynamic bending, blocking, and binding properties of the transcriptional regulators SeqA, Fis, and IHF provide an elegant solution to multiple problems that could disrupt properly timed initiation and eventually prevent equivalent inheritance of genomic DNA in daughter cells. One pitfall that must be avoided after every initiation event is a second triggering of the same origin, which can result in closely spaced replication forks that could collide and then collapse. SeqA helps prevent this immediately after initiation (Sutera and Lovett, 2006), by blocking DnaA's access to *oriC* and preventing reformation of the orisome. Coincidently, SeqA inhibition of new DnaA-ATP synthesis contributes to the reduction in initiation potential (Skarstad and Katayama, 2013). IHF also safeguards against re-initiation by binding and bending the *datA* locus to promote the formation of a bound DnaA-ATP complex that then can be inactivated by RIDA.

Asynchronous initiations are also problematic to replicating *E. coli* cells. If all origins in rapidly growing cells do not fire synchronously, when cells divide, the daughter cells will contain different amounts of DNA, which could result in eventual formation of anucleate cells. The dynamic binding of Fis and IHF at *oriC* promotes initiation synchrony because it delays DnaA occupation of *oriC* until sufficient DnaA-ATP has accumulated to fill multiple origins (described above, and in **Figure 3A**). Further, since Fis is growth rate regulated, its effect on the amount of DnaA-ATP needed for initiation is different at different growth rates, which helps solve the problem of ensuring that orisome assembly and DnaA-ATP availability

are compatible with specific cell cycle stages over a wide variety of growth rates (**Figure 3B**).

While transcriptional modulators clearly play a key role in regulating *E. coli* origin function during the cell cycle, far less is known about other bacterial types. There are clearly transcriptional regulators that bind to other bacterial origins and regulate initiation; these include CtrA, SpoOA, AdpA, and MtrA in *C. crescentus*, *B. subtilis*, *S. coelicolor*, and *M. tuberculosis*, respectively, reviewed in (Wolanski et al., 2014; Grimwade and Leonard, 2019). However, for the most part, how precise initiation timing is achieved in most bacterial types is not well understood. It is certainly possible that part of the immense diversity found among bacterial replication origins are caused by the need for different types of regulation, some of which could involve transcriptional modulators. Hopefully, expanding studies to greater numbers of bacterial types will determine whether analogs of SeqA, Fis, and IHF regulate initiation by similar strategies or will reveal new uses of transcriptional regulators in diverse bacterial orisome assemblies.

AUTHOR CONTRIBUTIONS

JG and AL contributed equally to the writing of this article.

FUNDING

The work from our laboratory that is cited in this review was supported by the Public Health Service grant no. GM54042. Publication of this article was funded by the Open Access Subvention Fund of the Florida Tech Libraries and by the Office of Dean of the Florida Tech College of Engineering and Science.

REFERENCES

Ball, C. A., Osuna, R., Ferguson, K. C., and Johnson, R. C. (1992). Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J. Bacteriol.* 174, 8043–8056. doi: 10.1128/jb.174.24.8043-8056.1992

Bell, S. P., and Kaguni, J. M. (2013). Helicase loading at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.* 5:a010124. doi: 10.1101/cshperspect.a010124

Biek, D. P., and Cohen, S. N. (1989). Involvement of integration host factor (IHF) in maintenance of plasmid pSC101 in *Escherichia coli*: characterization of pSC101 mutants that replicate in the absence of IHF. *J. Bacteriol.* 171, 2056–2065. doi: 10.1128/jb.171.4.2056-2065.1989

Bleichert, F., Botchan, M. R., and Berger, J. M. (2017). Mechanisms for initiating cellular DNA replication. *Science* 355:aah6317. doi: 10.1126/science.aah6317

Bogan, J. A., and Helmstetter, C. E. (1997). DNA sequestration and transcription in the oriC region of *Escherichia coli*. *Mol. Microbiol.* 26, 889–896. doi: 10.1046/j.1365-2958.1997.6221989.x

Boye, E., and Løbner-Olesen, A. (1991). Bacterial growth control studied by flow cytometry. *Res. Microbiol.* 142, 131–135. doi: 10.1016/0923-2508(91)90020-B

Brendler, T., Abeles, A., and Austin, S. (1995). A protein that binds to the P1 origin core and the oriC 13mer region in a methylation-specific fashion is the product of the host seqA gene. *EMBO J.* 14, 4083–4089. doi: 10.1002/j.1460-2075.1995.tb00080.x

Campbell, J. L., and Kleckner, N. (1990). *E. coli* oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* 62, 967–979. doi: 10.1016/0092-8674(90)90271-F

Cassler, M. R., Grimwade, J. E., and Leonard, A. C. (1995). Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.* 14, 5833–5841. doi: 10.1002/j.1460-2075.1995.tb00271.x

Cassler, M. R., Grimwade, J. E., McGarry, K. C., Mott, R. T., and Leonard, A. C. (1999). Drunken-cell footprints: nuclease treatment of ethanol-permeabilized bacteria reveals an initiation-like nucleoprotein complex in stationary phase replication origins. *Nucleic Acids Res.* 27, 4570–4576. doi: 10.1093/nar/27.23.4570

Cooper, S., and Helmstetter, C. E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. *J. Mol. Biol.* 31, 519–540. doi: 10.1016/0022-2836(68)90425-7

Ditto, M. D., Roberts, D., and Weisberg, R. A. (1994). Growth phase variation of integration host factor level in *Escherichia coli*. *J. Bacteriol.* 176, 3738–3748. doi: 10.1128/jb.176.12.3738-3748.1994

Donachie, W. D. (1968). Relationship between cell size and time of initiation of DNA replication. *Nature* 219, 1077–1079. doi: 10.1038/2191077a0

Dorman, C. J. (2014). Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. *J. Mol. Microbiol. Biotechnol.* 24, 316–331. doi: 10.1159/000368850

Dorman, C. J., and Deighan, P. (2003). Regulation of gene expression by histone-like proteins in bacteria. *Curr. Opin. Genet. Dev.* 13, 179–184. doi: 10.1016/S0959-437X(03)00025-X

Ellenberger, T., and Landy, A. (1997). A good turn for DNA: the structure of integration host factor bound to DNA. *Structure* 5, 153–157. doi: 10.1016/S0969-2126(97)00174-3

Engelhorn, M., and Geiselmann, J. (1998). Maximal transcriptional activation by the IHF protein of *Escherichia coli* depends on optimal DNA bending by the activator. *Mol. Microbiol.* 30, 431–441. doi: 10.1046/j.1365-2958.1998.01078.x

Filutowicz, M., and Appelt, K. (1988). The integration host factor of *Escherichia coli* binds to multiple sites at plasmid R6K gamma origin and is essential for replication. *Nucleic Acids Res.* 16, 3829–3843. doi: 10.1093/nar/16.9.3829

Filutowicz, M., and Roll, J. (1990). The requirement of IHF protein for extrachromosomal replication of the *Escherichia coli* oriC in a mutant deficient in DNA polymerase I activity. *New Biol.* 2, 818–827.

Filutowicz, M., Ross, W., Wild, J., and Gourse, R. L. (1992). Involvement of Fis protein in replication of the *Escherichia coli* chromosome. *J. Bacteriol.* 174, 398–407. doi: 10.1128/jb.174.2.398-407.1992

Finkel, S. E., and Johnson, R. C. (1992). The Fis protein: it's not just for DNA inversion anymore. *Mol. Microbiol.* 6, 3257–3265. doi: 10.1111/j.1365-2958.1992.tb02193.x

Flåtten, I., and Skarstad, K. (2013). The Fis protein has a stimulating role in initiation of replication in *Escherichia coli in vivo*. *PLoS One* 8:e83562. doi: 10.1371/journal.pone.0083562

Fossum, S., Crooke, E., and Skarstad, K. (2007). Organization of sister origins and replisomes during multifork DNA replication in *Escherichia coli*. *EMBO J.* 26, 4514–4522. doi: 10.1038/sj.emboj.7601871

Friedman, D. I. (1988). Integration host factor: a protein for all reasons. *Cell* 55, 545–554. doi: 10.1016/0092-8674(88)90213-9

Frimodt-Møller, J., Charbon, G., Krogfelt, K. A., and Løbner-Olesen, A. (2016). DNA replication control is linked to genomic positioning of control regions in *Escherichia coli*. *PLoS Genet.* 12:e1006286. doi: 10.1371/journal.pgen.1006286

Fujimitsu, K., Senriuchi, T., and Katayama, T. (2009). Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* 23, 1221–1233. doi: 10.1101/gad.1775809

Fuller, R. S., Funnell, B. E., and Kornberg, A. (1984). The dnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* 38, 889–900. doi: 10.1016/0092-8674(84)90284-8

Gao, F., Luo, H., and Zhang, C. T. (2013). DoriC 5.0: an updated database of oriC regions in both bacterial and archaeal genomes. *Nucleic Acids Res.* 41, D90–D93. doi: 10.1093/nar/gks990

Geier, G. E., and Modrich, P. (1979). Recognition sequence of the dam methylase of *Escherichia coli* K12 and mode of cleavage of Dpn I endonuclease. *J. Biol. Chem.* 254, 1408–1413. doi: 10.1016/S0021-9258(17)34217-5

Gille, H., Egan, J. B., Roth, A., and Messer, W. (1991). The FIS protein binds and bends the origin of chromosomal DNA replication, oriC, of *Escherichia coli*. *Nucleic Acids Res.* 19, 4167–4172. doi: 10.1093/nar/19.15.4167

Grimwade, J. E., and Leonard, A. C. (2019). Blocking the trigger: inhibition of the initiation of bacterial chromosome replication as an antimicrobial strategy. *Antibiotics* 8:111. doi: 10.3390/antibiotics8030111

Grimwade, J. E., Rozgaja, T. A., Gupta, R., Dyson, K., Rao, P., and Leonard, A. C. (2018). Origin recognition is the predominant role for DnaA-ATP in initiation of chromosome replication. *Nucleic Acids Res.* 46, 6140–6151. doi: 10.1093/nar/gky457

Grimwade, J. E., Ryan, V. T., and Leonard, A. C. (2000). IHF redistributes bound initiator protein, DnaA, on supercoiled oriC of *Escherichia coli*. *Mol. Microbiol.* 35, 835–844. doi: 10.1046/j.1365-2958.2000.01755.x

Grimwade, J. E., Torgue, J. J., McGarry, K. C., Rozgaja, T., Enloe, S. T., and Leonard, A. C. (2007). Mutational analysis reveals *Escherichia coli* oriC interacts with both DnaA-ATP and DnaA-ADP during pre-RC assembly. *Mol. Microbiol.* 66, 428–439. doi: 10.1111/j.1365-2958.2007.05930.x

Guarné, A., Brendler, T., Zhao, Q., Ghirlando, R., Austin, S., and Yang, W. (2005). Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. *EMBO J.* 24, 1502–1511. doi: 10.1038/sj.emboj.7600634

Hansen, F. G., and Atlung, T. (2018). The DnaA tale. *Front. Microbiol.* 9:319. doi: 10.3389/fmicb.2018.00319

Helmstetter, C. E., Eenhuis, C., Theisen, P., Grimwade, J., and Leonard, A. C. (1992). Improved bacterial baby machine: application to *Escherichia coli* K-12. *J. Bacteriol.* 174, 3445–3449. doi: 10.1128/jb.174.11.3445-3449.1992

Helmstetter, C. E., and Leonard, A. C. (1987). Coordinate initiation of chromosome and minichromosome replication in *Escherichia coli*. *J. Bacteriol.* 169, 3489–3494. doi: 10.1128/jb.169.8.3489-3494.1987

Hengen, P. N., Bartram, S. L., Stewart, L. E., and Schneider, T. D. (1997). Information analysis of Fis binding sites. *Nucleic Acids Res.* 25, 4994–5002. doi: 10.1093/nar/25.24.4994

Hengen, P. N., Lyakhov, I. G., Stewart, L. E., and Schneider, T. D. (2003). Molecular flip-flops formed by overlapping Fis sites. *Nucleic Acids Res.* 31, 6663–6673. doi: 10.1093/nar/gkg877

Hiasa, H., and Marians, K. J. (1994). Fis cannot support oriC DNA replication in vitro. *J. Biol. Chem.* 269, 24999–25003. doi: 10.1016/S0021-9258(17)31489-8

Hiraga, S., Ichinose, C., Onogi, T., Niki, H., and Yamazoe, M. (2000). Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of oriC copies in *Escherichia coli*. *Genes Cells* 5, 327–341. doi: 10.1046/j.1365-2443.2000.00334.x

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

Hwang, D. S., and Kornberg, A. (1992). Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. *J. Biol. Chem.* 267, 23083–23086. doi: 10.1016/S0021-9258(18)50059-4

Inoue, Y., Tanaka, H., Kasho, K., Fujimitsu, K., Oshima, T., and Katayama, T. (2016). Chromosomal location of the DnaA-reactivating sequence DARS2 is important to regulate timely initiation of DNA replication in *Escherichia coli*. *Genes Cells* 21, 1015–1023. doi: 10.1111/gtc.12395

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

Kasho, K., Tanaka, H., Sakai, R., and Katayama, T. (2017). Cooperative DnaA binding to the negatively supercoiled datA locus stimulates DnaA-ATP hydrolysis. *J. Biol. Chem.* 292, 1251–1266. doi: 10.1074/jbc.M116.762815

Katayama, T. (2017). Initiation of DNA replication at the chromosomal origin of *E. coli*, oriC. *Adv. Exp. Med. Biol.* 1042, 79–98. doi: 10.1007/978-981-10-6955-0_4

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

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

Kaur, G., Vora, M. P., Czerwonka, C. A., Rozgaja, T. A., Grimwade, J. E., and Leonard, A. C. (2014). Building the bacterial orisome: high-affinity DnaA recognition plays a role in setting the conformation of oriC DNA. *Mol. Microbiol.* 91, 1148–1163. doi: 10.1111/mmi.12525

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

Klungsøyr, H., and Skarstad, K. (2004). K. Positive supercoiling is generated in the presence of *Escherichia coli* SeqA protein. *Mol. Microbiol.* 54, 123–131. doi: 10.1111/j.1365-2958.2004.04239.x

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. doi: 10.1002/j.1460-2075.1989.tb08620.x

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

Leonard, A. C., and Grimwade, J. E. (2010). Regulating DnaA complex assembly: it is time to fill the gaps. *Curr. Opin. Microbiol.* 13, 766–772. doi: 10.1016/j.mib.2010.10.001

Leonard, A. C., and Grimwade, J. E. (2015). The orisome: structure and function. *Front. Microbiol.* 6:545. doi: 10.3389/fmicb.2015.00545

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

Løbner-Olesen, A., Hansen, F. G., Rasmussen, K. V., Martin, B., and Kuempel, P. L. (1994). The initiation cascade for chromosome replication in wild-type and dam methyltransferase deficient *Escherichia coli* cells. *EMBO J.* 13, 1856–1862. doi: 10.1002/j.1460-2075.1994.tb06454.x

Løbner-Olesen, A., Marinus, M. G., and Hansen, F. G. (2003). Role of SeqA and dam in *Escherichia coli* gene expression: a global/microarray analysis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4672–4677. doi: 10.1073/pnas.0538053100

Løbner-Olesen, A., Skarstad, K., Hansen, F. G., von Meyenburg, K., and Boye, E. (1989). The DnaA protein determines the initiation mass of *Escherichia coli* K-12. *Cell* 57, 881–889. doi: 10.1016/0092-8674(89)90802-7

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

Margulies, C., and Kaguni, J. M. (1998). The FIS protein fails to block the binding of DnaA protein to oriC, the *Escherichia coli* chromosomal origin. *Nucleic Acids Res.* 26, 5170–5175. doi: 10.1093/nar/26.22.5170

Martínez-Antonio, A., and Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol.* 6, 482–489. doi: 10.1016/j.mib.2003.09.002

McGarry, K. C., Ryan, V. T., Grimwade, J. E., and Leonard, A. C. (2004). Two discriminatory binding sites in the *Escherichia coli* replication origin are required for DNA strand opening by initiator DnaA-ATP. *Proc. Natl. Acad. Sci. U. S. A.* 101, 2811–2816. doi: 10.1073/pnas.0400340101

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: 10.1111/j.1574-6976.2002.tb00620.x

Messer, W., Bellekes, U., and Lother, H. (1985). Effect of dam methylation on the activity of the *E. coli* replication origin, oriC. *EMBO J.* 4, 1327–1332. doi: 10.1002/j.1460-2075.1985.tb03780.x

Messer, W., and Weigel, C. (2003). DnaA as a transcription regulator. *Methods Enzymol.* 370, 338–349. doi: 10.1016/S0076-6879(03)70030-5

Miller, D. T., Grimwade, J. E., Betteridge, T., Rozgaja, T., Torgue, J. J., and Leonard, A. C. (2009). Bacterial origin recognition complexes direct assembly of higher-order DnaA oligomeric structures. *Proc. Natl. Acad. Sci. U. S. A.* 106, 18479–18484. doi: 10.1073/pnas.0909472106

Molina, F., and Skarstad, K. (2004). Replication fork and SeqA focus distributions in *Escherichia coli* suggest a replication hyperstructure dependent on nucleotide metabolism. *Mol. Microbiol.* 52, 1597–1612. doi: 10.1111/j.1365-2958.2004.04097.x

Monteiro, L. M. O., Sanches-Medeiros, A., and Westmann, C. A. (2020). Silva-Rocha, R. Unraveling the complex interplay of Fis and IHF Through synthetic promoter engineering. *Front. Bioeng. Biotechnol.* 8:510. doi: 10.3389/fbioe.2020.00510

Morigen, , Odsbu, I., and Skarstad, K. (2009). Growth rate dependent numbers of SeqA structures organize the multiple replication forks in rapidly growing *Escherichia coli*. *Genes Cells* 14, 643–657. doi: 10.1111/j.1365-2443.2009.01298.x

Nash, H. A., and Robertson, C. A. (1981). Purification and properties of the *Escherichia coli* protein factor required for lambda integrative recombination. *J. Biol. Chem.* 256, 9246–9253. doi: 10.1016/S0021-9258(19)52537-6

Nievera, C., Torgue, J. J., Grimwade, J. E., and Leonard, A. C. (2006). SeqA blocking of DnaA-oriC interactions ensures staged assembly of the *E. coli* pre-RC. *Mol. Cell* 24, 581–592. doi: 10.1016/j.molcel.2006.09.016

Noguchi, Y., Sakiyama, Y., Kawakami, H., and Katayama, T. (2015). The Arg fingers of key DnaA Protomers are oriented inward within the replication origin oriC and stimulate DnaA subcomplexes in the initiation complex. *J. Biol. Chem.* 290, 20295–20312. doi: 10.1074/jbc.M115.662601

Ogawa, T., Yamada, Y., Kuroda, T., Kishi, T., and Moriya, S. (2002). The datA locus predominantly contributes to the initiator titration mechanism in the control of replication initiation in *Escherichia coli*. *Mol. Microbiol.* 44, 1367–1375. doi: 10.1046/j.1365-2958.2002.02969.x

Oka, A., Sasaki, H., Sugimoto, K., and Takanami, M. (1984). Sequence organization of replication origin of the *Escherichia coli* K-12 chromosome. *J. Mol. Biol.* 176, 443–458. doi: 10.1016/0022-2836(84)90171-2

Opel, M. L., Aeling, K. A., Holmes, W. M., Johnson, R. C., Benham, C. J., and Hatfield, G. W. (2004). Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. *Mol. Microbiol.* 53, 665–674. doi: 10.1111/j.1365-2958.2004.04147.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., 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

Polaczek, P. (1990). Bending of the origin of replication of *E. coli* by binding of IHF at a specific site. *New Biol.* 2, 265–271.

Rao, P., Rozgaja, T. A., Alqahtani, A., Grimwade, J. E., and Leonard, A. C. (2018). Low affinity DnaA-ATP recognition sites in *E. coli* oriC make non-equivalent and growth rate-dependent contributions to the regulated timing of chromosome replication. *Front. Microbiol.* 9:1673. doi: 10.3389/fmicb.2018.01673

Riber, L., Frimodt-Møller, J., Charbon, G., and Løbner-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

Rice, P. A. (1997). Making DNA do a U-turn: IHF and related proteins. *Curr. Opin. Struct. Biol.* 7, 86–93. doi: 10.1016/S0959-440X(97)80011-5

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

Roth, A., Urmoneit, B., and Messer, W. (1994). Functions of histone-like proteins in the initiation of DNA replication at oriC of *Escherichia coli*. *Biochimie* 76, 917–923. doi: 10.1016/0300-9084(94)90016-7

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., and Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol. Microbiol.* 82, 475–488. doi: 10.1111/j.1365-2958.2011.07827.x

Russell, D. W., and Zinder, N. D. (1987). Hemimethylation prevents DNA replication in *E. coli*. *Cell* 50, 1071–1079. doi: 10.1016/0092-8674(87)90173-5

Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E., and Leonard, A. C. (2004). *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Mol. Microbiol.* 51, 1347–1359. doi: 10.1046/j.1365-2958.2003.03906.x

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

Sakiyama, Y., Kasho, K., Noguchi, Y., Kawakami, H., and Katayama, T. (2017). Regulatory dynamics in the ternary DnaA complex for initiation of chromosomal replication in *Escherichia coli*. *Nucleic Acids Res.* 45, 12354–12373. doi: 10.1093/nar/gkx914

Samitt, C. E., Hansen, F. G., Miller, J. F., and Schaechter, M. (1989). *In vivo* studies of DnaA binding to the origin of replication of *Escherichia coli*. *EMBO J.* 8, 989–993. doi: 10.1002/j.1460-2075.1989.tb03462.x

Schaechter, M., Maaloe, O., and Kjeldgaard, N. O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced grown of salmonella typhimurium. *J. Gen. Microbiol.* 19, 592–606. doi: 10.1099/00221287-19-3-592

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

Schmid, M. B. (1990). More than just "histone-like" proteins. *Cell* 63, 451–453. doi: 10.1016/0092-8674(90)90438-K

Schneider, R., Travers, A., Kutateladze, T., and Muskhelishvili, G. (1999). A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol. Microbiol.* 34, 953–964. doi: 10.1046/j.1365-2958.1999.01656.x

Schneider, R., Travers, A., and Muskhelishvili, G. (1997). FIS modulates growth phase-dependent topological transitions of DNA in *Escherichia coli*. *Mol. Microbiol.* 26, 519–530. doi: 10.1046/j.1365-2958.1997.5951971.x

Seshasayee, A. S., Sivaraman, K., and Luscombe, N. M. (2011). An overview of prokaryotic transcription factors: a summary of function and occurrence in bacterial genomes. *Subcell. Biochem.* 52, 7–23. doi: 10.1007/978-90-481-9069-0_2

Shimizu, M., Noguchi, Y., Sakiyama, Y., Kawakami, H., Katayama, T., and Takada, S. (2016). Near-atomic structural model for bacterial DNA replication initiation complex and its functional insights. *Proc. Natl. Acad. Sci. U. S. A.* 113, E8021–E8030. doi: 10.1073/pnas.1609649113

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

Skarstad, K., and Katayama, T. (2013). Regulating DNA replication in bacteria. *Cold Spring Harb. Perspect. Biol.* 5:a012922. doi: 10.1101/cshperspect.a012922

Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K., and Kleckner, N. (1995). *E. coli* SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell* 82, 927–936. doi: 10.1016/0092-8674(95)90272-4

Slomińska, M., Wegrzyn, A., Konopa, G., Skarstad, K., and Wegrzyn, G. (2001). SeqA, the *Escherichia coli* origin sequestration protein, is also a specific transcription factor. *Mol. Microbiol.* 40, 1371–1379. doi: 10.1046/j.1365-2958.2001.02480.x

Speck, C., and Messer, W. (2001). Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *EMBO J.* 20, 1469–1476. doi: 10.1093/emboj/20.6.1469

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

Stepankiw, N., Kaidow, A., Boye, E., and Bates, D. (2009). The right half of the *Escherichia coli* replication origin is not essential for viability, but facilitates multi-forked replication. *Mol. Microbiol.* 74, 467–479. doi: 10.1111/j.1365-2958.2009.06877.x

Sugiyama, R., Kasho, K., Miyoshi, K., Ozaki, S., Kagawa, W., Kurumizaka, H., et al. (2019). A novel mode of DnaA-DnaA interaction promotes ADP dissociation for reactivation of replication initiation activity. *Nucleic Acids Res.* 47, 11209–11224. doi: 10.1093/nar/gkz795

Sutera, V. A., and Lovett, S. T. (2006). The role of replication initiation control in promoting survival of replication fork damage. *Mol. Microbiol.* 60, 229–239. doi: 10.1111/j.1365-2958.2006.05093.x

Taghbalout, A., Landoulsi, A., Kern, R., Yamazoe, M., Hiraga, S., Holland, B., et al. (2000). Competition between the replication initiator DnaA and the sequestration factor SeqA for binding to the hemimethylated chromosomal

origin of *E. coli in vitro*. *Genes Cells* 5, 873–884. doi: 10.1046/j.1365-2443.2000.00380.x

Theisen, P. W., Grimwade, J. E., Leonard, A. C., Bogan, J. A., and Helmstetter, C. E. (1993). Correlation of gene transcription with the time of initiation of chromosome replication in *Escherichia coli*. *Mol. Microbiol.* 10, 575–584. doi: 10.1111/j.1365-2958.1993.tb00929.x

Torheim, N., and Skarstad, K. (1999). K. *Escherichia coli* SeqA protein affects DNA topology and inhibits open complex formation at oriC. *EMBO J.* 18, 4882–4888. doi: 10.1093/emboj/18.17.4882

Travers, A. (1997). DNA-protein interactions: IHF--the master bender. *Curr. Biol.* 7, R252–R254. doi: 10.1016/S0960-9822(06)00114-X

Travers, A., Schneider, R., and Muskhelishvili, G. (2001). DNA supercoiling and transcription in *Escherichia coli*: The FIS connection. *Biochimie* 83, 213–217. doi: 10.1016/S0300-9084(00)01217-7

Von Freiesleben, U., Rasmussen, K. V., Atlung, T., and Hansen, F. G. (2000). Rifampicin-resistant initiation of chromosome replication from oriC in ihf mutants. *Mol. Microbiol.* 37, 1087–1093. doi: 10.1046/j.1365-2958.2000.02060.x

Waldminghaus, T., and Skarstad, K. (2009). The *Escherichia coli* SeqA protein. *Plasmid* 61, 141–150. doi: 10.1016/j.plasmid.2009.02.004

Wang, G., and Maier, R. J. (2015). Bacterial histone-like proteins: roles in stress resistance. *Curr. Genet.* 61, 489–492. doi: 10.1007/s00294-015-0478-x

Washington, T. A., Smith, J. L., and Grossman, A. D. (2017). Genetic networks controlled by the bacterial replication initiator and transcription factor DnaA in *Bacillus subtilis*. *Mol. Microbiol.* 106, 109–128. doi: 10.1111/mmi.13755

Wolanski, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwińska, J. (2014). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front. Microbiol.* 5:735. doi: 10.3389/fmicb.2014.00735

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

Wold, S., Skarstad, K., Steen, H. B., Stokke, T., and Boye, E. (1994). The initiation mass for DNA replication in *Escherichia coli* K-12 is dependent on growth rate. *EMBO J.* 13, 2097–2102. doi: 10.1002/j.1460-2075.1994.tb06485.x

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, 471–481. doi: 10.1042/BJ20050143

Zheng, H., Bai, Y., Jiang, M., Tokuyasu, T. A., Huang, X., Zhong, F., et al. (2020). General quantitative relations linking cell growth and the cell cycle in *Escherichia coli*. *Nat. Microbiol.* 5, 995–1001. doi: 10.1038/s41564-020-0717-x

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 27 September 2021
doi: 10.3389/fmicb.2021.733085

The Stress-Active Cell Division Protein ZapE Alters FtsZ Filament Architecture to Facilitate Division in *Escherichia coli*

*Eric C. DiBiasio, Rebecca A. Dickinson, Catherine E. Trebino, Colby N. Ferreira, Josiah J. Morrison and Jodi L. Camberg**

Department of Cell and Molecular Biology, The University of Rhode Island, Kingston, RI, United States

OPEN ACCESS

Edited by:
Jianping Xie,
Southwest University, China

Reviewed by:
Matthew Cabeen,
Oklahoma State University, United States
Handuo Shi,
Stanford University, United States
Maria A. Oliva,
Consejo Superior de Investigaciones Científicas (CSIC), Spain

***Correspondence:**
Jodi L. Camberg
cambergj@uri.edu

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *29 June 2021*
Accepted: *25 August 2021*
Published: *27 September 2021*

Citation:
DiBiasio EC, Dickinson RA, Trebino CE, Ferreira CN, Morrison JJ and Camberg JL (2021) The Stress-Active Cell Division Protein ZapE Alters FtsZ Filament Architecture to Facilitate Division in Escherichia coli. Front. Microbiol. 12:733085. doi: 10.3389/fmicb.2021.733085

During pathogenic infections, bacterial cells experience environmental stress conditions, including low oxygen and thermal stress. Bacterial cells proliferate during infection and divide by a mechanism characterized by the assembly of a large cytoskeletal structure at the division site called the Z-ring. The major protein constituting the Z-ring is FtsZ, a tubulin homolog and GTPase that utilizes the nucleotide to assemble into dynamic polymers. In *Escherichia coli*, many cell division proteins interact with FtsZ and modulate Z-ring assembly, while others direct cell wall insertion and peptidoglycan remodeling. Here, we show that ZapE, an ATPase that accumulates during late constriction, directly interacts with FtsZ and phospholipids *in vitro*. In the presence of adenosine triphosphate (ATP), ZapE induces bundling of GTP-induced FtsZ polymers; however, ZapE also binds FtsZ in the absence of GTP. The ZapE mutant protein ZapE(K84A), which is defective for ATP hydrolysis, also interacts with FtsZ and induces FtsZ filament bundling. *In vivo*, cultures of *zapE* deletion cells contain a low percentage of filamentous cells, suggesting that they have a modest division defect; however, they are able to grow when exposed to stress, such as high temperature and limited oxygen. When combined with the chromosomal deletion of *minC*, which encodes an FtsZ disassembly factor, Δ*zapE* Δ*minC* cells experience growth delays that slow proliferation at high temperature and prevent recovery. This synthetic slow growth phenotype after exposure to stress suggests that ZapE may function to ensure proliferation during and after stress, and this is exacerbated when cells are also deleted for *minC.* Expression of either ZapE or ZapE(K84A) complements the aberrant growth phenotypes *in vivo* suggesting that the division-associated role of ZapE does not require ZapE ATP hydrolysis. These results support that ZapE is a stress-regulated cell division protein that interacts directly with FtsZ and phospholipids, promoting growth and division after exposure to environmental stress.

Keywords: cytokinesis, constriction, ATPase, AAA+, stress, Z-ring-associated protein

INTRODUCTION

Proliferation via the highly conserved bacterial cell division pathway requires more than 30 different proteins acting in a well-coordinated manner (de Boer, 2010). While many of the individual cell division proteins involved have been characterized, knowledge of their functional interactions remains incomplete, including the interactions of late stage division proteins and those that direct peptidoglycan insertion or remodeling (Soderstrom et al., 2018, 2019). In *Escherichia coli* and most bacteria, cell division requires the highly conserved tubulin-like FtsZ protein (Bi and Lutkenhaus, 1991). FtsZ is a GTPase, and binding of GTP to FtsZ protomers leads to the formation of linear protofilaments arranged in a head-to-tail orientation (Du and Lutkenhaus, 2019). FtsZ protofilaments comprise the Z-ring at midcell during constriction and may act as a scaffold for assembly of the divisome (Soderstrom and Daley, 2017). A previous publication reported the presence of cell division protein ZapE in Gram-negative bacteria, including *E. coli*, and showed that ZapE was important for cell division at high temperature (42°C) and under anaerobic conditions (Marteyn et al., 2014). Moreover, ZapE was identified as essential in *Shigella flexneri* for colonization of the gastrointestinal tract (Marteyn et al., 2014).

FtsZ assembly into linear polymers and the dynamic exchange of FtsZ subunits at the Z-ring is regulated by many FtsZ-interacting proteins, including the "Z-ring-associated proteins" (Zaps). ZapA, -B, -C, and -D localize to the Z-ring early in the cell division pathway and regulate the development of the Z-ring (Schumacher et al., 2017). ZapE was characterized as a Z-associated protein, but ZapE is also a member of the AAA + ATPase family (ATPase associated with a variety of cellular activities) and contains both Walker A $_{78}$(GGVGRGKT)$_{85}$ and Walker B $_{141}$(CFDEF)$_{145}$ nucleotide-binding sites (Marteyn et al., 2014). The requirement for ZapE in *E. coli* grown under anaerobic and thermal stress conditions suggests that it may be a conditional cell division protein that is regulated at the transcriptional and/or translational levels in response to environmental conditions, or that it may become active *in vivo* during exposure to stress.

The mechanism by which ZapE impacts the cell division pathway through its direct interaction with FtsZ is largely unknown. ZapE is a late stage cell division protein that was reported to influence FtsZ polymerization *in vitro* by destabilizing FtsZ polymers in an adenosine triphosphate (ATP)-dependent manner (Marteyn et al., 2014). Another AAA + ATPase that has also been shown to influence FtsZ polymer assembly and destabilize FtsZ polymers is ClpX (Camberg et al., 2009; Viola et al., 2017). ClpX is a hexameric ATP-dependent unfoldase that, along with the serine protease ClpP, forms the two-component bacterial proteasome complex ClpXP. ClpXP processively degrades specific protein substrates, including FtsZ, after initiating ATP-dependent unfolding from the N- or C-terminus. During *E. coli* division, ClpXP degradation of polymerized FtsZ reduces polymer abundance and modifies the dynamic exchange of FtsZ subunits in the Z-ring by a polymer severing mechanism (Camberg et al., 2009, 2014; Viola et al., 2017). Most notably, however, ClpXP was previously reported to prevent the accumulation of intracellular FtsZ aggregates, which accumulate in *E. coli* grown under thermal stress conditions, including growth at 42°C (Tomoyasu et al., 2001; LaBreck et al., 2017). Other stress-induced protein chaperones, such as ClpB and HtpG, have also been implicated in modulating FtsZ aggregates and polymers (LaBreck et al., 2017; Balasubramanian et al., 2019).

To gain mechanistic insight into the functional roles and interactions of ZapE in *E. coli*, we monitored biochemical activities and direct interactions with polymerized and non-polymerized FtsZ *in vitro*. Our results clearly demonstrate a direct interaction between the two proteins and also show that ZapE can recruit a chimeric FtsZ protein to phospholipid (PL) vesicles. ZapE bundles FtsZ polymers in the presence of ATP, which is in contrast to disassembly activity that was previously reported (Marteyn et al., 2014). Consistent with the previous report, we also show that deletion of *zapE* induces cell filamentation in a small proportion of cells. Interestingly, we also show that deletion of *zapE* slows overall growth of a *minC* deletion strain, and this is complemented by reinserting wild-type *zapE* or *zapE(K84A)* back into the native locus. MinC is an FtsZ polymerization inhibitor that prevents division near the cell poles (Hu et al., 1999; Hu and Lutkenhaus, 2000; Dajkovic et al., 2008). Together, our results support that ZapE is critical for division under thermal and oxygen stress and may function to organize and maintain active populations of FtsZ *in vivo* to facilitate division.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions

Escherichia coli strains, listed in **Table 1**, were grown in LB (Lennox), Luria–Bertani broth with NaCl 5 g/L, media and/or LB (no salt) media and agar, plus ampicillin (100 μg ml^{-1}), or kanamycin (50 μg ml^{-1}), where indicated. Gene deletions were constructed by Lambda red recombination as described (Datsenko and Wanner, 2000), and chromosomal in-frame gene replacements with *zapE* or *zapE(K84A)* were constructed as described (LaBreck et al., 2019). Culture densities were monitored and measured by OD_{600}. For protein purification, *zapE* and *zapE*(K84A) expression plasmids were constructed by restriction cloning (*Nde*I, *Hin*dIII) and site-directed mutagenesis into pET-28a expression vector with an N-terminal 6 × -histidine tag.

Protein Expression and Purification

Vectors containing *zapE* and *zapE(K84A)* with an N-terminal 6 × -histidine tag were transformed and overexpressed in *E. coli* BL21 (λde3). Cell cultures were grown to an OD_{600} of 0.8 and induced with IPTG (0.5 mM) for 4 h at 30°C. Cultures were centrifuged, and pellets were lysed by French press in 25 mM Tris (pH 7.5) buffer containing 0.1 mM TCEP, 10 mM $MgCl_2$, 150 mM KCl, and 10% glycerol. Soluble cell extracts were collected by centrifugation at 35,000 × *g* for 30 min, applied to

TABLE 1 | *E. coli* strains and plasmids used for genetic analyses and constructions in this study.

E. coli strain or plasmid	Relevant genotype description	Source, reference or construction
Strains		
MG1655	*LAM-rph-1*	Blattner et al., 1997
BL21 (λde3)	F-*ompT gal dcm lon hsdSB(rB- mB-)* λ(de3[*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*])	EMD Millipore, United States
ED0011	MG1655 Δ*zapE*:*parE-kan*	MG1655; λred
JC0232	MG1655 Δ*minC*:*frt*	Camberg et al., 2011
ED0024	MG1655 Δ*minC*:*frt*, Δ*zapE*: *parE-kan*	MG1655 Δ*minC*:*frt*; λred
ED0118	MG1655 Δ*minC*:*frt* Δ*zapE*:*zapE*	MG1655 Δ*minC*:*frt* Δ*zapE*: *parE-kan*; λred
ED0131	MG1655Δ*minC*:*frt* Δ*zapE*:*zapE(K84A)*	MG1655 Δ*minC*:*frt* Δ*zapE*: *parE-kan*; λred
Plasmids		
pKD46	*Amp*, *recombinase*	Datsenko and Wanner, 2000
pKD267	*parE-kan, tet*	LaBreck et al., 2019
pET-H_6-ZapE	*kan*, *His_6-zapE*	This study
pET-ZapE-H_6	*kan*, *zapE-His_6*	This study
pET-H_6-ZapE(K84A)	*kan*, *His_6-zapE(K84A)*	This study
pET-Gfp-FtsZ	*kan*, *His_6-Gfp-ftsZ*	Viola et al., 2017
pET-FtsZ	*kan*, FtsZ	Camberg et al., 2009

a cobalt column, and eluted with an imidazole gradient (0 mM to 150 mM). Collected protein was fractionated on a sephadex G25 column to remove imidazole and small contaminants. Native FtsZ was expressed in *E. coli* BL21 (λde3) and purified as described (Camberg et al., 2009). ZapE containing a C-terminal 6 × -histidine tag was purified as described above. Histidine-tagged Gfp-FtsZ, which contains green fluorescent protein (Gfp) fused to the N-terminus of FtsZ, was expressed in *E. coli* BL21 (λde3) and purified as described (Viola et al., 2017). Where indicated, size exclusion chromatography was performed on purified H_6-ZapE. Briefly, H_6-ZapE was incubated at room temperature with or without ATP (5 mM) for 5 min and then passed through a Sephacryl S-200 (25 ml) column with 25 mM Tris (pH 7.5) buffer containing 0.1 mM TCEP, 10mM $MgCl_2$, 150 mM KCl, and 10% glycerol, and, where indicated, equilibrated with ATP (0.5 mM). Eluting fractions (250 μl) were collected (0.5 ml min^{-1}) and analyzed by the Bradford protein assay and the proteins were visualized by SDS-PAGE and Coomassie staining.

Nucleotide Hydrolysis Assays

Hydrolysis of ATP or GTP, where indicated, by ZapE or FtsZ in 50 mM HEPES buffer, pH 7.5, containing 150 mM KCl and 20 mM $MgCl_2$ in the absence or presence of SUVs (250 μg ml^{-1}), where indicated, was monitored by measuring the release of inorganic phosphate with time using the Biomol Green phosphate detection reagent (Enzo Life Sciences) at 23°C as described (Camberg et al., 2014). Free phosphate was quantitated by comparison with a phosphate standard curve. Where indicated, ATP concentration was titrated in hydrolysis reactions, and the results were analyzed in GraphPad Prism (version 8) by fitting data to the Michaelis–Menten equation.

FtsZ Sedimentation Assays

FtsZ polymers were assembled by incubating FtsZ (6 μM) in 50 mM HEPES buffer, pH 7.5, containing 150 mM KCl, 20 mM $MgCl_2$ with GTP (2 mM) in the absence and presence of ZapE or ZapE(K84A) (8 μM) with or without ATP or ATPγS (4 mM), where indicated, at 23°C for 10 min and collected by centrifugation for 20 min at 250,000 × *g*, Supernatants and pellets were resuspended in equivalent volumes of LDS sample buffer (Life Technologies) and analyzed by SDS-PAGE and Coomassie staining.

Light Scattering Assays

Light scattering (90° angle) was performed in 50 mM MES (pH 6.5) buffer containing 100 mM KCl and 10 mM $MgCl_2$. Where indicated, reaction mixtures (80 μl) containing ZapE (0 to 8 μM) or ZapE(K84A) (4 μM), with or without FtsZ (8 μM), were monitored at 23°C using an Agilent Eclipse fluorescence spectrophotometer with both excitation and emission wavelengths set to 450 nm with 5 nm slit widths, as described (Conti et al., 2018). Protein reactions were incubated at room temperature for 5 min, and where indicated, with ATP (4 mM) or ATPγS (2 mM). Baseline readings were collected for 5 min, GTP (1.5 mM) was added, and light scattering was measured at 23°C for up to 120 min, where indicated.

Binding and Retention Assay

Binding of ZapE to FtsZ was assayed by dot blot. Briefly, bovine serum albumin (BSA) (10 μM), ZapE (10 μM), and FtsZ (10 μM) alone or with GDP (2 mM) or GTP (2 mM) were spotted (3 μl) onto a nitrocellulose membrane. The membrane was blocked with BSA (2%) in tris-buffered saline (TBS) then incubated with ZapE (8 μM) and ATP (4 mM) in 50 mM HEPES (pH 7.5) buffer containing 150 mM KCl and 20 mM $MgCl_2$ at 23°C for 1 h to allow binding. After washing with TBS containing Tween-20 (0.01%), the membrane was probed with rabbit polyclonal antisera (anti-ZapE) and then goat anti-rabbit antibody coupled to horseradish peroxidase. The membrane was developed by chemiluminescence, and spot density was analyzed by NIH ImageJ. Anti-ZapE antiserum was generated in rabbits using purified H_6-ZapE (ThermoFisher).

Filter retention assays were performed to detect interactions between fluorescent Gfp-FtsZ and ZapE *in vitro*. ZapE (5 μM), ZapE(K84A) (5 μM), and/or Gfp-FtsZ (10 μM) reaction mixtures (40 μl), in 50 mM MES (pH 6.5) buffer containing 100 mM KCl and 10 mM $MgCl_2$, were incubated with ATP (4 mM) and, where indicated, GTP (2 mM) at 23°C for 10 min. Reactions were filtered through 100 kDa Nanosep polyethersulfone filters by centrifugation at 16,900 × *g* for 20 min. Protein present in filtrate and retentate were analyzed by SDS-PAGE and Coomassie staining, and densitometry was performed using NIH ImageJ.

Phospholipid Recruitment Assays

Phospholipid recruitment assays with ZapE were performed by incubating ZapE (1 μM) or ZapE K84A (1 μM) together with small unilamellar vesicles (SUVs) (500 μg ml^{-1}) for 5 min at 23°C in the presence or absence of ATP (4 mM), where indicated in 50 mM HEPES (pH 7.5) reaction buffer containing 150 mM KCl and 20 mM $MgCl_2$. Phospholipid recruitment assays with ZapE and Gfp-FtsZ (3 μM) were performed by incubating Gfp-FtsZ with GTP (4 mM), where indicated, in reaction buffer for 2 min and then added to a reaction containing ZapE (1 μM) pre-assembled with SUVs (500 μg ml^{-1}) and ATP (4 mM) and then incubated for an additional 2 min at room temperature. Phospholipid vesicles and associated proteins were collected by low-speed centrifugation at 21,000 × *g* for 15 min. Supernatants and pellets were analyzed by SDS-PAGE and Coomassie staining. Percent of protein in the pellet fraction was determined by densitometry using NIH ImageJ.

Microscopy

MG1655, MG1655 Δ*zapE*, MG1655 Δ*zapE*:*zapE, and* MG1655 Δ*zapE*:*zapE(K84A)* strains were grown overnight in LB (Lennox) media at 30°C. The next day, cultures were diluted into fresh LB (Lennox) medium standardized to OD_{600} of 0.05, grown for 3 h at 35°C, and then the final OD_{600} was measured. Cells were then harvested by centrifugation at 3,000 × *g* for 5 min, resuspended in 100 μl of 1 × PBS with EDTA (1 mM), and fixed to glass slides with poly-$_L$-lysine-coated coverslips or, where indicated, applied to glass slides with agarose gel pads (5%, wt/vol) containing M9 minimal medium supplemented with 0.4% glucose. Cells were imaged by DIC microscopy using a Zeiss AxioCam HRc high-resolution camera, processed using Adobe photoshop CS6, and measured using ImageJ software.

For electron microscopy imaging, reactions containing buffer (50 mM MES pH 6.5, 100 mM KCl, 10 mM $MgCl_2$), ZapE (8 μM), ZapE (K84A) (8 μM), FtsZ (8 μM), and ATP (4 mM), where indicated, were incubated for 5 min at 23°C, then GTP (1.5 mM), was added to induce complex formation, where indicated, and reactions were incubated for an additional 15 min at room temperature. Samples were applied to a 300-mesh carbon/formvar coated grid, fixed with glutaraldehyde (2.5%), and negatively stained with uranyl acetate (2%). Samples were visualized at × 10,000— × 25,000 direct magnification by transmission electron microscopy using a JEM-2100 80 Kev instrument.

Colony Growth Assays

MG1655, MG1655 Δ*zapE*, MG1655 Δ*minC*, MG1655 Δ*minC*Δ*zapE*, MG1655 Δ*minC*Δ*zapE*:*zapE*, *and* MG1655 Δ*minC*Δ*zapE*:*zapE(K84A)* strains were grown overnight in LB (Lennox) medium at 30°C. The next day, cultures were diluted into fresh LB (no salt) or LB (Lennox) NaCl 5 g/L medium, where indicated, standardized to an OD_{600} of 0.05 and grown for 3 h under mild heat shock conditions in static, closed culture at 42°C, or aerobically at 37°C and shaking. After the allotted time, growth was measured by OD_{600}, and cultures were diluted (2-log) into LB (no salt) medium. Dilutions were spot plated (5 μl) onto Lennox agar LB (no salt) agar, and plates were incubated at room temperature. Images of colony growth were collected, and biomass was quantified in ImageJ software to determine spot density in the zone of growth.

Western Blotting

MG1655, MG1655 Δ*zapE*, MG1655 Δ*minC*, MG1655 Δ*minC*Δ*zapE*, MG1655 Δ*minC*Δ*zapE*:*zapE*, *and* MG1655 Δ*minC*Δ*zapE*:*zapE(K84A)* strains were grown overnight in LB (Lennox) media at 30°C. The next day, cultures were diluted into fresh LB (no salt) to an OD_{600} of 0.05 and grown for 3 h under mild heat shock conditions in static, closed cultures at 42°C, or aerobically at 37°C with shaking. Cells were harvested, normalized to total number of cells for SDS-PAGE, and then cell extracts were analyzed by Western blotting. Membranes were blocked with BSA (2%) in TBS containing Tween-20 (0.005%), and ZapE was detected with rabbit polyclonal antisera raised against ZapE followed by goat anti-rabbit antibody coupled to horseradish peroxidase. Membranes were developed using chemiluminescence.

RESULTS

ZapE Interacts With FtsZ *In vitro* and Bundles FtsZ Polymers in an ATP-Dependent Manner

Escherichia coli ZapE was previously reported to engage FtsZ *in vitro* and localize to the Z-ring *in vivo* (Marteyn et al., 2014). To biochemically characterize the enzymatic activity and interaction between ZapE and FtsZ further, we purified ZapE as an N-terminal six histidine fusion protein by metal affinity chromatography and then assayed ZapE for ATP hydrolysis activity and a direct interaction with FtsZ (**Figure 1A** and **Supplementary Figure 1A**). First, we titrated substrate (ATP) concentration from 0 to 5 mM in reactions containing ZapE (12 μM) to determine the maximal reaction velocity (V_{max}). The reaction velocity plateaued near ∼1 pmol Pi min^{-1} $pmol^{-1}$. Curve fitting to the Michaelis–Menten equation calculated a V_{max} of 1.1 ± 0.09 pmol Pi min^{-1} $pmol^{-1}$ with a K_m of 1.7 ± 0.36 mM, under the conditions tested (**Figure 1A**). To determine the oligomeric state of ZapE, we fractionated ZapE by size exclusion chromatography in the absence and presence of ATP, and calculated the size of ZapE to be ∼40 kDa under both conditions based on the elution volume of the peak

fraction (**Supplementary Figures 1B,C**), which is consistent with the calculated molecular mass of 44,583 Da. However, we also observed that the elution peak was broader with an early eluting shoulder in the presence of ATP, which could suggest a minor population of dimers in equilibrium with the predominant monomer fraction (**Supplementary Figures 1B,C**). We detected no apparent positive cooperativity upon titrating protein or substrate concentration in ATP hydrolysis assays (**Figure 1A** and **Supplementary Figure 1D**), which is consistent with ZapE functioning as a monomer in these reactions. Notably, we also constructed a C-terminally tagged six histidine fusion protein, but it hydrolyzed ATP at an approximately 65% slower rate (0.31 ± 0.02 pmol Pi min^{-1} $pmol^{-1}$) than ZapE with the N-terminal histidine tag, suggesting that the addition of a tag sequence to the C-terminus of ZapE may interfere with ATPase activity (**Supplementary Figure 1E**). A detailed structural model of the complete ZapE protein is unavailable; therefore, to analyze ZapE functional regions in further detail, we modeled ZapE onto DnaA, which has 44% similarity across a region of ZapE (amino acids 74 through 181) that overlaps the Walker A motif (**Figure 1B**) (Ozaki et al., 2008). The substitution mutation K84A in the Walker A motif of ZapE was previously reported to impair ATP hydrolysis *in vitro* by thin layer chromatography; therefore, we purified ZapE(K84A) as an N-terminal histidine fusion protein to quantitatively measure ATP hydrolysis. We observed no phosphate released in reactions containing ZapE(K84A) (6 μM) and ATP under the conditions tested, in contrast to wild-type ZapE (**Figure 1C**). Together, these results demonstrate that ZapE is an ATPase and substitution of Lys 84 with Ala prevents activity, consistent with a previous report (Marteyn et al., 2014).

ZapE was implicated in regulating FtsZ polymerization in qualitative fluorescence-based microscopy assays *in vitro* (Marteyn et al., 2014), suggesting that there is a direct interaction with polymerized FtsZ. To detect the interaction, we first tested if ZapE co-pellets with FtsZ polymers assembled with GTP. We incubated ZapE with FtsZ in the presence of GTP and collected FtsZ polymers by ultracentrifugation. After fractionating supernatants and pellets, we analyzed both fractions by SDS-PAGE for the relative amounts of FtsZ and ZapE. We observed that FtsZ, incubated with GTP to promote polymerization, fractionates with the pellet, consistent with the assembly of GTP-dependent polymers (**Figure 1D**). When polymerized FtsZ was incubated with ZapE, pellet fractions also contained ZapE along with FtsZ. In a control reaction without FtsZ, only a small amount of ZapE was detected in the centrifugation pellet. Together, these results suggest that ZapE interacts with FtsZ polymers. Both protein bands migrated very closely by SDS-PAGE due to their size similarity complicating any quantitation by densitometry. When ATP or ATPγS was included in the reaction, the amount of ZapE co-pelleting with FtsZ and GTP increased, relative to reactions performed in the absence of ATP. These results show that ZapE co-pellets with FtsZ and GTP, and that ATP increases the amount of ZapE in the pellet with FtsZ. Next, as a control, we confirmed that FtsZ sediments only in the presence of GTP and does not fractionate with the pellet when GTP is omitted, consistent with specific GTP-dependent polymerization of FtsZ (**Figure 1E**). Finally, we tested the ZapE mutant protein ZapE(K84A), which is defective for ATP hydrolysis, and observed that it also co-pellets with FtsZ and GTP (**Figure 1E**).

To directly observe the effects of ZapE on FtsZ polymerization in real time *in vitro*, we used 90° light scattering to monitor GTP-dependent polymerization. First, we pre-incubated FtsZ (8 μM) with ZapE (0 to –8 μM) in the presence of ATP (4 mM) for 5 min, added GTP (1.5 mM) to induce polymerization, and monitored the change in light scatter with time. We observed that as ZapE concentration increases, the light scattered and maximum intensity recorded also increased, with a greater than twofold increase in maximum light scatter comparing ZapE (0 μM) and ZapE (8 μM) (**Figures 2A,B**). The increase in light scatter by FtsZ has been reported to correspond to the accumulation of FtsZ polymers after addition of GTP (Mukherjee and Lutkenhaus, 1994, 1999; Camberg et al., 2014). In a control assay, no increase in light scatter was detected for ZapE alone after addition of GTP (**Supplementary Figure 2A**). We also performed a longer time course of the FtsZ polymer assembly reaction to determine if polymers persisted or rapidly disassembled with ZapE. We observed that the kinetics of disassembly, which can be observed during minutes 40 through 90, appear similar with and without ZapE, although large complexes did persist longer with ZapE (**Supplementary Figure 2B**).

To determine if ZapE slows the GTPase activity of FtsZ, like other *E. coli* proteins reported to bundle FtsZ polymers (Gueiros-Filho and Losick, 2002; Low et al., 2004; Small et al., 2007; Hale et al., 2011; Durand-Heredia et al., 2012), we measured the rate of GTP hydrolysis in reactions containing FtsZ (6 μM) in the absence and presence of ZapE (12 μM). The GTP hydrolysis rate of FtsZ was determined to be 1.00 ± 0.28 pmol P_i min^{-1} $pmol^{-1}$ under the conditions tested, and the rate did not significantly change in the presence of ZapE 1.13 ± 0.31 pmol P_i min^{-1} $pmol^{-1}$ (**Supplementary Figure 2C**). Additionally, ZapE is unable to use GTP as a substrate under the conditions tested (data not shown). We also tested if FtsZ modifies ZapE ATPase activity but detected no change in the rate of ATP hydrolysis by ZapE when FtsZ was included in the reaction (**Supplementary Figure 2D**). Together, these results show that addition of ZapE to FtsZ polymerization reactions increases the scatter of protein complexes containing ZapE and FtsZ, suggesting the ZapE may bundle or functionally crosslink FtsZ polymers. However, we did not detect that ZapE can modify the GTP turnover by FtsZ under the conditions tested.

Next, to determine if ATP is necessary for the observed increase in light scatter stimulated by ZapE in FtsZ polymerization reactions, we repeated the assay without ATP. We did not detect ZapE-induced stimulation of GTP-dependent light scatter when ATP was omitted from FtsZ polymerization reactions containing ZapE (**Figure 2C**). Next, we tested ATPγS (2 mM) in ZapE and FtsZ 90° light scattering assays (**Figure 2C**) and observed a modest increase in light scatter compared with the increase in the absence of ATP (**Figure 2C**). Finally, we also tested if ZapE(K84A) stimulates FtsZ light scatter in GTP-dependent polymerization assays and observed an increase similar to wild-type ZapE with ATP (**Figure 2D**). These results suggest that nucleotide binding

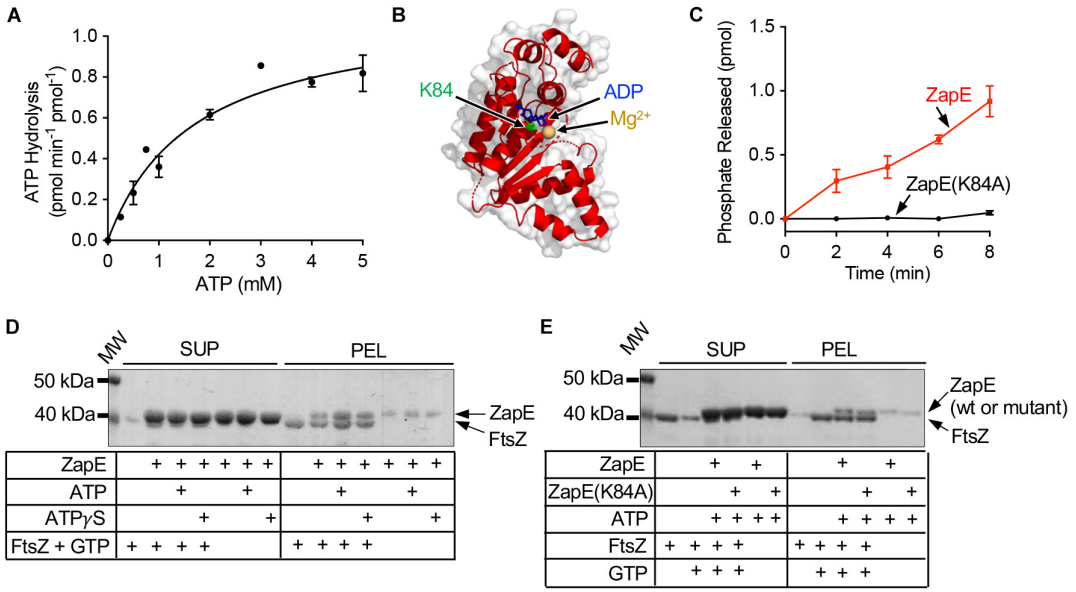


FIGURE 1 | ZapE hydrolyzes ATP and co-pellets with FtsZ polymers *in vitro*. **(A)** ZapE (12 μM) hydrolysis rates were measured by monitoring the release of phosphate with increasing ATP concentrations (0 to 5 mM) as described in *Materials and methods* section. Data shown are the average of at least three replicates represented as mean ± SEM. **(B)** Structural model of *E. coli* ZapE residues (51–275) modeled onto DnaA (pdb 2Z4R) (Ozaki et al., 2008) with nucleotide shown in blue (ADP). Lys 84 is shown as green Corey–Pauling–Koltun model. **(C)** Hydrolysis of ATP (4 mM) by ZapE (6 μM) and ZapE(K84A) (6 μM) was monitored with time (0, 2.0, 4.0, 6.0, and 8.0 min) and phosphate released (pmol) was reported. Data shown are the average of at least three replicates represented as mean ± SEM. **(D)** FtsZ (6 μM) polymers were assembled with GTP (2 mM) in the absence and presence of ZapE (8 μM) alone and with ATP or ATPγS (4 mM). FtsZ polymers were collected by ultracentrifugation, as described in *Materials and methods* section, and contents of supernatants and pellets were analyzed by SDS-PAGE to detect ZapE. Data shown are representative of at least three replicates. **(E)** FtsZ (6 μM) was incubated without GTP, where indicated, or assembled into polymers with GTP (2 mM) in the absence and presence of ZapE (8 μM) or ZapE(K84A) (6 μM) with ATP (4 mM). FtsZ polymers were collected by ultracentrifugation, as described in *Materials and methods* section, and contents of supernatants and pellets were analyzed by SDS-PAGE to detect ZapE. Data shown are representative of at least three replicates.

by ZapE, but not hydrolysis, is sufficient to induce assembly of large complexes with FtsZ polymers. To determine if the large complexes that we detected are bundled FtsZ polymers, we performed transmission electron microscopy (TEM) and negative staining of FtsZ polymers assembled with GTP in the absence and presence of ZapE and ATP, the condition that increases light scatter (**Figure 2E**). We observed that without ZapE, FtsZ polymers were long and thin, and many polymers detected appeared to be single-stranded with some thicker filaments present (**Figure 2E**). In contrast, in the reaction containing FtsZ polymers incubated with ZapE, polymers were grouped together as large bundles of approximately 100 nm in width, and no polymers or bundles were observed by ZapE without FtsZ (**Figure 2E**). Bundles were also observed when FtsZ polymers were incubated with ZapE(K84A), which is consistent with light scatter assays (**Figures 2D,E**).

Polymerization and light scattering assays both indicated that ZapE directly binds to FtsZ polymers; however, to determine if ZapE also binds to FtsZ in the absence of GTP, we performed dot blot binding assays by applying purified FtsZ to a nitrocellulose membrane in the presence of various nucleotides and then monitored ZapE recruitment. Briefly, BSA (10 μM), ZapE (10 μM), FtsZ (10 μM), FtsZ (10 μM) with GDP (2 mM), and FtsZ (10 μM) with GTP (2 mM) were applied as a spot to a nitrocellulose membrane and then blocked with BSA. A solution of ZapE (8 μM) with ATP (4 mM) was incubated with the membrane to allow for potential binding. Binding was then probed with a ZapE antibody. Our results show that ZapE binds FtsZ similarly under all conditions tested and that ZapE does not bind to the control spot containing BSA (**Figure 2F**). These results suggest that ZapE binds to both the unpolymerized and polymerized states of FtsZ. To more quantitatively compare ZapE binding to FtsZ and the absence and presence of GTP, we collected FtsZ–ZapE complexes by ultrafiltration on a polyethersulfone filter with a 100-kDa exclusion limit. To analyze retained proteins and clearly differentiate ZapE from FtsZ by SDS-PAGE, since they migrate similarly due to their close size, we utilized the chimeric fusion protein Gfp–FtsZ. Gfp–FtsZ, which has a molecular weight of 68 kDa, was previously shown to polymerize, hydrolyze GTP, and behave similarly to wild type FtsZ in *in vitro* assays and assemble into a Z-ring *in vivo* (Viola et al., 2017). In the presence of Gfp–FtsZ (10 μM) and ATP (4 mM), 63.4% of the total ZapE in the reaction (5 μM) was retained in complex with Gfp–FtsZ, whereas only 33.4% of ZapE was retained on the filter when Gfp–FtsZ was omitted, suggesting that ZapE forms a complex with FtsZ in the absence of nucleotide (**Supplementary Figure 2E**). The amount of ZapE that was retained did not change in the presence of GTP, which promotes FtsZ polymerization, suggesting that ZapE binds similarly to non-polymerized and polymerized

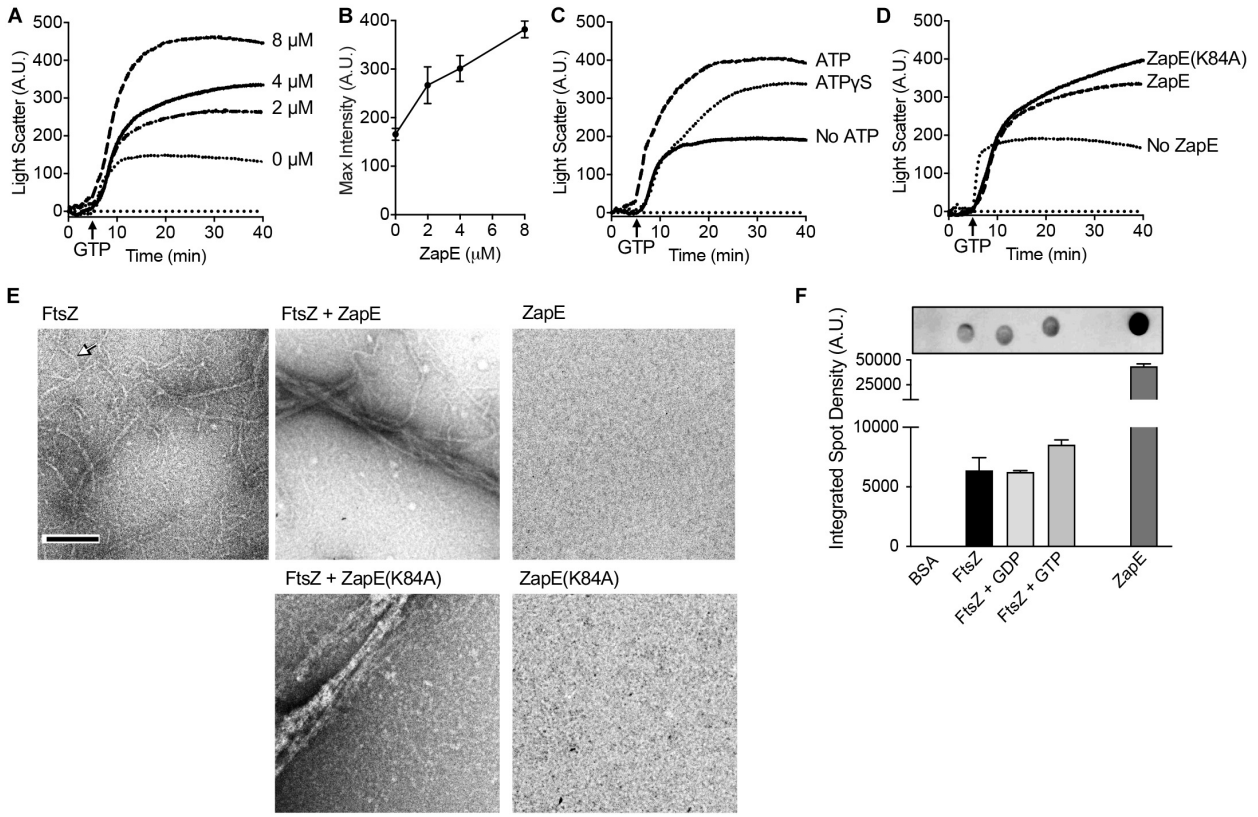


FIGURE 2 | ZapE promotes FtsZ polymer bundling. **(A)** GTP-dependent assembly of complexes containing FtsZ was monitored by 90° light scattering, as described in *Materials and methods* section. FtsZ (8 μM) was incubated with increasing concentration of ZapE (0 to 8 μM) in the presence of ATP (4 mM). After collecting a baseline for 5 min, GTP (1.5 mM) was added to stimulate FtsZ polymerization and light scatter was monitored over time. **(B)** Maximum intensity (A.U.) values after addition of GTP in **(A)** for light scattering reactions containing FtsZ (8.0 μM) and increasing concentrations of ZapE titration (0, 2.0, 4.0, and 8.0 μM) and ATP (4 mM). Data shown are the average of at least three replicates represented as mean ± SEM. **(C)** Light scattering assay showing reactions containing FtsZ (8.0 μM) and ZapE (8.0 μM) in the absence or presence of ATP (4 mM) or ATPγS (2 mM). Data shown are representative of at least three replicates. **(D)** Light scattering assay showing reactions containing FtsZ (8 μM), assembled with GTP (1.5 mM) in the absence and presence of ZapE (4 μM) or ZapE(K84A) (4 μM) with ATP (4 mM). Data shown are representative of at least three replicates. **(E)** ZapE (8 μM) or ZapE(K84A) (8 μM) were incubated with ATP (4 mM) and, where indicated, FtsZ (8 μM) and GTP (1.5 mM). Reactions were incubated for 15 min at 23°C and then visualized by transmission electron microscopy (TEM) as described in *Materials and methods* section. Size bars are 200 nm. Arrow indicates an FtsZ polymer. **(F)** Binding of ZapE to FtsZ was assayed by dot blot. Bovine serum albumin (BSA) (10 μM), ZapE (10 μM), and FtsZ (10 μM) alone or with GDP (2 mM) or GTP (2 mM) were spotted (3 μl) onto a nitrocellulose membrane and then membranes were blocked, incubated with ZapE (8 μM) and ATP (4 mM) to allow binding, and then probed with antibodies to ZapE. Spot densities were quantified by NIH ImageJ. Data shown are the average of at least three replicates represented as mean ± SEM.

Gfp–FtsZ (**Supplementary Figure 2E**). Finally, to determine if ATP hydrolysis is required for complex formation between ZapE and FtsZ, we performed ultrafiltration assays with ZapE(K84A) (5 μM), co-incubated with FtsZ (10 μM) and ATP (4 mM), and also observed that Gfp–FtsZ retained ZapE(K84A) at a higher value (80.2%), in comparison with ZapE(K84A) alone 50.3% (**Supplementary Figure 2F**), suggesting that ZapE nucleotide hydrolysis is not required for Gfp–FtsZ binding.

Phospholipid Binding and Recruitment of FtsZ

In addition to FtsZ, ZapE was reported to be present in complexes with FtsQ, FtsI, FtsL, and FtsN, which are all late cell division proteins and also integral membrane proteins (Marteyn et al., 2014). Therefore, we tested if ZapE could bind to *E. coli* PL directly in a phospholipid recruitment assay (**Figure 3A**). ZapE was incubated with PL vesicles (500 μg ml^{-1}) in the absence and presence of ATP (4 mM), and then supernatant and PL-pellet fractions were collected by low-speed centrifugation and analyzed by SDS-PAGE (**Figure 3A**). We observed that in the presence of ATP, 76% of ZapE fractionated with PLs; however, only 23.6% of ZapE was recruited to the PL pellet in the absence of ATP, suggesting that PL binding is modulated by ATP. In the absence of SUVs, ZapE remained in the supernatant with and without ATP (**Supplementary Figure 2G**). ZapE(K84A) showed a similar fractionation pattern where 80.3% of ZapE(K84A) was recruited to the PL pellet in the presence of ATP (**Figure 3A**). To determine if the interaction was disrupted by increasing ionic strength, we repeated the PL recruitment assay with ZapE and ZapE(K84A) in the presence of ATP and 400 mM NaCl. We observed that ATP-dependent association with PLs was tolerant to 400 mM

NaCl (**Supplementary Figure 2H**), but was disrupted at higher NaCl concentrations (data not shown). This suggests that the binding detected is likely due to an electrostatic interaction at the PL surface (i.e., phosphate head group). Next, we tested if the addition of PLs to ZapE modifies the ATP hydrolysis activity in ATPase activity assays; however, we observed no change in the rate of ATP hydrolysis by ZapE when PLs were included in the reaction (**Figure 3B**). Finally, to determine if ZapE recruits FtsZ to PLs, we incubated Gfp-FtsZ with GTP and PLs in the absence and presence of ZapE and ATP (**Figure 3C**). We observed that 21.6% of the total Gfp-FtsZ included in the reaction fractionated with PLs when ZapE was also present, but only 6% of the total Gfp-FtsZ was detected in the absence of ZapE. Together, our results show that ZapE binds to PLs with ATP, and ZapE recruits FtsZ to the PL surface *in vitro*; however, neither interaction with PLs or FtsZ modifies ZapE ATP hydrolysis activity under the conditions tested.

Deletion of *zapE* Impairs Division and Leads to a Synthetic Slow Growth Phenotype in a *minC* Deletion Strain

ZapE was reported to be an ATPase and also a cell division protein in *E. coli* that is important for division under low-oxygen conditions and in cultures grown at high temperature (42°C) (Marteyn et al., 2014). First, to determine if we observe a phenotypic change in morphology following deletion of *zapE*, we constructed a *zapE* deletion strain in the *E. coli* K-12 strain MG1655 by Lambda Red recombination. Wild-type MG1655 cells and MG1655 Δ*zapE* cells were cultured in liquid LB medium under aerobic conditions, grown to log phase for 3 h at 35°C, and analyzed by microscopy to observe any morphologic changes. Both strains grew to a similar OD_{600} of approximately 0.8 A.U. As expected, *E. coli* MG1655 cells were short rods, with a mean length of 2.78 ± 0.53 μm ($n = 200$); however, cultures of MG1655 Δ*zapE* had a mean cell length of 3.86 ± 6.15 μm ($n = 200$), which is 38.8% longer than wild-type cells (**Figures 4A,B**), with a small population (7%) of filamentous cells longer than 5 μm. These results show that a moderate division phenotype is observed when cells have been deleted for *zapE* and grown aerobically at 35°C (**Figure 4A**). Notably, a mild filamentous phenotype was also previously shown for MG1655 Δ*zapE* cultured under aerobic conditions, and this phenotype was further exacerbated by anaerobiosis and elevated temperature (Marteyn et al., 2014). Next, to determine if ZapE ATP hydrolysis is required to support normal division, we constructed a strain expressing ZapE(K84A) from the native locus on the chromosome by replacing a *parE-kan* cassette at the original locus with *zapE*, to restore the original genotype or *zapE(K84A)*. We performed microscopy on cells in log phase and determined that both strains, *zapE* + (restored) and *zapE(K84A)*, had a cell length distribution similar to wild-type MG1655 cells, and no filamentous cells were observed (**Figures 4C,D**). Additionally, MG1655 Δ*zapE*:*zapE* (restored) cells were observed to have a mean length of 2.26 ± 0.55 μm (n = 200), and MG1655 Δ*zapE*:*zapE(K84A)* cells were observed to have a mean length of 2.31 ± 0.60 μm ($n = 200$) (**Figure 4D**). These results suggest that deletion of *zapE* leads to a mild cell division defect in *E. coli* and that ATP hydrolysis does not appear to be critical for this function, since *zapE(K84A)* restores the moderate filamentation defect, similar to *zapE*.

The cell division defect detected in the deletion strain is moderate, therefore we tested if cultures exposed to high temperature and/or oxygen stress developed a more severe growth defect. First, we cultured wild type MG1655 cells and MG1655 Δ*zapE* cells on solid LB medium without O_2 by dilution spot-plating; however, we observed no major differences in the extent of growth of the two strains after 24 h (**Supplementary Figure 3A**). Then, we tested if exposure to both high temperature and oxygen stress exacerbated any growth defects. We incubated the cells at 42°C for 3 h in static cultures with no air contact to induce thermal and oxygen stress, then spotted culture dilutions onto LB agar plates, and incubated the plates at 23°C (**Figure 5A**). We observed that cultures of cells exposed to stress grew slower than cells cultured aerobically at 35°C after 3 h (**Supplementary Figures 3B,C**); however, this was independent of the absence or presence of *zapE*. We then spot plated the cultures after exposure to stress and observed that both wild type and *zapE* deletion cells grew to a similar extent on plates after the initial exposure to stress (**Figures 5A–C** and **Supplementary Figure 3D**). Moreover, the morphologies of stressed *zapE* deletion cells also appeared similar to morphologies of unstressed cells cultured aerobically at 30°C [2.19 ± 0.15 μm ($n = 100$) and 2.32 ± 0.12 μm ($n = 100$), respectively], with occasional cell filaments observed under both conditions (3% and 2%, respectively) (**Supplementary Figures 3E,F**).

Next, to determine if the deletion of *zapE* in combination with chromosomal deletion of another FtsZ assembly regulator, such as *minC*, further altered cell growth or morphology, we constructed a strain deleted for both *minC* and *zapE* by Lambda Red recombination. It was previously reported that deletion of either *slmA* or *clpX* caused a synthetic phenotype in combination with deletion of *minC* (Bernhardt and de Boer, 2005; Camberg et al., 2011) and that the *minC* deletion strain may be more sensitive to deletion of non-essential FtsZ assembly regulators. MinC is a cell division protein that functions in the cell to prevent assembly of the Z-ring near the cell poles by destabilizing FtsZ polymers (Hu et al., 1999; Hu and Lutkenhaus, 2000; Dajkovic et al., 2008; LaBreck et al., 2019). To evaluate the synthetic phenotypes of *minC* and *zapE* deletion strains, we first compared morphologies, since deletion of *minC* alone induces mild filamentation and minicell formation. We observed that cultures of cells deleted for both *minC* and *zapE* contained short filaments and minicells, similar to the cells deleted for *minC* alone, even after culturing cells at high temperature with oxygen stress (**Supplementary Figures 3E,F**). However, following exposure to stress, we observed that cells deleted for both *minC* and *zapE* recovered poorly, in contrast to the cells deleted for either *minC* or *zapE*. Analysis of cultured spot densities showed severe growth delays that were approximately 50% slower than wild type MG1655 cells (**Figures 5A–C**). Reinsertion of *zapE* or *zapE(K84A)* at the chromosomal *zapE* locus restored the ability to recover after exposure to stress (**Figure 5C**). These results show that after exposure to stress, the cells deleted for *zapE* and *minC* are less able to resume

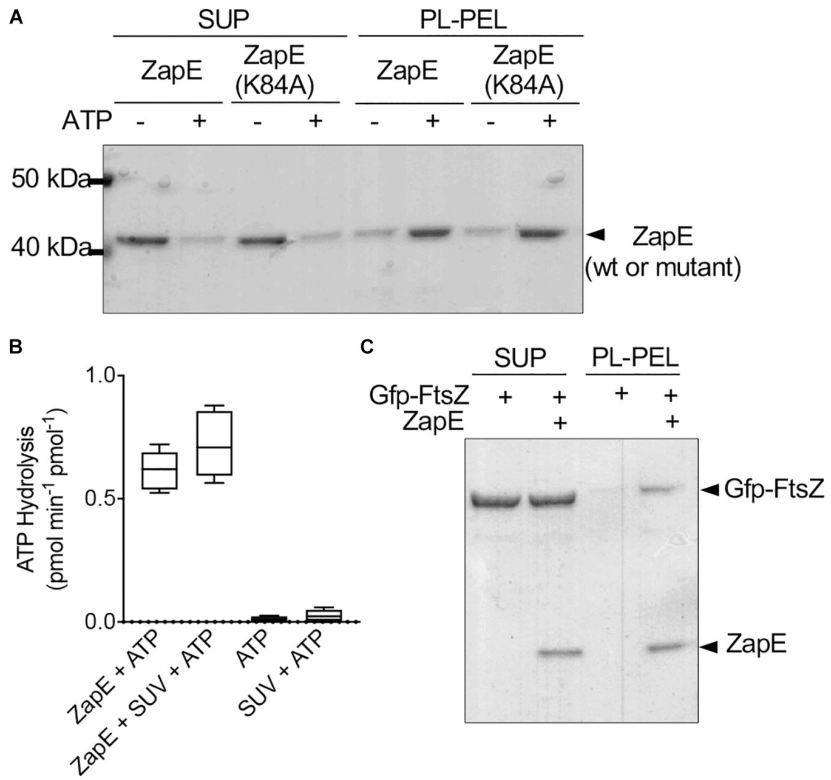


FIGURE 3 | ZapE recruits FtsZ to phospholipid (PL) vesicles *in vitro*. **(A)** ZapE (1 μM) or, where indicated, ZapE(K84A) (1 μM), was incubated with *E. coli* PL vesicles (500 μg ml^{-1}) in the absence or presence of ATP (4 mM). PL-associated protein was collected by low-speed centrifugation. Supernatants and pellets were analyzed by SDS-PAGE. Data shown are representative of at least three replicates. **(B)** ATP hydrolysis assays were performed in the presence of PL vesicles in reactions containing ZapE (6 μM), ATP (4 mM), and PL vesicles (250 μg ml^{-1}). Data shown are the average of at least three replicates represented as mean ± SEM. **(C)** Gfp-FtsZ (3 μM) was incubated with GTP (4 mM) and added to PL vesicles (500 μg ml^{-1}) with and without ZapE (1 μM) and ATP (4 mM). PL-associated protein was collected by low-speed centrifugation. Supernatants and pellets were analyzed by SDS-PAGE. Data shown are representatives of at least three replicates.

normal growth and experience a growth delay on solid LB agar. Finally, we tested if exposure to stress induces expression of *zapE* in wild type cells. We cultured wild type, *minC* and *zapE* deletion strains under stressed (42°C, oxygen stress) and non-stressed conditions (37°C, aerobic with shaking). We observed that ZapE was present in wild-type cells under both stressed and non-stressed conditions and not present in Δ*zapE* or Δ*zapE*Δ*minC* deletion strains (**Figure 5D**). Interestingly, ZapE appears to be present at higher levels in the cells deleted for *minC* in both stressed and non-stressed cells (**Figure 5D**); however, it is unclear if this is due to transcriptional regulation or modified post-translational degradation. Together, these results show that although ZapE is important for resuming division following stress, it does not appear that it is induced or synthesized to a large extent in response to exposure to environmental stress.

DISCUSSION

Here, we report that ZapE is an ATPase that binds directly to FtsZ and bundles GTP-induced FtsZ polymers in light scattering assays and by TEM, and this interaction requires ATP. We also show that ZapE binds to PL vesicles *in vitro* with ATP and can recruit the chimeric protein Gfp-FtsZ to PL vesicles. While FtsZ polymer bundling activity was observed *in vitro* in the absence of PLs, it is unclear if the bundling activity also occurs while ZapE is bound to PLs. To date, only two other proteins in *E. coli* have been reported to localize FtsZ to the PL membrane, FtsA, and ZipA (Hale and de Boer, 1997, 1999; Liu et al., 1999; Hale et al., 2000; Haney et al., 2001; Pichoff and Lutkenhaus, 2002; Conti et al., 2018). Although ZapE is not essential for survival, it is unclear if ZapE shares any redundant functions or overlapping roles with either FtsA or ZipA (Pichoff and Lutkenhaus, 2002; Geissler et al., 2003). The interaction between ZapE and PLs is likely electrostatic, since it is disrupted by high salt concentration; however, the interaction is tolerant to 400 mM NaCl, so it is likely relevant *in vivo* (**Supplementary Figure 2H**). Accordingly, ZapE was previously reported in association with the bacterial inner membrane by EM (Marteyn et al., 2014). As ATP is required for ZapE to bundle FtsZ polymers in light scatter assays and to bind to PLs ***in vitro*** (**Figures 2C**, **3A**), this suggests that a potential role for ZapE may be to stabilize FtsZ polymer assemblies adjacent to the membrane. Furthermore, this function appears to be important for ensuring growth during and after exposure to stress.

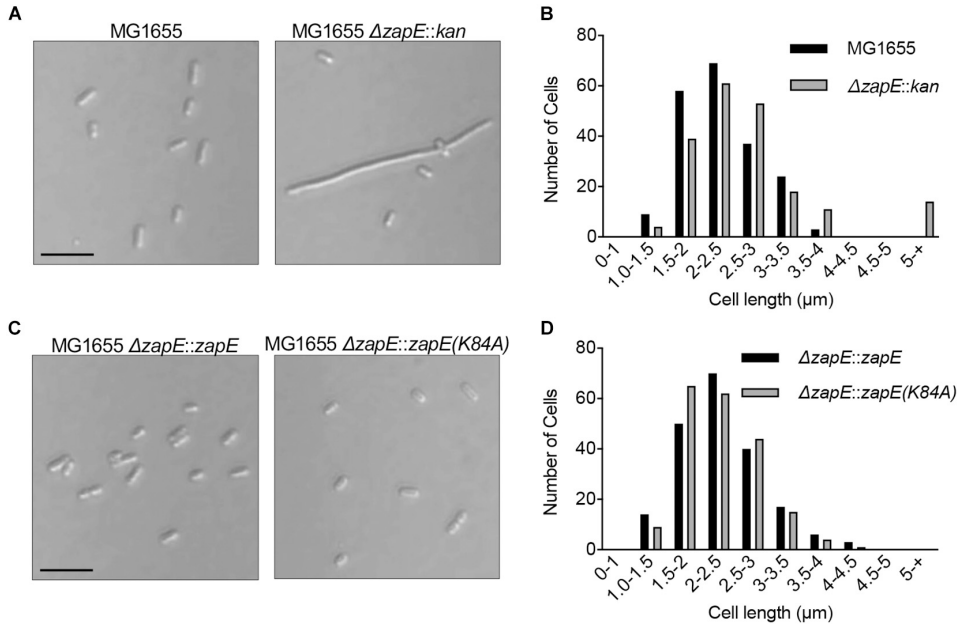


FIGURE 4 | Cell morphology of *E. coli* strains deleted for *zapE.* **(A)** Log phase cultures of MG1655, MG1655 Δ*zapE* grown at 35°C in LB. Cells were collected and added to poly-$_L$-lysine-coated coverslips and observed by DIC microscopy. Size bar is 5 μm. **(B)** Cell length distribution of MG1655 and MG1655 Δ*zapE* from morphology experiment in **(A)** (*n* = 200 cells). **(C)** Expression of ZapE wild type or mutant protein was restored in MG1655 Δ*zapE* with wild type *zapE* or *zapE(K84A)* at the native locus in the chromosome by recombination. Log phase cultures of MG1655 Δ*zapE*:*zapE* and MG1655 Δ*zapE*:*zapE(K84A)* were grown at 35°C in LB. Cells were collected and added to poly-$_L$-lysine-coated coverslips and visualized by DIC microscopy. Size bar is 5 μm. **(D)** Cell length distribution of MG1655 Δ*zapE*:*zapE* and MG1655 Δ*zapE*:*zapE(K84A)* from morphology experiment in **(C)** (*n* = 200 cells).

ZapE is member of the AAA + superfamily of ATPases and, in our study, we show that ZapE is a monomer with and without ATP (**Supplementary Figures 1B,C**). ZapE shares sequence homology with other AAA + ATPases, including DnaA, DnaC, FtsH, and VCP/p97 (Peters et al., 1992; Akiyama et al., 1995; Tomoyasu et al., 1995; Karata et al., 2001; Davies et al., 2008; Duderstadt and Berger, 2008; Mogk et al., 2008). AAA + ATPases have two fundamental regions important for ATP binding and subsequent hydrolysis, Walker A and Walker B motifs. The Walker A Lys (K84) residue is crucial for ATP hydrolysis of ZapE *in vitro* since we observed that ZapE(K84A) is defective for ATP hydrolysis; however, ZapE(K84A) binds to and bundles FtsZ polymers. Many AAA + ATPases (i.e., ClpX) form hexameric rings, and may be stabilized when bound to nucleotide (Grimaud et al., 1998) (i.e., FtsH and VCP/p97) (Akiyama et al., 1995; Tomoyasu et al., 1995; Karata et al., 2001; Davies et al., 2008; Duderstadt and Berger, 2008; Mogk et al., 2008). Although less common, there are examples of stable monomeric AAA + proteins, including the initiator of DNA replication, DnaA, and the eukaryotic cell division, and DNA replication regulator Cdc6 (cell division control protein 6) (Liu et al., 2000; Rozgaja et al., 2011). The mechanism of FtsZ polymer bundling by ZapE is still unclear; however, if ZapE functions as a monomer during the bundling activity, then it may contain two FtsZ-interaction sites per ZapE protomer that could bridge two protofilaments. Other Zaps (A–D) also interact with FtsZ and influence bundling of FtsZ protofilaments, although the Zaps are not structurally related. ZapA interacts directly with FtsZ, while ZapB does not, but is recruited to midcell by ZapA and enhances the activity of ZapA to promote FtsZ filament bundling (Low et al., 2004; Small et al., 2007; Galli and Gerdes, 2012; Buss et al., 2013; Roach et al., 2014; Caldas et al., 2019). *In vitro*, ZapA forms a tetramer and cross-links FtsZ polymers, while reducing the GTPase activity of FtsZ (Gueiros-Filho and Losick, 2002; Low et al., 2004; Small et al., 2007; Pacheco-Gomez et al., 2013). ZapC also promotes the bundling of FtsZ polymers and inhibits FtsZ GTP hydrolysis (Hale et al., 2011). Deletion of *zapC* leads to a minor cell elongation phenotype, which is exacerbated by additional deletions in *zapA* or *zapB* and further suggests a role for ZapC in Z-ring stability (Hale et al., 2011; Durand-Heredia et al., 2012). *In vitro*, ZapD cross-links FtsZ polymers and reduces the GTPase activity of FtsZ (Durand-Heredia et al., 2012). Deletion of *zapD* alone confers no defect in viability; however, in combination with temperature-sensitive mutation in the *ftsZ* gene, *ftsZ*(G105S), deletion of *zapD* leads to filamentation and a decrease in viability, suggesting the role of ZapD as a regulator of FtsZ (Durand-Heredia et al., 2012). While ZapA, ZapC, and ZapD all bundle FtsZ, leading toward a reduction in the rate of FtsZ in GTP hydrolysis, we detected no decrease in FtsZ GTP hydrolysis activity in the presence of ZapE under the conditions tested, which may suggest a novel mechanism for FtsZ polymer bundling or crosslinking. Moreover, the previous characterization of ZapE reported that ZapE disassembles FtsZ polymers in reactions containing Ca^{2+}, which is known to support the formation of bundled sheets of FtsZ and GTP

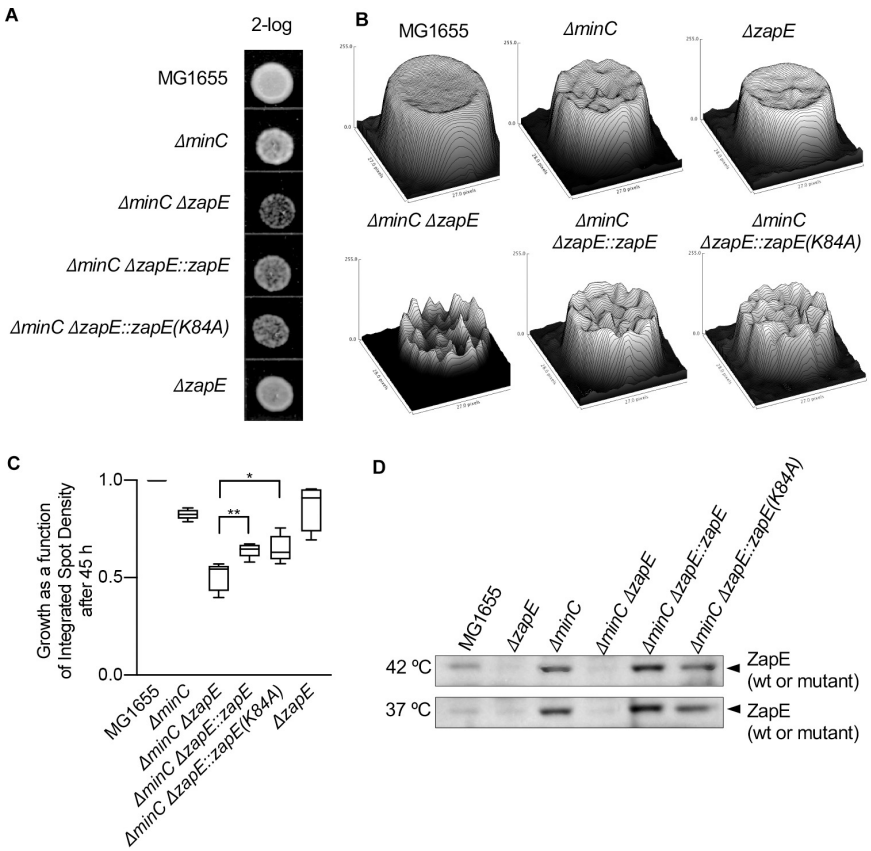


FIGURE 5 | Cells deleted for *zapE* and *minC* grow slowly after exposure to thermal and oxygen stress. **(A)** Overnight cultures MG1655, MG1655 Δ*zapE,* MG1655 Δ*minC,* MG1655 Δ*minC*Δ*zapE*:*kan,* MG1655 Δ*minC*Δ*zapE*:*zapE(K84A),* and MG1655 Δ*minC*Δ*zapE*:*zapE* were diluted in to fresh medium (LB broth) and exposed to thermal stress (42°C) and low oxygen conditions for 3 h. Then, dilutions (2-log) were spotted (5 μl) on no salt LB agar and incubated at room temperature. **(B)** Images of spots were collected and analyzed by surface mapping in ImageJ to compare biomass and visualized in a three-dimensional plot of density versus pixel area. Biomass was evaluated by comparing integrated spot density after 45 h of growth. **(C)** Growth analysis in **(B)** was collected across five replicate experiments. Average biomass as a function of integrated spot density is represented as a box and whiskers plot. (p-values are as follows: *<0.05 and **<0.01). **(D)** Overnight cultures MG1655, MG1655 Δ*zapE,* MG1655 Δ*minC,* MG1655 Δ*minC*Δ*zapE*:*kan,* MG1655 Δ*minC*Δ*zapE*:*zapE(K84A),* and MG1655 Δ*minC*Δ*zapE*:*zapE* were diluted into fresh medium (LB broth) and exposed to thermal stress (42°C) and low oxygen conditions, or at 37°C and shaking, for 3 h. Cells were harvested and analyzed for ZapE expression by immunoblotting.

(Yu and Margolin, 1997; Lowe and Amos, 1999; Marteyn et al., 2014). We did not detect FtsZ protofilament disassembly activity by sedimentation, light scatter, or TEM under the conditions tested. However, we did detect that ZapE binds to unassembled FtsZ and suspect that under some conditions, ZapE could sequester FtsZ monomers and prevent them from participating in polymerization, or that ZapE modulates Ca^{2+}-dependent sheet formation of FtsZ in the presence of GTP. Our work shows that the FtsZ polymer bundling function of ZapE is calcium independent (**Figures 2A,E**). Additional work will be necessary to determine if ZapE performs a holdase or destabilizing function that modifies FtsZ polymer dynamics *in vivo*. A previous report demonstrated that the exchange of Gfp–FtsZ polymer subunits within the Z-ring in dividing cells is slower in cells deleted for *zapE,* indicating that ZapE influences FtsZ structures at the division site and supporting the model that ZapE contributes to the regulation of FtsZ polymer assemblies during division (Viola et al., 2017).

In *E. coli*, deletion of *zapE* was previously reported to produce an elongated cell phenotype exacerbated by anaerobic conditions or elevated temperature (42°C) (Marteyn et al., 2014). Our study also demonstrated conditions in which deletion of *zapE* generated a subpopulation of filamentous cells under the conditions tested. We also observed a synthetic slow growth phenotype in cells deleted for *minC* and *zapE* after exposure to mild heat shock, which is more severe than deletion of either gene individually. Synthetic phenotypes have also been observed in other strains when *minC* was deleted in combination with *slmA* (synthetic lethal) or *clpX* (synthetic filamentous) (Bernhardt and de Boer, 2005; Camberg et al., 2011). MinC spatially regulates placement of the Z-ring *in vivo* by preventing Z-ring assembly near the cell poles (Hu et al., 1999; Hu and Lutkenhaus, 2000; Dajkovic et al., 2008; LaBreck et al., 2019). In cells deleted for *minC*, ZapE may function to organize FtsZ near the membrane to support future rounds of proliferation. Our results establish

that when division site selection is impaired by deletion of *minC*, or when cells are under stress, other FtsZ assembly regulators including ZapE are important for promoting division.

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

ED, RD, CT, and JC contributed to conception and design of the study. ED, RD, CT, and CF contributed to assay development and performed experiments. JM performed electron microscopy. ED wrote the first draft of the manuscript. CT, CF, JM, and JC wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

Research reported in this publication was supported in part by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM118927 to JC. The funders had no role in the study design, data collection, and interpretation, or the decision to submit the work for publication.

ACKNOWLEDGMENTS

We thank Ben Piraino, Negar Rahmani, and Victor Hunt for their helpful discussions and critical reading of the manuscript, and Janet Atoyan for microscopy and sequencing assistance. Microscopy and sequencing were performed at the Rhode Island Genomics and Sequencing Center, supported in part by the National Science Foundation (MRI Grant No. DBI-0215393 and EPSCoR Grant Nos. 0554548 and EPS-1004057), the US Department of Agriculture (Grant Nos. 2002-34438-12688, 2003-34438-13111, and 2008-34438-19246), and the University of Rhode Island. The TEM data was acquired at the RI Consortium for Nanoscience and Nanotechnology, a URI College of Engineering core facility partially funded by the National Science Foundation EPSCoR, Cooperative Agreement #OIA-1655221.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.733085/full#supplementary-material

REFERENCES

Akiyama, Y., Yoshihisa, T., and Ito, K. (1995). FtsH, a membrane-bound ATPase, forms a complex in the cytoplasmic membrane of *Escherichia coli*. *J. Biol. Chem.* 270, 23485–23490. doi: 10.1074/jbc.270.40.23485

Balasubramanian, A., Markovski, M., Hoskins, J. R., Doyle, S. M., and Wickner, S. (2019). Hsp90 of *E. coli* modulates assembly of FtsZ, the bacterial tubulin homolog. *Proc. Natl. Acad. Sci. U.S.A.* 116, 12285–12294. doi: 10.1073/pnas.1904014116

Bernhardt, T. G., and de Boer, P. A. (2005). SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Mol. Cell* 18, 555–564. doi: 10.1016/j.molcel.2005.04.012

Bi, E. F., and Lutkenhaus, J. (1991). FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354, 161–164. doi: 10.1038/354161a0

Blattner, F. R., Plunkett, G. III, 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

Buss, J., Coltharp, C., Huang, T., Pohlmeyer, C., Wang, S. C., Hatem, C., et al. (2013). In vivo organization of the FtsZ-ring by ZapA and ZapB revealed by quantitative super-resolution microscopy. *Mol. Microbiol.* 89, 1099–1120. doi: 10.1111/mmi.12331

Caldas, P., Lopez-Pelegrin, M., Pearce, D. J. G., Budanur, N. B., Brugues, J., and Loose, M. (2019). Cooperative ordering of treadmilling filaments in cytoskeletal networks of FtsZ and its crosslinker ZapA. *Nat. Commun.* 10:5744. doi: 10.1038/s41467-019-13702-4

Camberg, J. L., Hoskins, J. R., and Wickner, S. (2009). ClpXP protease degrades the cytoskeletal protein, FtsZ, and modulates FtsZ polymer dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10614–10619. doi: 10.1073/pnas.0904886106

Camberg, J. L., Hoskins, J. R., and Wickner, S. (2011). The interplay of ClpXP with the cell division machinery in *Escherichia coli*. *J. Bacteriol.* 193, 1911–1918. doi: 10.1128/JB.01317-10

Camberg, J. L., Viola, M. G., Rea, L., Hoskins, J. R., and Wickner, S. (2014). Location of dual sites in *E. coli* FtsZ important for degradation by ClpXP; one at the C-terminus and one in the disordered linker. *PLoS One* 9:e94964. doi: 10.1371/journal.pone.0094964

Conti, J., Viola, M. G., and Camberg, J. L. (2018). FtsA reshapes membrane architecture and remodels the Z-ring in *Escherichia coli*. *Mol. Microbiol.* 107, 558–576. doi: 10.1111/mmi.13902

Dajkovic, A., Lan, G., Sun, S. X., Wirtz, D., and Lutkenhaus, J. (2008). MinC spatially controls bacterial cytokinesis by antagonizing the scaffolding function of FtsZ. *Curr. Biol.* 18, 235–244. doi: 10.1016/j.cub.2008.01.042

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

Davies, J. M., Brunger, A. T., and Weis, W. I. (2008). Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure* 16, 715–726. doi: 10.1016/j.str.2008.02.010

de Boer, P. A. (2010). Advances in understanding *E. coli* cell fission. *Curr. Opin. Microbiol.* 13, 730–737. doi: 10.1016/j.mib.2010.09.015

Du, S., and Lutkenhaus, J. (2019). At the heart of bacterial cytokinesis: the Z ring. *Trends Microbiol.* 27, 781–791. doi: 10.1016/j.tim.2019.04.011

Duderstadt, K. E., and Berger, J. M. (2008). AAA+ ATPases in the initiation of DNA replication. *Crit. Rev. Biochem. Mol. Biol.* 43, 163–187. doi: 10.1080/10409230802058296

Durand-Heredia, J., Rivkin, E., Fan, G., Morales, J., and Janakiraman, A. (2012). Identification of ZapD as a cell division factor that promotes the assembly of FtsZ in *Escherichia coli*. *J. Bacteriol.* 194, 3189–3198. doi: 10.1128/JB.00176-12

Galli, E., and Gerdes, K. (2012). FtsZ-ZapA-ZapB interactome of *Escherichia coli*. *J. Bacteriol.* 194, 292–302. doi: 10.1128/JB.05821-11

Geissler, B., Elraheb, D., and Margolin, W. (2003). A gain-of-function mutation in *ftsA* bypasses the requirement for the essential cell division gene *zipA* in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4197–4202. doi: 10.1073/pnas.0635003100

Grimaud, R., Kessel, M., Beuron, F., Steven, A. C., and Maurizi, M. R. (1998). Enzymatic and structural similarities between the *Escherichia coli* ATP-

dependent proteases, ClpXP and ClpAP. *J. Biol. Chem.* 273, 12476–12481. doi: 10.1074/jbc.273.20.12476

Gueiros-Filho, F. J., and Losick, R. (2002). A widely conserved bacterial cell division protein that promotes assembly of the tubulin-like protein FtsZ. *Genes Dev.* 16, 2544–2556. doi: 10.1101/gad.1014102

Hale, C. A., and de Boer, P. A. (1997). Direct binding of FtsZ to ZipA, an essential component of the septal ring structure that mediates cell division in *E. coli*. *Cell* 88, 175–185. doi: 10.1016/S0092-8674(00)81838-3

Hale, C. A., and de Boer, P. A. (1999). Recruitment of ZipA to the septal ring of *Escherichia coli* is dependent on FtsZ and independent of FtsA. *J. Bacteriol.* 181, 167–176. doi: 10.1128/JB.181.1.167-176.1999

Hale, C. A., Rhee, A. C., and de Boer, P. A. (2000). ZipA-induced bundling of FtsZ polymers mediated by an interaction between C-terminal domains. *J. Bacteriol.* 182, 5153–5166. doi: 10.1128/JB.182.18.5153-5166.2000

Hale, C. A., Shiomi, D., Liu, B., Bernhardt, T. G., Margolin, W., Niki, H., et al. (2011). Identification of *Escherichia coli* ZapC (YcbW) as a component of the division apparatus that binds and bundles FtsZ polymers. *J. Bacteriol.* 193, 1393–1404. doi: 10.1128/JB.01245-10

Haney, S. A., Glasfeld, E., Hale, C., Keeney, D., He, Z., and de Boer, P. (2001). Genetic analysis of the *Escherichia coli* FtsZ.ZipA interaction in the yeast two-hybrid system. Characterization of FtsZ residues essential for the interactions with ZipA and with FtsA. *J. Biol. Chem.* 276, 11980–11987. doi: 10.1074/jbc.M009810200

Hu, Z., and Lutkenhaus, J. (2000). Analysis of MinC reveals two independent domains involved in interaction with MinD and FtsZ. *J. Bacteriol.* 182, 3965–3971. doi: 10.1128/JB.182.14.3965-3971.2000

Hu, Z., Mukherjee, A., Pichoff, S., and Lutkenhaus, J. (1999). The MinC component of the division site selection system in *Escherichia coli* interacts with FtsZ to prevent polymerization. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14819–14824. doi: 10.1073/pnas.96.26.14819

Karata, K., Verma, C. S., Wilkinson, A. J., and Ogura, T. (2001). Probing the mechanism of ATP hydrolysis and substrate translocation in the AAA protease FtsH by modelling and mutagenesis. *Mol. Microbiol.* 39, 890–903. doi: 10.1046/j.1365-2958.2001.02301.x

LaBreck, C. J., Conti, J., Viola, M. G., and Camberg, J. L. (2019). MinC N- and C-domain interactions modulate FtsZ assembly, division site selection, and MinD-dependent oscillation in *Escherichia coli*. *J. Bacteriol.* 201:e00374-18. doi: 10.1128/JB.00374-18

LaBreck, C. J., May, S., Viola, M. G., Conti, J., and Camberg, J. L. (2017). The protein chaperone ClpX targets native and non-native aggregated substrates for remodeling, disassembly, and degradation with ClpP. *Front. Mol. Biosci.* 4:26. doi: 10.3389/fmolb.2017.00026

Liu, J., Smith, C. L., DeRyckere, D., DeAngelis, K., Martin, G. S., and Berger, J. M. (2000). Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol. Cell* 6, 637–648. doi: 10.1016/S1097-2765(00)00062-9

Liu, Z., Mukherjee, A., and Lutkenhaus, J. (1999). Recruitment of ZipA to the division site by interaction with FtsZ. *Mol. Microbiol.* 31, 1853–1861. doi: 10.1046/j.1365-2958.1999.01322.x

Low, H. H., Moncrieffe, M. C., and Lowe, J. (2004). The crystal structure of ZapA and its modulation of FtsZ polymerisation. *J. Mol. Biol.* 341, 839–852. doi: 10.1016/j.jmb.2004.05.031

Lowe, J., and Amos, L. A. (1999). Tubulin-like protofilaments in Ca2+-induced FtsZ sheets. *EMBO J.* 18, 2364–2371. doi: 10.1093/emboj/18.9.2364

Marteyn, B. S., Karimova, G., Fenton, A. K., Gazi, A. D., West, N., Touqui, L., et al. (2014). ZapE is a novel cell division protein interacting with FtsZ and modulating the Z-ring dynamics. *mBio* 5:e00022-14. doi: 10.1128/mBio.00022-14

Mogk, A., Haslberger, T., Tessarz, P., and Bukau, B. (2008). Common and specific mechanisms of AAA+ proteins involved in protein quality control. *Biochem. Soc. Trans.* 36, 120–125. doi: 10.1042/BST0360120

Mukherjee, A., and Lutkenhaus, J. (1994). Guanine nucleotide-dependent assembly of FtsZ into filaments. *J. Bacteriol.* 176, 2754–2758. doi: 10.1128/jb.176.9.2754-2758.1994

Mukherjee, A., and Lutkenhaus, J. (1999). Analysis of FtsZ assembly by light scattering and determination of the role of divalent metal cations. *J. Bacteriol.* 181, 823–832. doi: 10.1128/JB.181.3.823-832.1999

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

Pacheco-Gomez, R., Cheng, X., Hicks, M. R., Smith, C. J., Roper, D. I., Addinall, S., et al. (2013). Tetramerization of ZapA is required for FtsZ bundling. *Biochem. J.* 449, 795–802. doi: 10.1042/BJ20120140

Peters, J. M., Harris, J. R., Lustig, A., Muller, S., Engel, A., Volker, S., et al. (1992). Ubiquitous soluble Mg(2+)-ATPase complex. A structural study. *J. Mol. Biol.* 223, 557–571. doi: 10.1016/0022-2836(92)90670-F

Pichoff, S., and Lutkenhaus, J. (2002). Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. *EMBO J.* 21, 685–693. doi: 10.1093/emboj/21.4.685

Roach, E. J., Kimber, M. S., and Khursigara, C. M. (2014). Crystal structure and site-directed mutational analysis reveals key residues involved in *Escherichia coli* ZapA function. *J. Biol. Chem.* 289, 23276–23286. doi: 10.1074/jbc.M114.561928

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., and Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. *Mol. Microbiol.* 82, 475–488. doi: 10.1111/j.1365-2958.2011.07827.x

Schumacher, M. A., Huang, K. H., Zeng, W., and Janakiraman, A. (2017). Structure of the Z ring-associated protein, ZapD, bound to the C-terminal domain of the tubulin-like protein, FtsZ, suggests mechanism of Z ring stabilization through FtsZ cross-linking. *J. Biol. Chem.* 292, 3740–3750. doi: 10.1074/jbc.M116.773192

Small, E., Marrington, R., Rodger, A., Scott, D. J., Sloan, K., Roper, D., et al. (2007). FtsZ polymer-bundling by the *Escherichia coli* ZapA orthologue, YgfE, involves a conformational change in bound GTP. *J. Mol. Biol.* 369, 210–221. doi: 10.1016/j.jmb.2007.03.025

Soderstrom, B., Badrutdinov, A., Chan, H., and Skoglund, U. (2018). Cell shape-independent FtsZ dynamics in synthetically remodeled bacterial cells. *Nat. Commun.* 9:4323. doi: 10.1038/s41467-018-06887-7

Soderstrom, B., Chan, H., and Daley, D. O. (2019). Super-resolution images of peptidoglycan remodelling enzymes at the division site of *Escherichia coli*. *Curr. Genet.* 65, 99–101. doi: 10.1007/s00294-018-0869-x

Soderstrom, B., and Daley, D. O. (2017). The bacterial divisome: more than a ring? *Curr. Genet.* 63, 161–164. doi: 10.1007/s00294-016-0630-2

Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., et al. (1995). *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. *EMBO J.* 14, 2551–2560. doi: 10.1002/j.1460-2075.1995.tb07253.x

Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P., and Bukau, B. (2001). Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. *Mol. Microbiol.* 40, 397–413. doi: 10.1046/j.1365-2958.2001.02383.x

Viola, M. G., LaBreck, C. J., Conti, J., and Camberg, J. L. (2017). Proteolysis-dependent remodeling of the tubulin homolog FtsZ at the division septum in *Escherichia coli*. *PLoS One* 12:e0170505. doi: 10.1371/journal.pone.0170505

Yu, X. C., and Margolin, W. (1997). Ca2+-mediated GTP-dependent dynamic assembly of bacterial cell division protein FtsZ into asters and polymer networks in vitro. *EMBO J.* 16, 5455–5463. doi: 10.1093/emboj/16.17.5455

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 26 October 2021
doi: 10.3389/fmicb.2021.721899

Quantitative Examination of Five Stochastic Cell-Cycle and Cell-Size Control Models for *Escherichia coli* and *Bacillus subtilis*

Guillaume Le Treut[1], Fangwei Si[2], Dongyang Li[3] and Suckjoon Jun[2,4]**

[1] Theory Group, Chan Zuckerberg Biohub, San Francisco, CA, United States, [2] Department of Physics, University of California, San Diego, San Diego, CA, United States, [3] Division of Biology and Biological Engineering, Broad Center, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA, United States, [4] Section of Molecular Biology, Division of Biology, University of California, San Diego, San Diego, CA, United States

OPEN ACCESS

Edited by:
Monika Glinkowska, University of Gdańsk, Poland

Reviewed by:
Shiladitya Banerjee, Carnegie Mellon University, United States
Ramon Grima, University of Edinburgh, United Kingdom

***Correspondence:**
Guillaume Le Treut
guillaume.letreut@czbiohub.org
Suckjoon Jun
suckjoon.jun@gmail.com

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *07 June 2021*
Accepted: *06 October 2021*
Published: *26 October 2021*

Citation:
Le Treut G, Si F, Li D and Jun S (2021) Quantitative Examination of Five Stochastic Cell-Cycle and Cell-Size Control Models for Escherichia coli and Bacillus subtilis. Front. Microbiol. 12:721899. doi: 10.3389/fmicb.2021.721899

We examine five quantitative models of the cell-cycle and cell-size control in *Escherichia coli* and *Bacillus subtilis* that have been proposed over the last decade to explain single-cell experimental data generated with high-throughput methods. After presenting the statistical properties of these models, we test their predictions against experimental data. Based on simple calculations of the defining correlations in each model, we first dismiss the stochastic Helmstetter-Cooper model and the Initiation Adder model, and show that both the Replication Double Adder (RDA) and the Independent Double Adder (IDA) model are more consistent with the data than the other models. We then apply a recently proposed statistical analysis method and obtain that the IDA model is the most likely model of the cell cycle. By showing that the RDA model is fundamentally inconsistent with size convergence by the adder principle, we conclude that the IDA model is most consistent with the data and the biology of bacterial cell-cycle and cell-size control. Mechanistically, the Independent Adder Model is equivalent to two biological principles: (i) balanced biosynthesis of the cell-cycle proteins, and (ii) their accumulation to a respective threshold number to trigger initiation and division.

Keywords: adder, bacterial cell cycle, bacterial cell size control, quantitative microbial physiology, bacterial physiology

INTRODUCTION

Quantitative microbial physiology is marked by close interactions between experiment and modeling since its birth in the mid twentieth century (see Jun et al., 2018 for a review of the history with extensive literature). In particular, bacterial cell-size and cell-cycle control has enjoyed rejuvenated interests in modeling with the advent of microfluidics techniques that allow tracking of thousands of individual cells over a hundred division cycles (see, for example, Wang et al., 2010; Moffitt et al., 2012; Long et al., 2013; Vashistha et al., 2021). Re-emerged from the new single-cell data is the adder principle (Campos et al., 2014; Jun and Taheri-Araghi, 2015; Taheri-Araghi et al., 2015), which states that individual cells grow by adding a fixed size from birth to division,

independently from their size at birth. This principle has characteristic repercussions on cell size homeostasis. Specifically, upon perturbation, the cell size at birth relaxes toward its steady-state value according to a first-order recurrence relation with a correlation coefficient equal to 1/2 (Voorn et al., 1993; Amir, 2014; Taheri-Araghi et al., 2015).

Although the adder principle was originally proposed and statistically tested almost three decades ago by Voorn et al. (1993) before its recent revival, its mechanistic origin has remained elusive until recently because direct experimental tests were not available for a long time (Jun et al., 2018; Si et al., 2019). Several models have been proposed so far (Campos et al., 2014; Iyer-Biswas et al., 2014; Taheri-Araghi et al., 2015; Harris and Theriot, 2016; Wallden et al., 2016; Amir, 2017; Micali et al., 2018; Si et al., 2019; Witz et al., 2019; Bertaux et al., 2020; Serbanescu et al., 2020; Zheng et al., 2020), and we expect a consensus to emerge as more experimental data become available.

The main purpose of this article is to derive and present steady-state statistical properties of quantitative bacterial cell-cycle and cell-size control models that we are currently aware of and, where relevant, critically examine them against single-cell data from our lab's mother machine experiments accumulated over the last decade in *E. coli* and *B. subtilis*. These models are (i) the stochastic Helmstetter-Cooper model (sHC) (Si et al., 2019), (ii) the initiation adder (IA) model (Amir, 2014, 2017; Ho and Amir, 2015), (iii) the Replication Double Adder (RDA) model (Witz et al., 2019), (iv) the Independent Double Adder (IDA) model (Si et al., 2019), and (v) the concurrent cell-cycle processes (CCCP) model (Boye and Nordström, 2003) and its stochastic version (Micali et al., 2018).

Some of these models are graphically illustrated in **Figure 1**. Briefly, the sHC model is a literal extension of the textbook Helmstetter Cooper model by allowing independent Gaussian fluctuations to each of the initiation mass, the τ_{cyc} = C+D period (from initiation to division), and the cell elongation rate. The IA model assumes that replication initiation is the sole implementation point of cell-size control, and division is strictly coupled to initiation such that division is triggered after fixed τ_{cyc} = C+D minutes have elapsed since initiation. The RDA model is similar to the IA model in that it also assumes that initiation is the reference point for cell-size control. Its main difference from the IA model is that it assumes division is triggered after the cell elongates a constant length per origin of replication δ_{id}, rather than a constant time, since initiation. In other words, δ_{id} is the added size during the C+D period. Both the IA and the RDA models assume the initiation adder, i.e., the cell growth by a nearly fixed size per replication origin between two consecutive initiation cycles, irrespective of the cell size at initiation (initiation mass).

The IDA model states that initiation and division are independently controlled by their respective initiator proteins. However, the IDA model is based on mechanistic assumptions that these proteins are produced in a balanced manner (i.e., for every protein, the mass synthesis rate is a fixed fraction of the total mass synthesis rate Scott et al., 2010), and initiation and division are triggered when the cell has accumulated their respective initiator proteins to their respective threshold numbers. The CCCP model states that replication cycle and division cycles progress independently, but checkpoints or their equivalent are activated to ensure cell division (Boye and Nordström, 2003).

This article is structured as follows. In section "Statistical Properties of Five Bacterial Cell-Size and Cell-Cycle Control Models," we summarize the five models and derive some of their statistical properties. In section "Test of the Models Against Data," we test the predictions of these models against the data. In section "Discussion," we critically examine one of the recent correlation analysis methods (the *I*-value analysis) used to justify the RDA model. We conclude that the IDA model is as of today the model most consistent with data, which also provides a falsifiable mechanistic picture.

STATISTICAL PROPERTIES OF FIVE BACTERIAL CELL-SIZE AND CELL-CYCLE CONTROL MODELS

The Stochastic Helmstetter-Cooper Model

The original HC model (Cooper and Helmstetter, 1968) is based on the experimental observation that the average duration of chromosome replication ("*C* period") can be longer than the average doubling time of the cells in fast-growing *E. coli*. In such growth conditions, *E. coli* must initiate a new round of replication before the ongoing replication cycle is completed. The core of the HC model is the recipe to trace replication initiation backward by $\tau_{cyc} = C+D > \tau$ minutes starting from cell division during overlapping cell cycles (**Figure 1**).

Thus, the HC model introduces three control parameters for a complete description of replication and division cycles: two temporal parameters (the doubling time τ and the duration of cell cycle $\tau_{cyc} = C+D$) and one spatial parameter (e.g., cell size at division or initiation). It was Donachie who showed that, if (i) $\tau_{cyc} = C+D$ is invariant under different nutrient conditions and (ii) the average cell size increases exponentially with respect to the nutrient-imposed growth rate $\lambda = \ln 2/\tau$ as $S = s_i \exp(\alpha \lambda)$ (where s_i and α are constant, and S is the average cell size of a steady-state population), then the cell size at initiation per replication origin s_i (or, the "initiation mass") must be mathematically invariant in all growth conditions (Donachie, 1968). This result was later generalized to all steady-state growth conditions with and without growth inhibition (Si et al., 2017).

Since the original HC model is deterministic and can be defined in terms of λ, τ_{cyc} and s_i (or division size S_d), one possible extension to a stochastic version is by making the three physiological variables stochastic. Together, they completely determine cell sizes including the size at division, assuming perfectly symmetric division (**Figure 1**). For simplicity, we draw λ, τ_{cyc} and s_i at cell birth from a multivariate Gaussian distribution, which also encodes cross- and mother-daughter correlations in the covariances matrix (Si et al., 2019).

The recursion relation for the cell size at division in this "stochastic" Helmstetter-Cooper (sHC) model can be written as

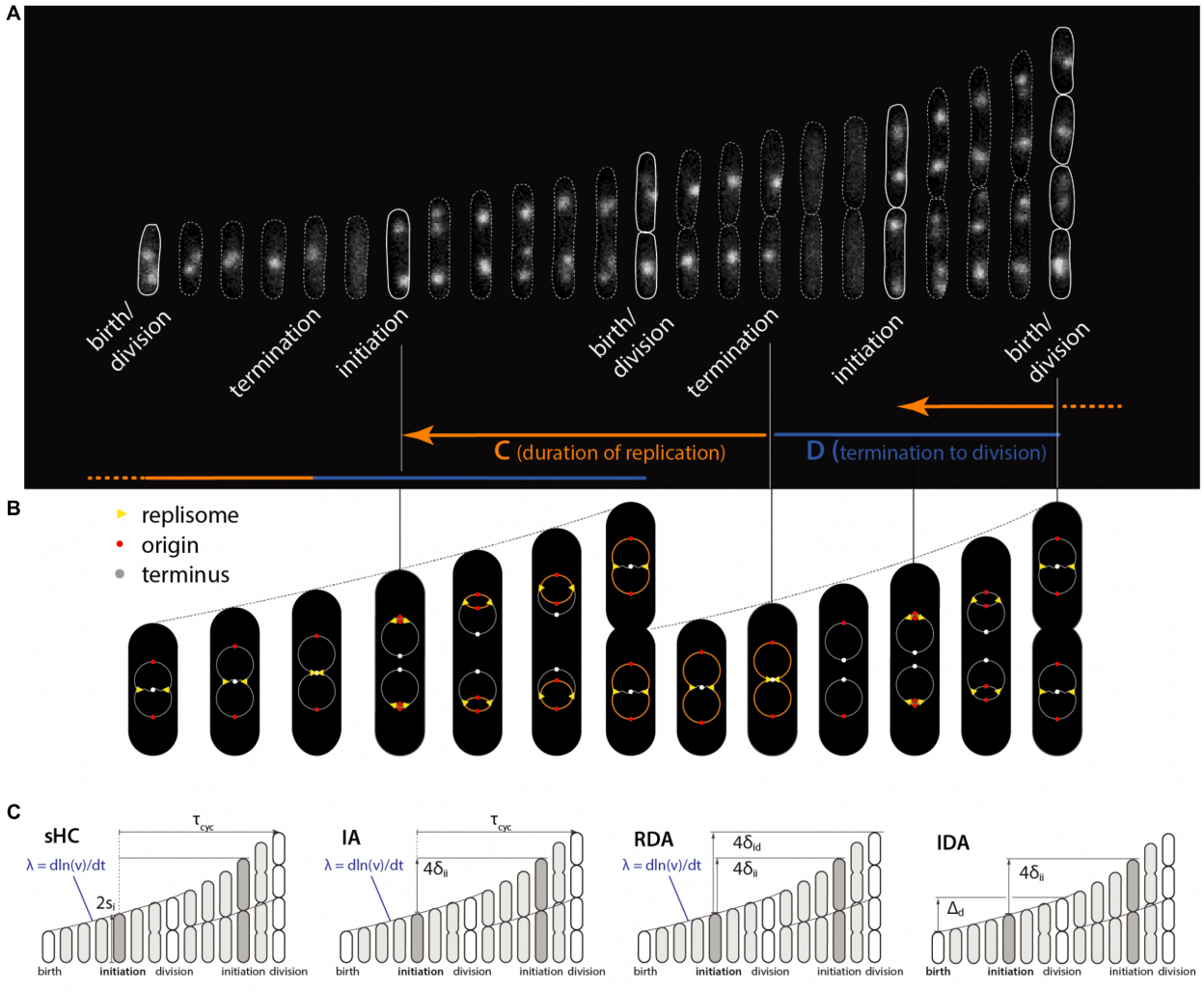


FIGURE 1 | Physiological parameters that can be measured from single-cell experiments. **(A)** Time-lapse images of a single *Escherichia coli* cell growing in a microfluidics channel. The cell boundaries are segmented from phase contrast images whereas the replication forks are visualized using a functional fluorescently labeled replisome protein (DnaN-YPet). **(B)** Multifork replication: in most growth conditions, several replication cycles overlap. The direction of the arrows is not the direction of time, but to illustrate that the HC model's core idea is to trace replication initiation backward in time by C+D from division. **(C)** Four models of *E. coli* cell cycle and their control variables, which can be measured from single-cell experiments. The sHC model describes cell size and cell cycle using three parameters: elongation rate $\lambda = d\ln(l)/dt$, where l is the cell length (not shown), $\tau_{cyc} = C+D$, and the initiation size per origin of replication s_i. The IA model uses λ, τ_{cyc} and the added size per origin of replication between consecutive replication initiation events δ_{ii}. The RDA model uses λ, δ_{ii} and the added size per origin of replication from initiation to division δ_{id}. The IDA model uses λ, δ_{ii} and the added size from birth to division Δ_d. Note that both δ_{id} and τ_{cyc} can span multiple generations. The prefactor before s_i, δ_{ii}, δ_{id} reflects multiple replication origins at initiation as depicted above.

follows:

$$S_d^{(n)} = s_i^{(n)} e^{\lambda \tau_{cyc}^{(n)}}, \quad (1)$$

where n denotes the generation index. If we assume that cells elongate exponentially at the growth rate λ (Wang et al., 2010), the number of overlapping cell cycles $p+1$ is completely determined by the relation:

$$S_d = 2\, S_b \Leftrightarrow e^{\lambda \tau_{cyc}} = 2^{p+1} e^{-\lambda a_i}. \quad (2)$$

where a_i is the time duration elapsed between cell birth and replication initiation. It follows that p is the integer part of τ_{cyc}/τ, where $\tau = \ln 2/\lambda$ is the generation time, so that $p+1$ is the number of overlapping cell cycles [unless noted otherwise we will adopt the convention that $X^{(n)}$ denotes the value of a physiological variable in generation n whereas X is the average over the whole lineage].

In the sHC model, consecutive sizes at initiation are correlated through $\rho_i = \rho(s_i^{(n)}, s_i^{(n+1)})$, where $\rho(A,B)$ stands for the Pearson correlation coefficient between variables A and B. In the absence of mother-daughter correlations for all three physiological variables, the cell should behave as a sizer, $\rho_d = \rho(S_d^{(n)}, S_d^{(n+1)}) = 0$. However, additional cross- or auto-correlations among λ, τ_{cyc} and s_i [such as cross-correlations between $s_i^{(n)}$ and $\tau_{cyc}^{(n)}$ and/or mother-daughter

correlations between $s_i^{(n)}$ and $s_i^{(n+1)}$] can have a non-trivial effect on size homeostasis. Analytical expressions for ρ_i, ρ_d and ρ_{id}=(s_i, S_d) are derived in **Supplementary Appendix C**. Importantly, ρ_d is particularly sensitive to the mother-daughter initiation-size correlation $\rho_i = \rho(s_i^{(n)}, s_i^{(n+1)})$ in the sHC model (Si et al., 2019). This prediction motivated an experimental study aiming at perturbing ρ_i by periodic expression of DnaA in order to break balanced biosynthesis for the DnaA protein (one of the two conditions to produce an adder phenotype) and thus break the adder phenotype in *E. coli*. Experiments rejected this prediction from sHC model (Si et al., 2019), since *E. coli* maintained its size homeostasis following the adder behavior despite periodic oscillations of *dnaA* expression level (Si et al., 2019). An important conclusion from the oscillation experiments is that replication initiation and cell division are independently controlled in steady-state conditions in both *E. coli* and *Bacillus subtilis*, thus firmly refuting the particular version of the sHC model.

The Initiation Adder Model

The IA model is a variant of the sHC model in which the constraint on the initiation mass (s_i = constant in all growth conditions) is replaced by an adder mechanism running between consecutive replication initiations (Sompayrac and Maaloe, 1973; Ho and Amir, 2015). Specifically, the cell initiates replication following the adder principle, i.e., the size added per origin between two consecutive initiation cycles, δ_{ii}, is independent of the cell size at initiation (Amir, 2017). Yet, as in the sHC model, the IA model assumes that division is triggered after a fixed duration of time, τ_{cyc}, has elapsed since initiation. The three stochastic control parameters in the IA model are therefore: λ, τ_{cyc} and δ_{ii}. A given cell size at replication initiation determines the next replication initiation event and one division event.

The recursion relation for cell size at division is the same as in the sHC model (Eq. 1). However, this relation is complemented with the following adder recursion relation determining the cell size per origin at replication initiation:

$$s_i^{(n+1)} = \frac{1}{2} s_i^{(n)} + \delta_{ii}^{(n)}. \quad (3)$$

As before, $\lambda^{(n)}$, $\tau_{cyc}^{(n)}$, and $\delta_{ii}^{(n)}$ are random variables associated with the n-th generation. To derive statistical properties of the IA model (**Table 1**), we will assume that the $\delta_{ii}^{(n)}$ are independent Gaussian variables. At steady-state, Eq. (3) implies that $s_i = 2\delta_{ii}$. Therefore, the number of overlapping cell-cycles is also determined by Eq. (2) in the IA model (namely, p+1).

Cell sizes at consecutive initiations are correlated as $\rho_i = 1/2$ (**Supplementary Appendix A**). Therefore the IA model can be seen as a specific case of the sHC model, for which there is no cross-correlations between physiological variables, and for which the only non-zero auto-correlation is $\rho_i = 1/2$. In general, $\rho_d < 1/2$, and it only reproduces the adder correlation in the deterministic limit where $\lambda\tau_{cyc}$ is a constant.

The Replication Double Adder Model

The RDA model states that the cell simultaneously follows two types of adder. The first adder is between two consecutive initiation cycles ("initiation adder"), same as in the IA model. The second adder states that the size added between initiation and division is independent of the cell size at initiation ("initiation-to-division" adder). This second initiation-to-division adder makes the RDA model different from the IA model, although both models can be considered initiation-centric. This model was developed to explain one specific data set with non-overlapping cell cycles in *E. coli* (Witz et al., 2019). In section "Test of the Models Against Data," we will use the same statistical analysis method that was used in Witz et al. (2019) to establish the RDA model.

TABLE 1 | Summary of a few models of the *E. coli* cell cycle.

Model	Definition	Control parameters	Agreement with the adder principle	ρ_i	ρ_d	ρ_{id}
sHC (Si et al., 2019)	$S_d = s_i \exp(\lambda\, \tau_{cyc})$	τ_{cyc}, s_i	Requires presence of cross- or auto-correlations between control parameters.	ρ_i	$\frac{\rho_i\eta_i^2 + \rho_\alpha\eta_\alpha^2 + \rho_i\rho_\alpha\eta_i^2\eta_\alpha^2}{\eta_i^2+\eta_\alpha^2+\eta_i^2\eta_\alpha^2}$	$\frac{\eta_i}{\sqrt{\eta_i^2+\eta_\alpha^2+\eta_i^2\eta_\alpha^2}}$
IA (Ho and Amir, 2015)	$s_i^{(n+1)} - s_i^{(n)}/2 = \delta_{ii}$ $S_d = s_i \exp(\lambda\, \tau_{cyc})$	τ_{cyc}, δ_{ii}	Only when $\lambda\tau_{cyc}$ is non-stochastic.	1/2	$\frac{1}{2}\frac{\eta_i^2}{\eta_i^2+\eta_\alpha^2+\eta_i^2\eta_\alpha^2}$	$\frac{\eta_i}{\sqrt{\eta_i^2+\eta_\alpha^2+\eta_i^2\eta_\alpha^2}}$
CCCP (Micali et al., 2018)	$\ln(s_i^{(n+1)}) = \ln(s_i^{(n)})/2 + A$ $\ln(S_R) = \ln(s_i) + \lambda C$ $\ln(S_H) = \ln(S_H)/2 + B$ $\ln(S_d) = \max(\ln(S_R), \ln(S_H))$	A, B, C	Yes (by adjusting f, σ_H and σ_R).	1/2	$\frac{\sigma_H^2 f^2/2 + \sigma_i^2(1-f)^2/2}{\sigma_H^2 f + \sigma_R^2(1-f) + f(1-f)(\mu_H-\mu_R)^2}$	$\frac{(1-f)\sigma_i}{\sqrt{\sigma_H^2 f + \sigma_R^2(1-f) + f(1-f)(\mu_H-\mu_R)^2}}$
IDA (Si et al., 2019)	$s_i^{(n+1)} - s_i^{(n)}/2 = \delta_{ii}$ $S_d^{(n+1)} - S_d^{(n)}/2 = \Delta_d$	δ_{ii}, Δ_d	Yes.	1/2	1/2	0
RDA (Witz et al., 2019)	$s_i^{(n+1)} - s_i^{(n)}/2 = \delta_{ii}$ $S_d^{(n+1)} - s_i^{(n)} = 2\delta_{id}$	δ_{ii}, δ_{id}	Only when δ_{id} is non-stochastic.	1/2	$\frac{1}{2}\left(1+3\frac{\sigma_{id}^2}{\sigma_{ii}^2}\right)^{-1}$	$\left(1+3\frac{\sigma_{id}^2}{\sigma_{ii}^2}\right)^{-1/2}$

The definition column indicates the equations defining the division and replication cycles. The control parameters are summarized in the next column. In the three rightmost columns we give the three correlations ρ_i, ρ_d, and ρ_{id}. We have used the following variables: (i) σ_{ii}^2: variance of δ_{ii}, (ii) σ_{id}^2: variance of δ_{id}, (iii) μ_i: mean of s_i, (iv) σ_i^2: variance of s_i, (v) μ_α: mean of $\alpha=\exp(\lambda\tau_{cyc})$, (vi) σ_α^2: variance of α, (vii) $\eta_i=\sigma_i/\mu_i$ is the coefficient of variation (CV) of s_i, (viii) $\eta_\alpha=\sigma_\alpha/\mu_\alpha$ is the CV of α, (ix) μ_H: mean of $\ln(S_H)$, (x) σ_H^2: variance of $\ln(S_H)$, (xi) μ_R: mean of $\ln(S_R)$, (xii) σ_R^2: variance of $\ln(S_R)$.

In the RDA model, the cell size per origin of replication, s_i, follows the same recursion relation as in the IA model (Eq. 3). As for the initiation-to-division adder, the cell size at division is determined by the following recursion relation:

$$S_d^{(n)} = s_i^{(n)} + 2\delta_{id}^{(n)}, \quad (4)$$

where δ_{id} represents the added size per origin of replication from initiation to division.

While Eq. (4) is straightforward to understand for a non-overlapping cell cycle, it is worth checking its validity for overlapping cell cycles. Let $S_i^{(n)}$ be the cell size at initiation for the n-th generation. Let us first emphasize how $S_i^{(n)}$ is measured. In the sHC model, $S_i^{(n)}$ is measured at a duration of time $\tau_{cyc}^{(n)}$ before division occurs. For a non-overlapping cell cycle, $\tau_{cyc}^{(n)} < \tau^{(n)}$, therefore $S_i^{(n)}$ is measured in generation n. For two overlapping cell cycles $2\tau^{(n)} > \tau_{cyc}^{(n)} > \tau^{(n)}$, initiation therefore occurs in the (n-1)-th generation, meaning that $S_i^{(n)}$ refers to a size measured in the mother cell (i.e., generation n-1) as shown in **Figure 1**. Cells are born with 2 origins of replications, therefore we have $S_i^{(n)} = 2s_i^{(n)}$. In this example, the mass synthesized between two consecutive replication initiation events must take into account one division event. Back to the RDA model, and using the same convention, the total added size from initiation to division is $2S_d^{(n)} - S_i^{(n)} = 4\,\delta_{id}^{(n)}$. The factor of 4 accounts for the 4 origin of replications present after replication initiation. Dividing by 2, we obtain Eq. (4). This reasoning generalizes to any number of overlapping cell cycles. From Eq. (4), we also obtain that the average cell size at division is $S_d = 2\,(\delta_{ii} + \delta_{id})$. An argument similar to Eq. (2) yields the number of overlapping cell cycles p+1 as a function of the mean of the physiological variables: p is the integer part of $\log_2(1 + \delta_{id}/\delta_{ii})$.

The IA model is not compatible with size convergence by the adder principle. While ρ_i =1/2 as in the IA model, the division size mother-daughter correlation is given by:

$$\rho_d = \frac{1}{2}\left(1 + 3\frac{\sigma_{id}^2}{\sigma_{ii}^2}\right)^{-1} < \frac{1}{2}, \quad (5)$$

where σ_{ii}^2 and σ_{id}^2 are the variances for δ_{ii} and δ_{id}, respectively (see **Supplementary Appendix A**). Since the adder principle is equivalent to ρ_d = 1/2 (see **Supplementary Appendix A**), the IA model converges to the adder only in the deterministic limit $\sigma_{id} \to 0$. In addition, we can also compute the correlation between initiation size per origin and division size $\rho_{id} = \rho(s_i^{(n)}, S_d^{(n)})$ and obtain:

$$\rho_{id} = \left(1 + 3\frac{\sigma_{id}^2}{\sigma_{ii}^2}\right)^{-1/2}. \quad (6)$$

The Independent Double Adder Model

The IDA model states that, in steady state, initiation and division independently follow the adder principle. That is, the size added between two consecutive initiations is independent of the size at initiation (as in the IA and RDA models), whereas the size added between two division cycles is independent of the cell size at birth (or division). The recursion relation for the division size can be written as:

$$S_d^{(n+1)} = \frac{1}{2}S_d^{(n)} + \Delta_d^{(n)}. \quad (7)$$

It follows that the average cell size at division is $S_d = 2\,\Delta_d$. An argument similar to Eq. (2) yields the number of overlapping cell cycles p+1, where p is the integer part of $\log_2(\Delta_d/\delta_{ii})$.

We have ρ_i =1/2 and ρ_d = 1/2 as expected by the definition of the model. Furthermore, since initiation and division follow two independent processes (Eqs. 3 and 7), division and initiation sizes are independent from each other, namely $\rho_{id} = 0$.

Mechanistically, the IDA model is based on (i) balanced biosynthesis of cell-cycle proteins and (ii) their accumulation to respective threshold numbers to trigger initiation and division (Si et al., 2019).

The Concurrent Cell-Cycle Processes Model

The CCCP model is an adaptable model with several adjustable parameters (as in the sHC model) and lies somewhere in between the IA and the IDA model. The adaptability is analogous to the presentation by Amir (2014) so that the model can be continuously adjusted between sizer and timer depending on the mother-daughter size correlations between −1 and +1. To ensure 1-1 correspondence between the replication cycle and the division cycle, the model explicitly implements a constraint that division must wait until after replication termination. Biologically, the model follows the view by Boye and Nordström (2003).

We discuss the specific case of the adder by fixing the mother-daughter size correlation coefficients to 1/2 as explained throughout this section "Statistical Properties of Five Bacterial Cell-Size and Cell-Cycle Control Models." That is, the cell size at initiation follows the recursion relation:

$$\ln(s_i^{(n+1)}) = \frac{1}{2}\ln(s_i^{(n)}) + A^{(n)}, \quad (8)$$

where $A^{(n)}$ is the logarithmic added size between consecutive replication initiations. As mentioned above, the CCCP model was originally introduced in a more general form than Eq. (8), with an adjustable correlation parameter (see **Supplementary Appendix D**). However, as explained by the authors a value of 1/2 is the most consistent with experimental data. Equation (8) is very similar to Eq. (3): it is an adder on the logarithmic sizes rather than on the actual sizes at replication initiation. Denoting C the time to replicate the chromosome, a candidate size for the division size is:

$$\ln(S_R^{(n)}) = \ln(s_i^{(n)}) + \lambda C^{(n)}, \quad (9)$$

where as before λ is the elongation rate. If chromosome replication was the only process determining the size at division, $S_R^{(n)}$ would be the division size. However, another process, namely the division adder, is constraining the division size, resulting in a second candidate size for division:

$$\ln(S_H^{(n)}) = \frac{1}{2}\ln(S_H^{(n-1)}) + B^{(n)}, \quad (10)$$

where B is the added logarithmic size between consecutive division adder cycles. Equation (10) is similar to Eq. (7) and represents the division adder. Finally, cell size at division is determined by the slowest of the two processes from Eqs. (9) and (10):

$$\ln(S_d^{(n)}) = \max(\ln(S_R^{(n)}), \ln(S_H^{(n)})). \quad (11)$$

Equation (11) simply means that division should start only after replication termination. Denoting f as the fraction of cases in which division size is limiting (namely $S_R < S_H = S_d$), the average time elapsed between replication initiation and cell division can be expressed as [assuming that $< \ln(x) > \approx \ln(< x >)$ (Ho and Amir, 2015)]:

$$\tau_{cyc} = (1-f)C + f\frac{2(B-A)}{\lambda}, \quad (12)$$

where A, B, C stand for means. Therefore, the number of overlapping cell cycles $p+1$ is determined by Eq. (2). Equation (12) has a functional dependence on growth rate compatible with experimental reports (Wallden et al., 2016).

Similarities and Differences Between the Stochastic Helmstetter-Cooper, Initiation Adder, Replication Double Adder, Independent Double Adder, and the Concurrent Cell-Cycle Processes Models

The question of implementation point for cell size control has been controversial in the past. In the sHC, IA, and RDA models, replication initiation is the implementation point of cell size control. By contrast, the IDA and CCCP models assume that the division and replication cycles are controlled by independent processes.

These models reflect a major challenge for identifying a cell-size control model that is compatible with the new plethora of high-throughput single-cell data (Wang et al., 2010; Wallden et al., 2016). Although the sHC and IA models can be dismissed by experimental evidence (section "Comparison of Correlations From Model Predictions and From Experimental Data"), the other models require more thorough analysis. For example, in contrast to the IDA model, the RDA model only ensures the initiation adder and it only reproduces the division adder behavior in the deterministic limit where δ_{id} is constant. The essential difference between the IDA and RDA models comes from the correlation between the size per oriC at initiation and the added size per oriC from initiation to division. Specifically, $\rho(s_i, \delta_{id})$ is zero for the RDA model whereas it takes negative values for the IDA.

In the next section, we test these models against data in more detail.

TEST OF THE MODELS AGAINST DATA

Description of the Experimental Data Used in This Study

We use datasets from our previous studies for *E. coli* and *B. subtilis* (Sauls et al., 2019; Si et al., 2019). We have also performed additional experiments for this study (see **Supplementary Methods**). All data and numerical analysis are available (Supplementary Information: Numerical Methods). In total, we have 15 experimental datasets from our studies. We have also analyzed the 4 experimental datasets made available by Witz et al. (2019).

Comparison of Correlations From Model Predictions and From Experimental Data

We first set out to test the different cell-cycle and cell-size control models. Specifically, we computed the four correlations

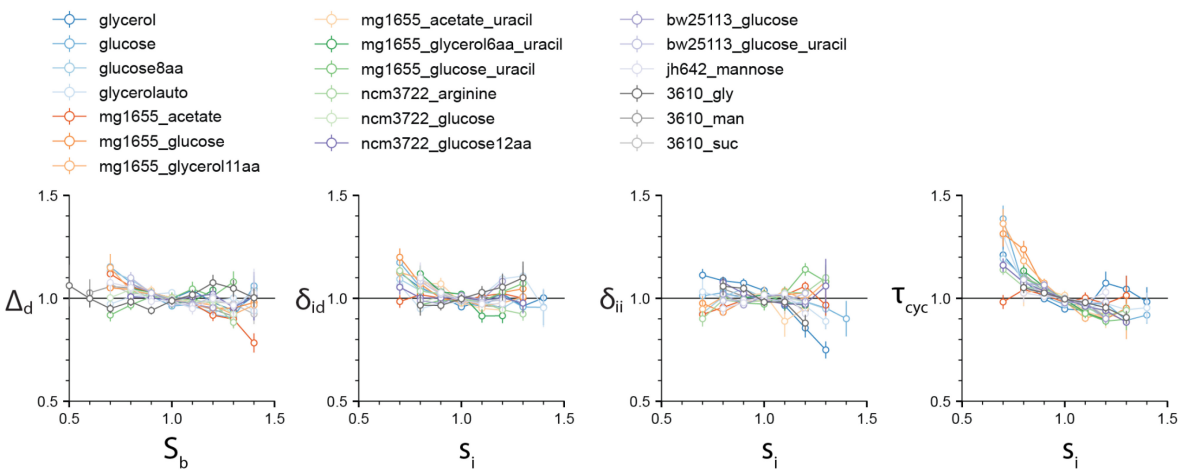


FIGURE 2 | The four correlations $\rho(S_b, \Delta_d)$, $\rho(s_i, \delta_{id})$, $\rho(s_i, \delta_{ii})$, and $\rho(s_i, \tau_{cyc})$ are computed for the 4 experimental datasets by Witz et al. (2019), and 15 experimental datasets that we produced (see **Supplementary Methods**). While the first three vanish for most experimental data, the $\rho(s_i, \tau_{cyc})$ displays a consistent negative correlation, inconsistent with the sHC and IA models. Variables in each dataset were normalized by their mean. Numerical values for the Pearson correlation coefficients are given in **Supplementary Data 2** file. Slopes can be inferred from the Pearson correlation coefficients and CVs in the approximation of bivariate Gaussian variables.

$\rho(S_b, \Delta_d)$, $\rho(s_i, \delta_{ii})$, $\rho(s_i, \delta_{id})$, and $\rho(s_i, \tau_{cyc})$ (**Figure 2**). The correlation $\rho(S_b, \Delta_d)$ is important because $\rho(S_b, \Delta_d)$=0 defines the adder-based cell-size homeostasis. Indeed, $\rho(S_b, \Delta_d)$ is zero in virtually all experimental data. The $\rho(s_i, \delta_{ii})$ is also close to zero, although deviations are seen for some experiments. These results suggest that both the IDA and RDA models are possible. By contrast, the $\rho(s_i, \tau_{cyc})$ correlation shows consistently a negative value. This refutes the sHC and IA models, which both assume that the initiation-to-division duration and the initiation size per origin are independent control parameters, and thus predict $\rho(s_i, \tau_{cyc}) = 0$. In addition, $\rho(s_i, \delta_{id})$ is also close to zero, in favor of the RDA model, but it is also slightly negative for several conditions, in agreement with the IDA model (**Supplementary Appendix B**). We did not test the CCCP model because it is a model to be adjusted to the data. The candidate models are therefore the RDA and IDA models. Hereafter, we focus on these two models.

Statistical Analysis and the Case Study of the *I*-value Analysis

In their recent paper (Witz et al., 2019), Witz et al. tracked replication and division cycles at the single-cell level, using experimental methods similar to previous works (Adiciptaningrum et al., 2015; Wallden et al., 2016; Si et al., 2019). They computed correlations between all pairs of measured physiological parameters, and attempted to identify the set of most mutually uncorrelated physiological variables by computing the "*I*-value," a metric to measure the statistical independence of the measured variables. They then assumed that statistically uncorrelated physiological parameters must represent biologically independent controls. Such approaches previously facilitated the discovery of the adder principle and its formal description (Taheri-Araghi et al., 2015). Based on this correlation analysis or *I*-value analysis, Witz et al. (2019) concluded that the RDA model is the most likely model of the cell-cycle and cell-size control.

To compute the *I*-value for a given model, one needs to identify the control parameters of the model and their characteristic features. For the RDA model, they are (i) the absence of correlation between the cell size at initiation and the added size between initiation and division, namely $\rho(s_i, \delta_{id}) = 0$, and (ii) the absence of correlation between the cell size at initiation and the added size between consecutive initiations, namely $\rho(s_i, \delta_{ii}) = 0$. In addition to these size variables, it is known that the growth rate is mostly independent of the other physiological variables, namely: $\rho(\lambda, \delta_{id}) = 0$ and $\rho(\lambda, \delta_{ii}) = 0$. Witz et al. (2019) hence proposed a scalar metric that summarizes these four correlations being equal to zero, namely the determinant I (or I-value) of the matrix of correlations between the 4 variables s_i, δ_{id}, δ_{ii} and λ (Eq. 13). When $I \ll 1$, some cross-correlations exists and both $\rho(\lambda, \delta_{id})$ and $\rho(\lambda, \delta_{ii})$ cannot vanish. On the other hand, when $I = 1$, the RDA model holds. Although since the work by Cooper and Helmstetter (1968) it has been known that the progression of cell size and cell cycle can be completely described using three variables (Wallden et al., 2016; Si et al., 2017), four variables are necessary here to encompass the correlation structure characterizing the RDA model, as explained by Witz et al. (2019, 2020). In summary, to measure the statistical independence of each set of parameters, the *I*-value analysis needs a correlation matrix of the following form (Eq. 13).

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{pmatrix} \tag{13}$$

The diagonal elements are 1's and the off-diagonal elements are cross-correlations between pairs of parameters. Therefore, if all parameters are statistically independent of each other, the off-diagonal elements should be 0, and the determinant I of the matrix should be 1. Based on this observation, Witz et al. (2019) used the determinant $I \leq 1$ of the matrix as a metric for statistical independence of the hypothetical control parameters, with $I = 1$ being the set of most independent parameters.

In section "Discussion," we will come back to some of the limitations of the *I*-value analysis.

Test of the *I*-value Analysis With Various Models

Our 4-variable *I*-value analysis of the models described in section "Statistical Properties of Five Bacterial Cell-Size and Cell-Cycle Control Models" (except the CCCP) for all 19 datasets is shown in **Figure 3** (top). We computed *I*-values for the following four models: RDA, IDA, IA, and sHC (see **Supplementary Methods**). The results of the analysis indicate that 7 out of our 15 experiments support the IDA rather than the RDA model, whereas 1 supports the IA model. Furthermore, when we applied the same analysis to all 4 datasets from Witz et al. (2019), we found that all 4 experiments support the IDA model. Note that Witz et al. (2019) had only analyzed one dataset. The sHC and IA models were included for completeness, and they show systematically lower *I*-values (except in one *B. subtilis* condition in which the IA models had the largest score for reasons we do not understand). Overall, the results of this analysis suggest that the IDA model is most consistent with the 19 datasets we have analyzed.

DISCUSSION

Limitations of the *I*-value Analysis

We noticed that the *I*-values for the RDA and IDA models were in general very close, although these two models point to two fundamentally different mechanisms of the cell cycle. We therefore asked to what extent the *I*-value analysis could be used to effectively identify meaningful models of the cell cycle. In the spirit of the ranking performed in the study by Witz et al. (2019), we considered all possible combinations of 4 among 18 physiological variables (see **Supplementary Methods**; see **Supplementary Figure 1**), and computed the *I*-values for each of Witz et al. (2019) 4 datasets and for each of our 15 datasets. Although the IDA and RDA models have high scores, we found

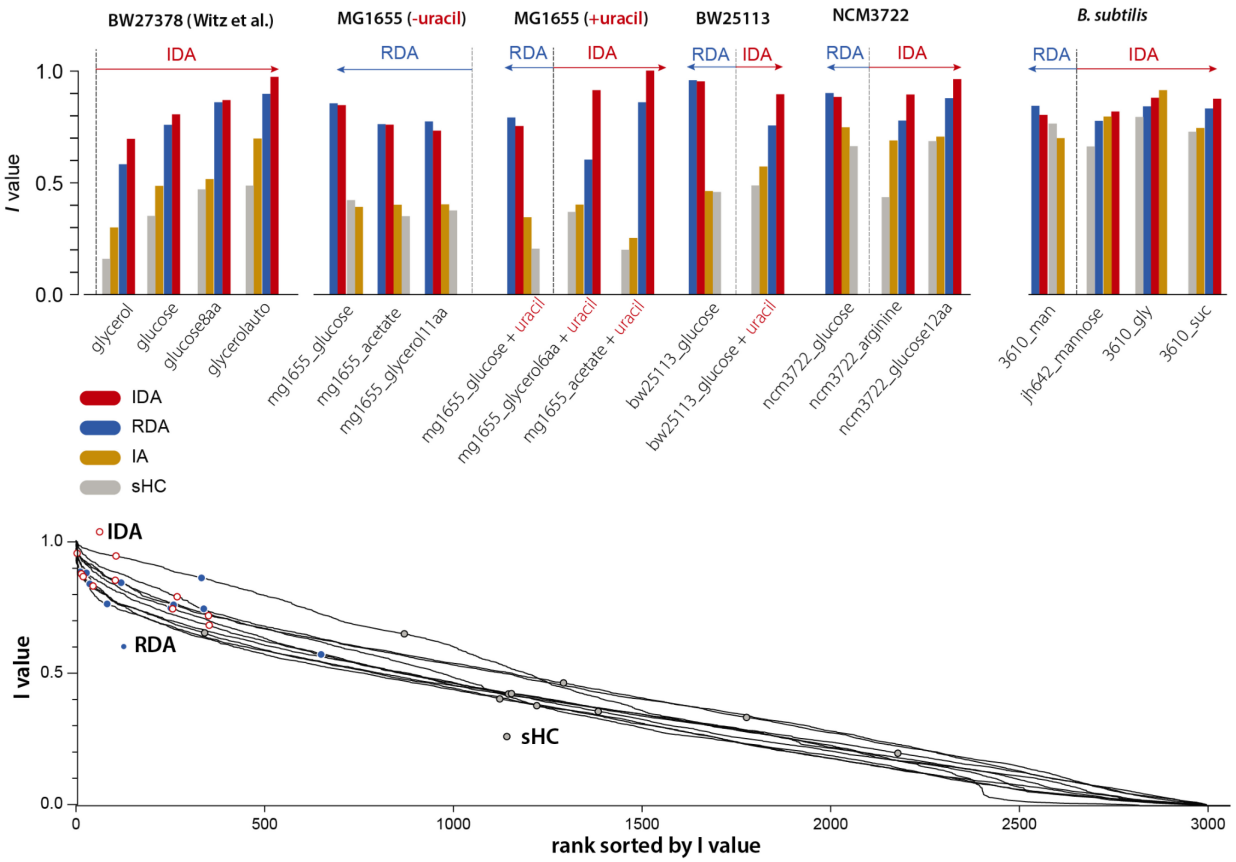


FIGURE 3 | We applied the *I*-value analysis proposed by Witz et al. (2019) to 4 models (RDA, IDA, IA, and sHC) using 15 experimental datasets that we produced (see **Supplementary Methods**), and the 4 datasets published by Witz et al. (2019). In the top bar graphs, the 4 variables are λ, Δ_d, δ_{ii}, S_b for both the IDA model and λ, δ_{id}, δ_{ii}, s_i for the RDA model (see the choice of variables in **Supplementary Methods**). The length of the arrows indicate how much IDA or RDA model is favored by the data.

that many other combinations have higher *I*-values, including combinations that do not correspond to any meaningful model of the cell cycle (**Figure 3** bottom). Since *I*-values cannot be used to distinguish sound from unsound models of the cell-cycle, we conclude that this analysis lacks predictive power.

Furthermore, the *I*-value analysis can only be employed to compare models with the same number of variables in defining correlations. To see this, let us consider the RDA and IDA models in **Figure 4**. The RDA model can be defined by the 3 parameters $\{\delta_{ii}, \delta_{id}, \lambda\}$ (**Figure 4A**). Indeed, from an initial condition consisting of an initiation size, only those 3 parameters need to be known at each generation to construct a whole lineage. Yet the defining correlations of the RDA model are given by $\rho(\lambda, \delta_{id}) = 0$ and $\rho(\lambda, \delta_{ii}) = 0$. Thus, we need a total of 4 variables $\{s_i, \delta_{ii}, \delta_{id}, \lambda\}$ to characterize the RDA model, which leads to the 4x4 correlation matrix shown in **Figure 4C**.

The problem with the above procedure is that the size of the correlation matrix becomes model dependent, and thus the *I*-value analysis cannot compare different models. For example, following the same reasoning, the IDA model would require 5 variables for the *I*-value analysis, because of the two defining correlations (s_i, δ_{ii}) and (S_b, Δ_d) (**Figure 4B**). Therefore, in addition to the three independent control parameters $\{\delta_{ii}, \Delta_d, \lambda\}$, the *I*-value analysis would require two additional parameters $\{s_i, S_b\}$ from the defining correlations. The resulting correlation matrices would then be 5x5 from $\{s_i, S_b, \delta_{ii}, \Delta_d, \lambda\}$ instead of 4x4 (**Figure 4C**). Since *I*-values obtained from correlation matrices of different sizes cannot be meaningfully compared, the *I*-value analysis is fundamentally limited to compare a specific class of models.

Finally, it is worthwhile mentioning that the *I*-value analysis employed by Witz et al. (2020) is only valid for non-overlapping cell cycles.

The Replication Double Adder Model Does Not Produce the Adder Principle

From Eq. (5), size convergence according to the RDA model is incompatible with the adder principle. More specifically, in the presence of fluctuations, the RDA model is skewed toward a sizer behavior. Using experimentally measured values for the variance of δ_{ii} and δ_{id}, we computed ρ_d according to Eq. (5) (**Figure 5**). The RDA model would predict a deviation from the adder principle, in contradiction to several experimental results

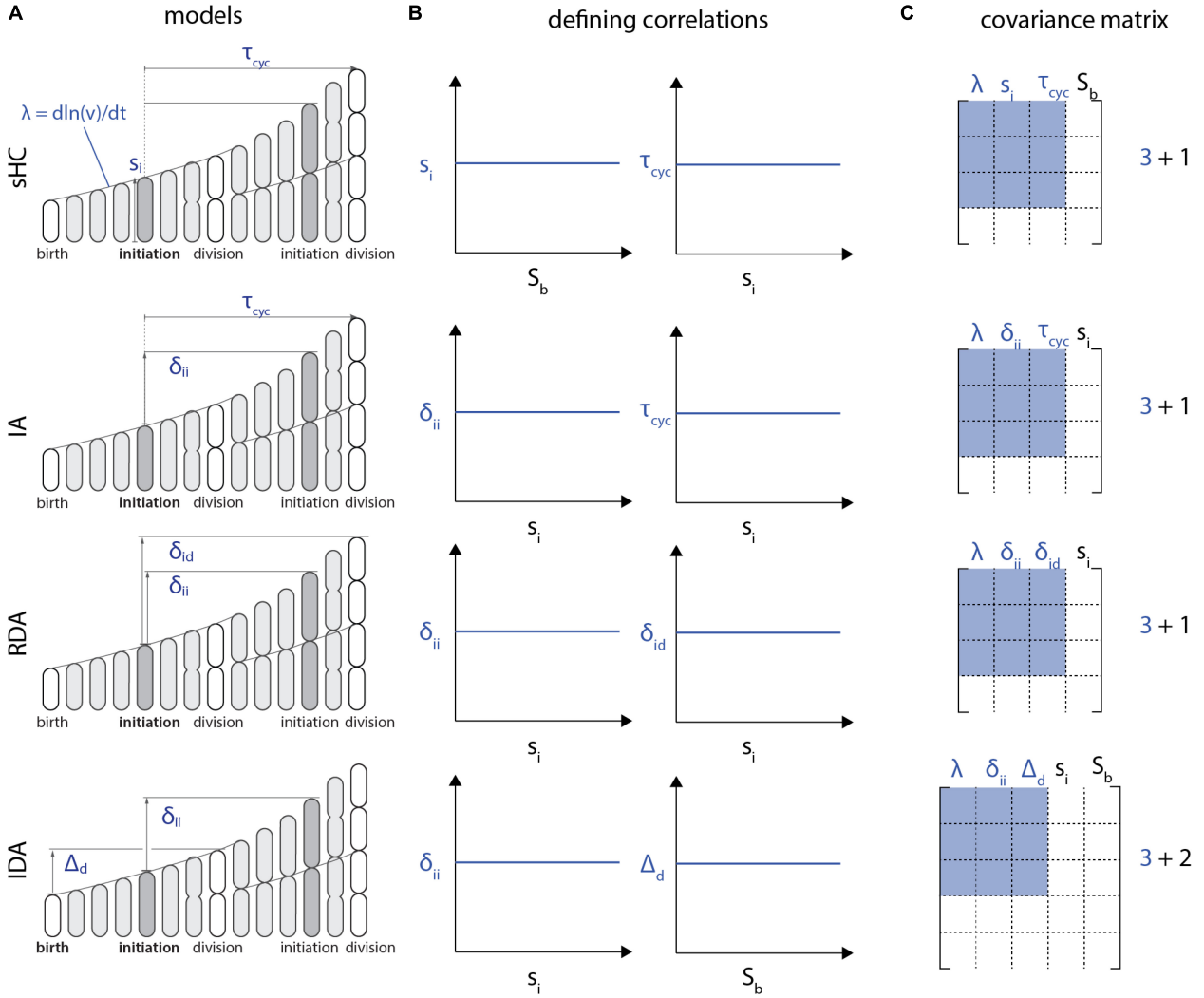


FIGURE 4 | Each model **(A)** is characterized by a set of two correlations **(B)**. Note that these defining correlations require 1 additional parameter S_b for the sHC model, 1 additional parameter s_i for the RDA and the IA models, whereas the IDA model requires 2 additional parameters, s_i and S_b. **(C)** As a result, the covariance matrix for the *l*-value analysis of these models, according to Witz et al. (2019), would be 4x4 for the sHC, IA and RDA models and 5x5 for the IDA model. Therefore, the *l*-values of these models cannot be meaningfully compared.

(Campos et al., 2014; Taheri-Araghi et al., 2015; Le Treut et al., 2020; Treut et al., 2020).

Witz et al. (2019) simulation of the RDA showed a good agreement with the experimental data. Yet further investigation showed that the agreement was a direct consequence of introducing yet another adjustable parameter, namely the variance of the septum position (Le Treut et al., 2020; Treut et al., 2020). Indeed, for perfectly symmetric division, the simulation results also show deviatiations from experimental adder behavior (**Supplementary Figure 2**), in agreement with Eq. (5).

Mechanistic Origin of the Independent Double Adder Model

The IDA model is a mechanistic model based on two experimentally verified hypotheses. First, the cell cycle proteins are produced in a balanced manner (i.e., the synthesis rate of each protein is the same as the growth rate of the cell):

$$\frac{dN}{dt} = c^* \frac{dV}{dt}, \tag{14}$$

where N is the protein copy number in the cell, V is the cell volume and c^* is the steady-state protein concentration. Second, initiation or division is triggered when the respective initiator protein reaches a threshold, namely:

$$N = N_0. \tag{15}$$

The adder phenotype is a natural consequence of these two assumptions, provided that the initiator proteins are equally partitioned at division between daughter cells (Si et al., 2019). The requirement that $N_0/2$ proteins must be synthesized between

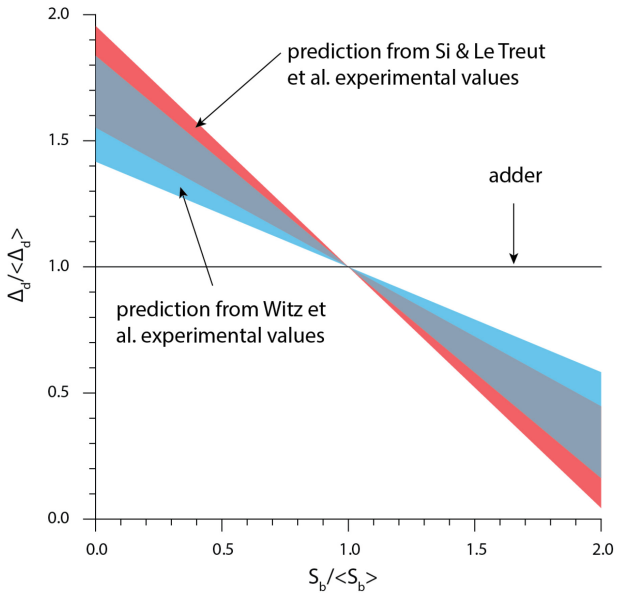


FIGURE 5 | The theory based on the initiation-centric model predicts a more sizer-like behavior. We used Eq. (1) and the experimental values of σ_{id}/σ_{ii} from Si et al. (2019), Witz et al. (2019).

birth and division results in the added volume from birth to division to be:

$$\Delta_d = \frac{N_0}{2c^*}. \quad (16)$$

This model was substantiated in an experimental study showing that perturbing the first condition, namely balanced biosynthesis, was enough to break the adder phenotype (Si et al., 2019). Balanced biosynthesis was perturbed in two orthogonal ways: (i) by oscillating the production rate of the FtsZ protein through periodical induction and (ii) by relieving FtsZ degradation through ClpX inhibition.

A similar mechanism is thought to apply to the initiation process, through the initiator protein DnaA, which accumulates at the origin of replication to trigger replication initiation. An important difference with the division mechanism, however, is that this results in a threshold to be reached at each origin of replication, thus (Eq. 15) is modified to:

$$N = (\#oriC) \times N_0. \quad (17)$$

Provided that once again proteins are equally partitioned at division, Eqs. (2) and (4) result in a fixed added volume per origin of replication between consecutive initiation events, hence the initiation-to-initiation adder δ_{ii}.

Agreement of the Models With the Full Correlation

We computed the three experimental correlations ρ_i, ρ_d, and ρ_{id} and compared them to the predictions of the models shown in **Table 1** (we didn't include the CCCP model because of the extra-parameter f which left the expressions undetermined). Unfortunately, this analysis failed to discriminate between the IDA and RDA models (see **Supplementary Data 2**). The IDA model accounted better for the experimental ρ_d correlation while the RDA model accounted better for the experimental ρ_{id} correlation. This suggests that this correlation study is not sufficient to discriminate between the proposed models. We do not discuss the ρ_i correlation because all 3 out of the 4 models considered predicted the same value of 1/2.

Is the (s_i, S_d) Correlation Real?

The decoupling between replication initiation and cell division was shown by performing independent perturbations to each of those two processes (Si et al., 2019). Specifically, replication initiation was periodically delayed by knocking down the expression of the initiator protein DnaA, yet cell division was left unaffected. Similarly, division was periodically delayed by repressing the expression of the division protein FtsZ, yet replication initiation was left unaffected. This decoupling supports the replication and division process being independent processes, as advocated in the IDA model. However, several nutrient-limitation growth conditions show a (s_i, S_d) correlation which is slightly positive, somewhat in between the zero correlation predicted by the IDA model and the value predicted by the RDA model (**Supplementary Data 2**). This suggests that refinements to the IDA model are still needed to perfectly agree with all experimental correlations. In that regard, some recent developments in cell-cycle modeling are promising (Nieto et al., 2020; Jia et al., 2021). For example, the cell-cycle is divided into several stages, and transitions between consecutive stages occur with a cell volume-dependent rate. Such theories can reproduce more finely the observed correlations at the expense of a larger number of parameters, such as for example the mild deviation from the adder principle toward the sizer behavior in slow growth conditions (Wallden et al., 2016). Similarly, although *B. subtilis* follows the adder principle, it appears the cell cycle can be divided into two phases, one exhibiting a sizer correlation and the other one exhibiting a timer correlation (Nordholt et al., 2020).

PERSPECTIVE AND CONCLUDING REMARKS

While applying a recently proposed correlation analysis, namely the *I*-value analysis, we became aware of some of its limitations, as explained above. In our view, this illustrates some of the caveats one may encounter when applying correlation analysis. While valuable in various contexts, correlation analysis can lead to erroneous conclusions when additional sources of variability such as experimental and measurement errors are not properly taken into account in the analysis. For example, adder correlations can emerge from non-adder mechanisms due to measurement errors in the cell radius (Facchetti et al., 2019). Therefore, while correlation analysis is useful to confront models to experimental data, we believe it is important to seek a molecular understanding of a model of the cell cycle.

The question of whether the implementation point of the cell cycle is birth or replication initiation has a long history. Although cell-size control was initially thought to be division centric because the CV of the division size was smaller than that of the doubling time (Schaechter et al., 1962), many interpreted the HC model (Cooper and Helmstetter, 1968) and Donachie's theoretical observation (Donachie, 1968) as an initiation-centric view for long. The rediscovery of the adder principle, which cannot be explained by the sHC model, has revealed the need to revisit models of the bacterial cell cycle. In this article, we have reviewed one of the latest controversies that has emerged in the field of quantitative bacterial physiology, namely the question of the implementation point of the cell cycle. Based on recent results and single-cell experimental data that we have generated over the last decade, we favor a mixed implementation strategy with two independent adders namely the IDA, as the most likely mechanism ruling the *E. coli* and *B. subtilis* cell cycle. Furthermore, the fact that both initiation and division share the same adder phenotype suggest to us that they also must share the same mechanistic principles. That is, initiation and division must require (i) balanced biosynthesis of their initiator proteins such as DnaA for initiation and FtsZ for division, and (ii) their accumulation to a respective threshold number to trigger initiation and division (Si et al., 2019). Ultimately, these predictions should be tested experimentally to gain mechanistic understanding and their generality beyond correlation analysis.

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/s.

AUTHOR CONTRIBUTIONS

GLT performed research and wrote the manuscript. FS performed research. DL performed research. SJ supervised the project and wrote the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank Martin Howard, Sandeep Krishna, Vahid Shahrezaei, and members of Hwa and Jun labs at UCSD for critical feedback and discussions. This work was supported by National Science Foundation (MCB 2016090).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.721899/full#supplementary-material

REFERENCES

Adiciptaningrum, A., Osella, M., Moolman, M. C., Cosentino Lagomarsino, M., and Tans, S. J. (2015). Stochasticity and homeostasis in the *E. coli* replication and division cycle. *Sci. Rep.* 5:18261. doi: 10.1038/srep18261

Amir, A. (2014). Cell size regulation in bacteria. *Phys. Rev. Lett.* 112:208102. doi: 10.1103/PhysRevLett.112.208102

Amir, A. (2017). Point of view: is cell size a spandrel? *Elife* 6:e22186. doi: 10.7554/eLife.22186

Bertaux, F., von Kügelgen, J., Marguerat, S., and Shahrezaei, V. (2020). A bacterial size law revealed by a coarse-grained model of cell physiology. *PLoS Comput. Biol.* 16:e1008245. doi: 10.1371/journal.pcbi.1008245

Boye, E., and Nordström, K. (2003). Coupling the cell cycle to cell growth. *EMBO Rep.* 4, 757–760. doi: 10.1038/sj.embor.embor895

Campos, M., Surovtsev, I. V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, S. E., et al. (2014). A constant size extension drives bacterial cell size homeostasis. *Cell* 159, 1433–1446. doi: 10.1016/j.cell.2014.11.022

Cooper, S., and Helmstetter, C. E. (1968). Chromosome replication and the division cycle of *Escherichia coli*. *J. Mol. Biol.* 31, 519–540. doi: 10.1016/0022-2836(68)90425-7

Donachie, W. D. (1968). Relationship between cell size and time of initiation of DNA replication. *Nature* 219, 1077–1079. doi: 10.1038/2191077a0

Facchetti, G., Knapp, B., Chang, F., and Howard, M. (2019). Reassessment of the basis of cell size control based on analysis of cell-to-cell variability. *Biophys. J.* 117, 1728–1738. doi: 10.1016/j.bpj.2019.09.031

Harris, L. K., and Theriot, J. A. (2016). relative rates of surface and volume synthesis set bacterial cell size. *Cell* 165, 1479–1492. doi: 10.1016/j.cell.2016.05.045

Ho, P.-Y., and Amir, A. (2015). Simultaneous regulation of cell size and chromosome replication in bacteria. *Front. Microbiol.* 6:662. doi: 10.3389/fmicb.2015.00662

Iyer-Biswas, S., Wright, C. S., Henry, J. T., Lo, K., Burov, S., Lin, Y., et al. (2014). Scaling laws governing stochastic growth and division of single bacterial cells. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15912–15917. doi: 10.1073/pnas.1403232111

Jia, C., Singh, A., and Grima, R. (2021). Cell size distribution of lineage data: analytic results and parameter inference. *iScience* 24:102220. doi: 10.1016/j.isci.2021.102220

Jun, S., and Taheri-Araghi, S. (2015). Cell-size maintenance: universal strategy revealed. *Trends Microbiol.* 23, 4–6. doi: 10.1016/j.tim.2014.12.001

Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology-history, recent progress, and the future with focus on cell size control: a review. *Rep. Prog. Phys.* 81:056601. doi: 10.1088/1361-6633/aaa628

Le Treut, G., Si, F., Li, D., and Jun, S. (2020). Comment on "Initiation of chromosome replication controls both division and replication cycles in *E. coli* through a double-adder mechanism. *bioRxiv* [Preprint]. doi: 10.1101/2020.05.08.084376

Long, Z., Nugent, E., Javer, A., Cicuta, P., Sclavi, B., Cosentino Lagomarsino, M., et al. (2013). Microfluidic chemostat for measuring single cell dynamics in bacteria. *Lab Chip* 13, 947–954. doi: 10.1039/c2lc41196b

Micali, G., Grilli, J., Osella, M., and Lagomarsino, M. C. (2018). Concurrent processes set E. coli cell division. *Sci. Adv.* 4:eaau3324. doi: 10.1126/sciadv.aau3324

Moffitt, J. R., Lee, J. B., and Cluzel, P. (2012). The single-cell chemostat: an agarose-based, microfluidic device for high-throughput, single-cell studies of bacteria and bacterial communities. *Lab Chip* 12, 1487–1494. doi: 10.1039/c2lc00009a

Nieto, C., Arias-Castro, J., Sánchez, C., Vargas-García, C., and Pedraza, J. M. (2020). Unification of cell division control strategies through continuous rate models. *Phys. Rev. E* 101:022401. doi: 10.1103/PhysRevE.101.022401

Nordholt, N., van Heerden, J. H., and Bruggeman, F. J. (2020). Biphasic cell-size and growth-rate homeostasis by single *Bacillus subtilis* cells. *Curr. Biol.* 30, 2238.e–2247.e. 2238–2247.e5*, doi: 10.1016/j.cub.2020.04.030

Sauls, J. T., Cox, S. E., Do, Q., Castillo, V., Ghulam-Jelani, Z., and Jun, S. (2019). Control of *Bacillus subtilis* replication initiation during physiological transitions and perturbations. *MBio* 10, e02205–19. doi: 10.1128/mBio.02205-19

Schaechter, M., Williamson, J. P., Hood, J. R. Jr., and Koch, A. L. (1962). Growth, cell and nuclear divisions in some bacteria. *J. Gen. Microbiol.* 29, 421–434. doi: 10.1099/00221287-29-3-421

Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. *Science* 330, 1099–1102. doi: 10.1126/science.1192588

Serbanescu, D., Ojkic, N., and Banerjee, S. (2020). Nutrient-dependent trade-offs between ribosomes and division protein synthesis control bacterial cell size and growth. *Cell Rep.* 32:108183. doi: 10.1016/j.celrep.2020.108183

Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., et al. (2017). Invariance of initiation mass and predictability of cell size in *Escherichia coli*. *Curr. Biol.* 27, 1278–1287. doi: 10.1016/j.cub.2017.03.022

Si, F., Treut, G. L., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic origin of cell-size control and homeostasis in bacteria. *Curr. Biol.* 29, 1760–1770.e7. doi: 10.1016/j.cub.2019.04.062

Sompayrac, L., and Maaloe, O. (1973). Autorepressor model for control of DNA replication. *Nat. New Biol.* 241, 133–135. doi: 10.1038/newbio241133a0

Supplementary Information: Numerical Methods (2021). Available online at: https://github.com/junlabucsd/DoubleAdderArticle/tree/frontiers (commit 4d3b79dc8492324f2ee4fc789d6e9a24a8c98c82)

Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, Js, et al. (2015). Cell-size control and homeostasis in bacteria. *Curr. Biol.* 25, 385–391. doi: 10.1016/j.cub.2014.12.009

Treut, G. L., Le Treut, G., Si, F., Li, D., and Jun, S. (2020). Single-cell data and correlation analysis support the independent double adder model in both *Escherichia coli* and *Bacillus subtilis*. *bioRxiv* [Preprint] doi: 10.1101/2020.10.06.315820

Vashistha, H., Kohram, M., and Salman, H. (2021). Non-genetic inheritance restraint of cell-to-cell variation. *Elife* 10:e64779. doi: 10.7554/eLife.64779.sa2

Voorn, W. J., Koppes, L. J. H., and Grover, N. B. (1993). Mathematics of cell division in *Escherichia coli*: comparison between sloppy-size and incremental-size kinetics. *Curr. Top. Mol. Gen.* 1, 187–194.

Wallden, M., Fange, D., Lundius, E. G., Baltekin, Ö, and Elf, J. (2016). The synchronization of replication and division cycles in individual *E. coli* cells. *Cell* 166, 729–739. doi: 10.1016/j.cell.2016.06.052

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., et al. (2010). Robust growth of *Escherichia coli*. *Curr. Biol.* 20, 1099–1103. doi: 10.1016/j.cub.2010.04.045

Witz, G., Julou, T., and van Nimwegen, E. (2020). Response to comment on "Initiation of chromosome replication controls both division and replication cycles in *E. coli* through a double-adder mechanism. *bioRxiv* [Preprint] doi: 10.1101/2020.08.04.227694

Witz, G., van Nimwegen, E., and Julou, T. (2019). Initiation of chromosome replication controls both division and replication cycles in through a double-adder mechanism. *Elife* 8:e48063. doi: 10.7554/eLife.48063.sa2

Zheng, H., Bai, Y., Jiang, M., Tokuyasu, T. A., Huang, X., Zhong, F., et al. (2020). General quantitative relations linking cell growth and the cell cycle in *Escherichia coli*. *Nat. Microbiol.* 5, 995–1001. doi: 10.1038/s41564-020-0717-x

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 October 2021
doi: 10.3389/fmicb.2021.733344

DNA-Binding Properties of YbaB, a Putative Nucleoid-Associated Protein From *Caulobacter crescentus*

Parul Pal[1], Malvika Modi[1], Shashank Ravichandran[2], Ragothaman M. Yennamalli[2†] and Richa Priyadarshini[1†]*

[1] Department of Life Sciences, School of Natural Sciences, Shiv Nadar University, Gautam Buddha Nagar, India,
[2] Department of Bioinformatics, School of Chemical and Biotechnology, SASTRA Deemed to be University, Thanjavur, India

OPEN ACCESS

Edited by:
Jianping Xie,
Southwest University, China

Reviewed by:
Catherine Ayn Brissette,
University of North Dakota, United States
Soumitra Ghosh,
University of Lausanne, Switzerland

***Correspondence:**
Richa Priyadarshini
richa.priyadarshini@snu.edu.in

†ORCID:
Richa Priyadarshini
orcid.org/0000-0002-8613-9438
Ragothaman M. Yennamalli
orcid.org/0000-0002-3327-1582

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *30 June 2021*
Accepted: *22 September 2021*
Published: *28 October 2021*

Citation:
Pal P, Modi M, Ravichandran S, Yennamalli RM and Priyadarshini R (2021) DNA-Binding Properties of YbaB, a Putative Nucleoid-Associated Protein From Caulobacter crescentus. Front. Microbiol. 12:733344. doi: 10.3389/fmicb.2021.733344

Nucleoid-associated proteins (NAPs) or histone-like proteins (HLPs) are DNA-binding proteins present in bacteria that play an important role in nucleoid architecture and gene regulation. NAPs affect bacterial nucleoid organization via DNA bending, bridging, or forming aggregates. EbfC is a nucleoid-associated protein identified first in *Borrelia burgdorferi*, belonging to YbaB/EbfC family of NAPs capable of binding and altering DNA conformation. YbaB, an ortholog of EbfC found in *Escherichia coli* and *Haemophilus influenzae*, also acts as a transcriptional regulator. YbaB has a novel tweezer-like structure and binds DNA as homodimers. The homologs of YbaB are found in almost all bacterial species, suggesting a conserved function, yet the physiological role of YbaB protein in many bacteria is not well understood. In this study, we characterized the YbaB/EbfC family DNA-binding protein in *Caulobacter crescentus. C. crescentus* has one YbaB/EbfC family gene annotated in the genome ($YbaB_{Cc}$) and it shares 41% sequence identity with YbaB/EbfC family NAPs. Computational modeling revealed tweezer-like structure of $YbaB_{Cc}$, a characteristic of YbaB/EbfC family of NAPs. N-terminal–CFP tagged $YbaB_{Cc}$ localized with the nucleoid and is able to compact DNA. Unlike *B. burgdorferi* EbfC protein, $YbaB_{Cc}$ protein is a non-specific DNA-binding protein in *C. crescentus*. Moreover, $YbaB_{Cc}$ shields DNA against enzymatic degradation. Collectively, our findings reveal that $YbaB_{Cc}$ is a small histone-like protein and may play a role in bacterial chromosome structuring and gene regulation in *C. crescentus.*

Keywords: DNA binding protein, *Caulobacter crescentus*, nucleoid associated protein, YbaB/EbfC family, gene regulation

INTRODUCTION

Similar to eukaryotic organisms, bacteria also pack their genetic material in a very small space. DNA-binding proteins known as nucleoid-associated proteins (NAPs) play a crucial role in nucleoid structuring and controlling gene expression. Although NAPs are referred to as histone-like proteins (HLPs), they are functionally very distinct from eukaryotic histones. Bacterial cells contain a wide variety of NAPs and their expression in the cell varies with growth phase of the culture. NAPs are small, basic proteins that can have the ability to bind to DNA, either as monomers, dimers, or tetramers or in complex with other DNA modulating proteins (Dorman, 2009;

Dillon and Dorman, 2010). The DNA binding occurs either through direct physical interaction or indirectly via binding with other accessory proteins (Dorman, 2014).

Major NAPs involved in nucleoid structuring are HU, IHF, H-NS, Lrp, Fis, and Dps. HU is small, basic, non-sequence specific DNA-binding protein, well conserved in eubacteria (Kamashev and Rouviere-Yaniv, 2000; Bahloul et al., 2001). In dimorphic bacterium, *Caulobacter crescentus*, HU is a 20-kDa heterodimer consisting of HU1 and HU2 subunits (Le et al., 2013; Le and Laub, 2016). HU protein colocalizes with the nucleoid and has uniform distribution among swarmer and stalked cells but is highly clustered in predivisional cells (Lee et al., 2011). Recently, a new NAP, GapR, was discovered in *C. crescentus* (Ricci et al., 2016; Arias-Cartin et al., 2017; Taylor et al., 2017). GapR binds AT-rich regions of the nucleoid (Ricci et al., 2016) and absence of GapR leads to morphological and cell division defects (Arias-Cartin et al., 2017; Taylor et al., 2017).

A new family of NAPs, EbfC/YbaB, was first discovered in *Borrelia* (Babb et al., 2006). The structure of EbfC/YbaB homodimer has been described to have a tweezer-like conformation, with tweezer region ascribed to alpha-helical DNA-binding domain and giving spacer region the ability to fit around double-stranded DNA (Lim et al., 2002; Riley et al., 2009). EbfC protein binding facilitates DNA bending (Riley et al., 2009). In *B. burgdorferi*, EbfC displays sequence specific DNA binding and binds to a palindromic DNA sequence, 5′-GTnAC-3′, where "n" can be any nucleotide (Riley et al., 2009). However, the DNA binding sites EbfC/YbaB orthologs of *E. coli* and *H. influenzae* are not known and are believed to be distinct from that of *B. burgdorferi* EbfC (Cooley et al., 2009). While EbfC/YbaB is almost present in all eubacteria, its physiological role in most bacterial species remains largely unknown. Here, we characterized $YbaB_{Cc}$, the homolog of YbaB/EbfC family DNA-binding protein in *C. crescentus*. $YbaB_{Cc}$ protein shares 41% sequence identity and has distinctive tweezer-like conformation of 'YbaB/EbfC family proteins. Moreover, $YbaB_{Cc}$ colocalizes with nucleoid and is able to compact DNA. $YbaB_{Cc}$ is a non-sequence specific DNA-binding protein with nucleoid-associated function in *C. crescentus.*

MATERIALS AND METHODS

Strains, Media, and Growth Conditions

C. crescentus strains were grown in peptone yeast extract (PYE) medium or M2G minimal medium at 30°C (Stove Poindexter', 1964) and *E. coli* cells were grown in LB medium at 37°C. Xylose (0.3%), glucose (0.2%), or arabinose (0.2%) was added to the growth media as required. Plasmids and strains used in this study are detailed in **Table 1**. Details of strain construction are mentioned in **Supplementary Text**. All media were purchased from Hi-Media Laboratories (Mumbai, India) and antibiotics were obtained from Sigma (United States).

Microscopy and Image Analysis

Expression from *Pxyl* promoter was obtained by addition of xylose to the growth media at an optical density OD_{600} 0.2. For studying localization of $YbaB_{Cc}$, expression of CFP-fused $YbaB_{Cc}$ was obtained from pXCFPN-5 vector in PBP3 temperature-sensitive mutant *C. crescentus* by addition of xylose (0.2%) at an optical density OD_{600} 0.1 in PYE broth. Cultures were incubated at 30°C for 1 h and then shifted to 37°C and grown for 4.5 h. After this, cells were harvested and washed with M2 medium and processed for 4′, 6′-diamidino-2-phenylindole (DAPI) staining. Cell samples (5 μl) were imaged as described before (Dubey and Priyadarshini, 2018) using Nikon Eclipse Ti microscope (United States) equipped with Nikon DS-U3 camera. Images were processed with Adobe Photoshop CS6. Cell length, nucleoid length, and cell width were analyzed with Fiji (ImageJ) software (Schindelin et al., 2012) and Oufti software (Paintdakhi et al., 2015).

4′, 6′-diamidino-2-phenylindole Staining

For DNA compaction studies, $YbaB_{Cc}$ was ectopically expressed in wild-type *E. coli* MG1655 cells, as described previously with necessary modifications (Ghosh et al., 2013; Oliveira Paiva et al., 2019). *C. crescentus* $YbaB_{Cc}$ was cloned in pBAD18 and transformed into *E. coli* MG1655. *E. coli* cells containing $YbaB_{Cc}$ and *E. coli* control strain carrying empty pBAD18 vector were grown overnight at 37°C in LB medium supplemented with 0.2% glucose and kanamycin. Overnight cells were washed twice with LB broth and secondary cultures were induced with 0.2% arabinose or glucose at OD_{600} 0.1. The cells were allowed to grow for 6 h and harvested. Pellets were washed with 1X PBS, fixed in 70% ethanol for 2 min (10% ethanol used in case of colocalization studies with CFP) at room temperature and washed again with 1X PBS. DAPI (1 mg/ml) was added to the cells (1:1,000 dilution) and incubated in dark for 15 min at room temperature. After final wash with 1X PBS, cells were observed by microscopy. For DNA compaction studies, images were captured by Nikon Eclipse Ti 2 Confocal Microscope (AR1MP; United States). Images were processed with Adobe Photoshop CS6. Cell length, nucleoid length, and cell width were analyzed with Fiji (ImageJ) software (Schindelin et al., 2012).

Protein Purification

The coding regions of $YbaB_{Cc}$ and $YbaB_{Cc}$ (Tn) were amplified by PCR and inserted into pET28b(+) (Novagen) between *Sac*I and *Bam*HI sites. These recombinant vectors were used to express 6X His-tagged $YbaB_{Cc}$ and $YbaB_{Cc}$ (Tn) proteins in *E. coli* BL21 (DE3) pLysS cells. Recombinant protein expression was obtained by adding 10 μM IPTG to 0.4 OD_{600} bacterial culture for 5 h at 37°C in LB medium supplemented with kanamycin. Cells were harvested by centrifugation (10,000 rpm, 5 min, 4°C) and processed for protein purification as described previously with minor modifications (Dubey and Priyadarshini, 2018). The washed pellets were resuspended in the lysis buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5% glycerol, 5 mM imidazole, 1 mM PMSF) and incubated with 1 mg/ml lysozyme for 45 min in shaker incubator at 37°C. Cells were lysed by sonication and the lysate was clarified by centrifugation (16,000 rpm, 20 min, 4°C). The supernatant was treated with DNaseI (New England Biolabs, United States)

TABLE 1 | Strains and plasmids.

	Relevant genotype or description	Sources/References
C. crescentus		
CB15N	Synchronizable derivative of wild-type strain CB15 (NA 1000)	Evinger and Agabian, 1977
CJW2141	CB15N *cc2570*:pTC67 *ftsI*(Ts)	Costa et al., 2008
RP42	CB15N *xylX*:pXCFPN-5-$ybab_{Cc}$	This work
RP43	CB15N/pPAL4	This work
RP45	CB15N *cc2570*:pTC67 *ftsI*(Ts) *xylX*:pXCFPN-5-$ybab_{Cc}$	This work
E. coli		
DH5α	Φ80 Δ*lacZ*Δ*M15*Δ(lacZYA-argF)U169 *deoR recA1 endAhsdR17 (rk-,mk +) phoA supE44 thi-1 gyrA96 relA1*	Laboratory strain collection
S17	RP4-2, Tc:Mu, KM-Tn7	Simon et al., 1983
BL21 (DE3) pLysS	F^- *ompT* $hsdS_B$ (r_B^-, m_B^-) *gal dcm* (DE3) pLysS(Cam^R)	Laboratory strain collection
MG1655	K-12 F^- λ^- *ilvG*$^-$ *rfb-50 rph-1*	Laboratory strain collection
RP40	MG1655/pPAL1	This work
RP41	BL21 (DE3) pLysS/pPAL3	This work
RP44	BL21 (DE3) pLysS/pPAL5	This work
RP46	MG1655/pBAD18	This work
Plasmids		
pBAD18	Kan^r, DH5α containing empty pBAD18	Guzman et al., 1995
pXCFPN-5	Tetr, vector used for generating N-terminal protein fusions encoded at the *xylX* locus	Thanbichler et al., 2007
pET28b	Protein expression vector, Kan^r	Laboratory strain collection
pJS14	High copy number vector, pBR1MCS derivative, Cmr	Laboratory strain collection
pPAL1	pBAD18 carrying $ybab_{Cc}$, Kan^r	This work
pPAL2	pXCFPN-5 carrying $ybab_{Cc}$fused to CFP	This work
pPAL3	pET28b carrying $ybab_{Cc}$fused to 6X-His at N-terminus	This work
pPAL4	pJS14 carrying $ybab_{Cc}$	This work
pPAL5	pET28b carrying truncated $ybab_{Cc}$fused to 6X-His at N-terminus	This work

for 15 min at 4°C. Supernatant was incubated with Ni^{2+}–NTA agarose beads (Qiagen, Germany) equilibrated in binding buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5% glycerol and 10 mM imidazole) for 6 h at 4°C. The mixture was passed through a gravity-flow column and washed with wash buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5% glycerol, and 20 mM imidazole). 6X His-tagged $YbaB_{Cc}$ was eluted with elution buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5% glycerol, and 300 mM imidazole) in multiple fractions. Eluted fractions were visualized on SDS–PAGE by Coomassie staining (**Supplementary Figure 3**). Protein containing fractions were pooled and given a buffer exchange for Storage buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5% glycerol), using PD-10 Desalting column (Cytiva, United States). Finally, the protein fraction was concentrated using Amicon® Ultra-4 spin-columns (Millipore, United States). Dot blot and Western blot analysis were performed with 1:5,000 dilution anti-His-antibody (Invitrogen, United States; **Supplementary Figure 3**). Protein concentration was estimated by Bradford Method (Bradford, 1976).

Western Blot and Dot Blot Analysis

Western blots were performed as described (Dubey and Priyadarshini, 2018). Cell lysates were separated on SDS–PAGE gels and transferred to a polyvinylidene difluoride (PVDF) membrane using a semi-dry transfer apparatus (Bio-Rad, United States). Dot blot was performed as previously described (Bhat and Rao, 2020), with appropriate modifications. Concentrated and diluted purified $YbaB_{Cc}$ was spotted onto nitrocellulose membrane and samples were allowed to completely air dry. Membranes were incubated with PBST blocking solution containing 5% non-fat milk at room temperature for 3 h. The membranes were then washed and subsequently incubated with 1:5,000 diluted anti-His antibody (Invitrogen, United States) overnight at 4°C. The next day, membranes were given four 10-min washes in PBST and incubated with horseradish peroxidase-conjugated anti-rabbit antibody (Invitrogen, United States) for 1–2 h at room temperature. Blots were washed with PBST and developed with the Bio-Rad Clarity and Clarity Max ECL Western Blotting Substrates according to the manufacturer's protocols. 6X His-tagged (**Supplementary Figures 3B,C, 4**) fusion proteins were confirmed by western blotting for stability.

DNA Protection Assay

To investigate DNA binding and protection ability of $YbaB_{Cc}$ against degradative action of DNases, pBAD18-kan vector was incubated with increasing $YbaB_{Cc}$ protein concentrations (20 min, 25°C) in Storage Buffer. The experiments were performed as described earlier (Datta et al., 2019). In brief, enzyme treatment was given at 37°C for 1 min using 1 Unit of DNaseI (New England Biolabs, United States). The enzyme was inactivated by incubation at 75°C for 15 min followed by protein denaturation at 95°C for 20 min. Reactions were electrophoresed

in 1% agarose gel containing ethidium bromide stain and DNA was visualized in a UV transilluminator.

Electrophoretic Mobility Shift Assay

EMSA was performed according to LightShift Chemiluminescent EMSA Kit Protocols, (Thermo Scientific, United States). In brief, oligonucleotides b-WT (124 bp) from *Borrelia burgdorferi* strain B31 (Riley et al., 2009) and cc-WT (200 bp of $ybab_{Cc}$ gene) from *C. crescentus* (this work) were labeled at 3′ end with Biotin-11-UTP according to the manufacturer's protocol (Biotin 3′ end DNA Labeling Kit, Thermo Scientific, United States). Labeled target probes were incubated with increasing concentrations of purified $YbaB_{Cc}$ in a reaction containing binding buffer (LightShift Chemiluminescent EMSA Kit, Thermo Scientific) and 50% glycerol at 25°C for 30 min. The reactions were electrophoresed in 6% DNA retardation gel (Invitrogen, United States) in 0.5x TBE. Separated DNA products were then electroblotted onto a positively charged Nylon membrane (Thermo Fisher Scientific, United States) using a semi-dry transfer apparatus (Bio-Rad, United States). Cross-linking by UV was done at 120 mJ/cm^2 for 1 min immediately after electroblotting. Detection of DNA and DNA-protein complexes was carried using Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific, United States).

Building Computational Model of $YbaB_{Cc}$ and Its Complex With DNA

The protein sequence of $YbaB_{Cc}$ was retrieved from NCBI (accession id: YP_002515644.1) and was given as input to Robetta tool (Kim et al., 2004),[1] where the comparative modeling approach was used to generate 1,000 sample models and obtained the top five models (Kim et al., 2004). The five models were analyzed to select one model using root mean square deviation as a selection measure. Additionally, the model's confidence score was also taken into consideration for building the protein-DNA complex. The DNA sequence (5′-ATGTAACAGCTGAATGTAACAA-3′) was used to construct a double-stranded B-DNA and the 3D coordinates were obtained using the conformational parameters extracted from fiber-diffraction experimental studies.[2] Subsequently, the protein and DNA were given as input to HADDOCK (van Zundert et al., 2016).

Protein-DNA Docking Studies

The parameters of protein-DNA docking were selected in the manner to perform a blind docking approach, where specific interaction site was not specified (Heté Nyi and van der Spoel, 2002; Teotia et al., 2019; Gondil et al., 2020). Additional parameters were set, such as force constant for center of mass contact restraints (1.0), force constant for surface contact restraints (1.0), radius of gyration (17.78), number of structures of rigid body docking (1,000), number of trials of rigid body minimization (5), sample 180 rotated solutions during rigid body EM (Yes), number of structures for semi-flexible refinement (200), sample 180 rotated solutions during semi-flexible SA (No), Perform final refinement (Yes), number of structures for the final refinement (200), number of structures to analyze (200), Fraction of Common Contacts (FCC) method of clustering method, RMSD cutoff for clustering (0.6), minimum cluster size (4), Non-bonded parameters (OPLX), include electrostatic during rigid body docking (Yes), Cutoff distance to define an hydrogen bind (2.5), cutoff distance to define a hydrophobic contact (3.9), Perform cross-docking (Yes), randomize starting orientations (Yes), perform initial rigid body minimization (Yes), allow translation in rigid body transformation (Yes). The top ranked protein-DNA complex from HADDOCK was analyzed using NUCPLOT (Luscombe et al., 1997).

[1] http://new.robetta.org/

[2] http://www.scfbio-iitd.res.in/software/drugdesign/bdna.jsp

RESULTS

Protein–DNA Complex Indicates Preferential Binding of $YbaB_{Cc}$

EbfC/YbaB homologs are ubiquitous in nearly all eubacterial genomes. *C. crescentus* genome CCNA_00269 is annotated as DNA-binding protein and BLAST analysis revealed 41% sequence identity and 62% similarity with EbfC/YbaB family NAPs. EbfC/YbaB family of proteins act as NAPs and have dimerization domains flanked on either side by DNA-binding domains (Babb et al., 2006; Riley et al., 2009). CCNA_00269 revealed similar domain arrangement (**Figure 1A**). Based on these results we annotated CCNA_00269 gene as $ybab_{Cc}$ and characterized its DNA binding properties in this study. EbfC/YbaB family of NAPs colocalize with the nucleoid and probably play a role in structuring of the bacterial chromosome (Jutras et al., 2012; Wang et al., 2012). To investigate DNA compaction activity of $YbaB_{Cc}$ protein, we introduced full-length $YbaB_{Cc}$ on pBAD18 plasmid into *E. coli*. Ectopic expression of $YbaB_{Cc}$ was induced by addition of arabinose and cells were subsequently stained with DAPI to label the nucleoid (**Figure 1B**). As evident from **Figures 1B–D**, only 51.3% of cell length is occupied by the nucleoid in arabinose-induced cells. In comparison, *E. coli* cells grown with glucose showed 68.4% of the cell length being occupied by nucleoid. Nucleoid compaction observed in presence of glucose could be due to leaky expression from high copy number pBAD18 plasmid (Dubey and Priyadarshini, 2018). In comparison, control RP46 cells carrying empty pBAD plasmid grown in presence of arabinose displayed almost 99% of the cell length being occupied by nucleoid. Our results suggest that $YbaB_{Cc}$ protein is able to compact DNA.

We further searched for YbaB homologs in other alpha-proteobacteria. YbaB homologs were found in most alpha-proteobacteria except, intracellular pathogens belonging to the family Rickettsiaceae (**Figure 2E**). *Rhodospirillum rubrum* and *Paracoccus denitrificans* genome were also devoid of YbaB homologs (**Figure 2E**). EbfC/YbaB family proteins function as homodimers having a unique "tweezer"-like structure (Lim et al., 2002; Cooley et al., 2009). We obtained a 3D model of $YbaB_{Cc}$ protein using Robetta program (Kim et al., 2004). Robetta samples nearly 1,000 models obtained after superposing the

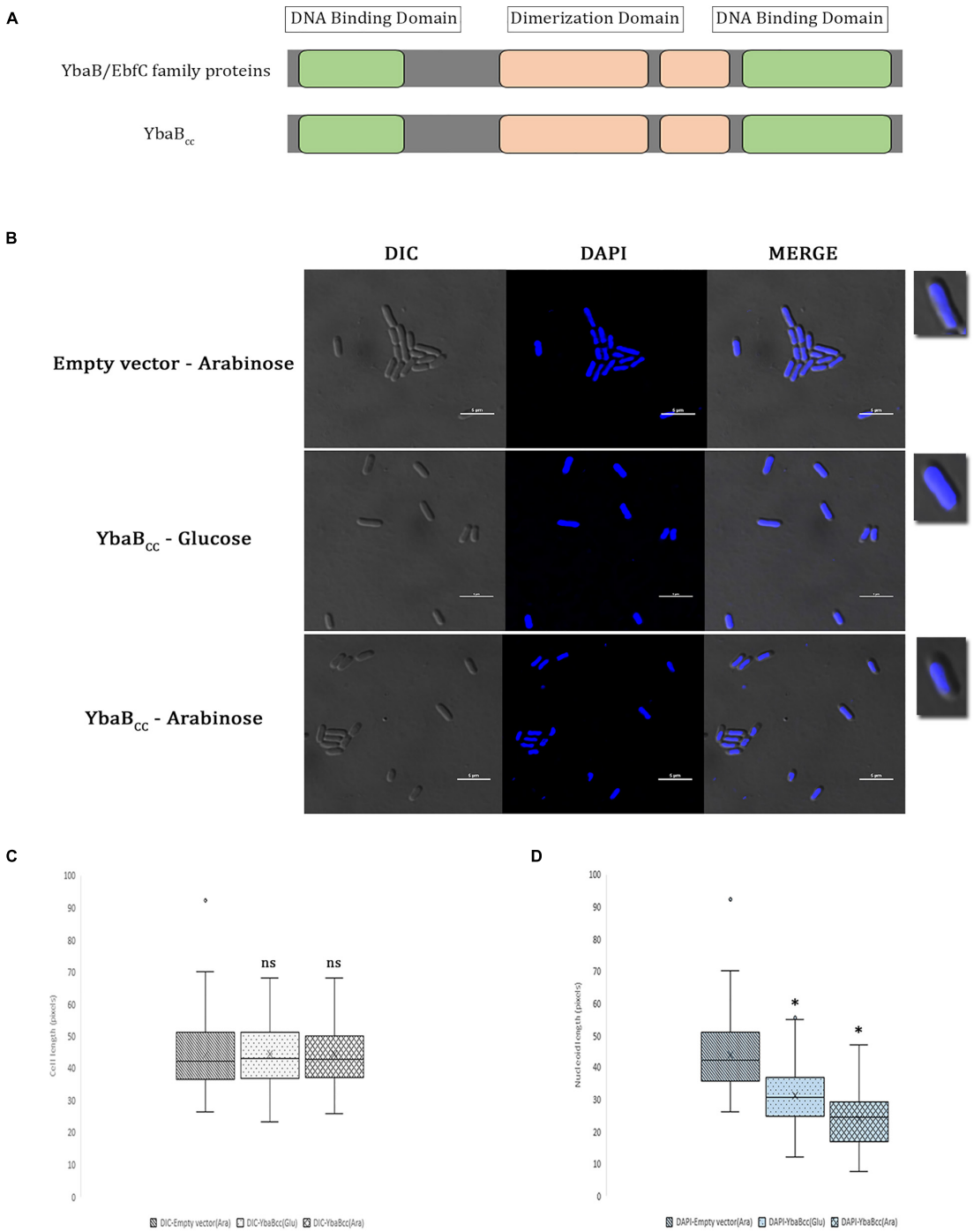


FIGURE 1 | YbaB$_{Cc}$ binds and compacts DNA in *E. coli*. **(A)** Schematic presentation of YbaB$_{Cc}$ protein domain organization. YbaB/EbfC family proteins show similarity in domain structures with YbaB$_{Cc}$. There are two DNA-binding domains present at either terminus of YbaB/EbfC family proteins, which together form a unique tweezer-like structure in order to bind to the target DNA. **(B)** YbaB$_{Cc}$ overexpression in *E. coli* cells. YbaB$_{Cc}$ was expressed ectopically from pBAD promoter in RP40 cells. In wild-type *E. coli*, nucleoid occupies the entire cytoplasmic space, whereas a compacted nucleoid tends to be away from the cell periphery and more toward the center of the bacterial cell. The scale bar represents 5 μM. **(C)** Box and whisker plots of mean cell length. Whiskers represent minimum and maximum cell lengths observed in each strain (1 μM = 32 pixels). Cross (x) represents mean values and horizontal line across boxes represent the median values. Analysis was done using Fiji (ImageJ) software (n, number of cells analyzed = 99 per strain, per condition). **(D)** Box and whisker plots of mean nucleoid length. Whiskers represent minimum and maximum nucleoid lengths observed in each strain (1 μM = 32 pixels). Cross (x) represents mean values and horizontal line across boxes represent the median values. Analysis was done using Fiji (ImageJ) software (n, number of cells analyzed = 99 per strain, per condition). **(C,D)** Same cells were analyzed for measuring cell lengths and nucleoid lengths. YbaB$_{Cc}$ protein bound nucleoid reduces in length as compared to the protein-free nucleoid in non-overexpressing conditions. Box and whisker plots (n = 99 per strain) were generated for statistical analysis with the whiskers including all data points within 1.5*IQR. ns = non-significant. *p < 0.0001 by one-way ANOVA compared to the empty vector pBAD (ara) control.

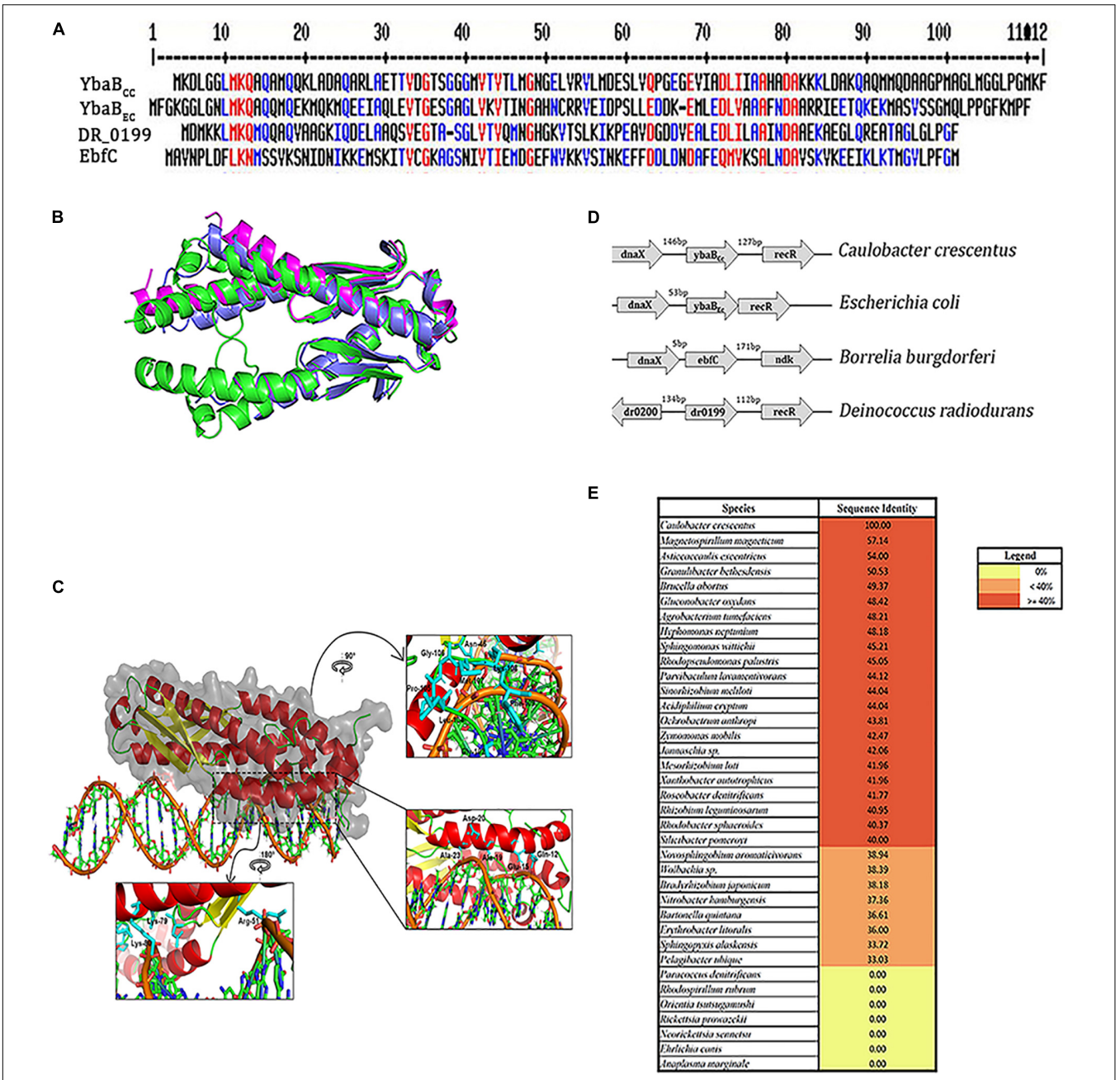


FIGURE 2 | YbaB$_{Cc}$ is conserved across multiple bacterial species. **(A)** Multiple sequence alignment of YbaB/EbfC proteins from different bacterial species using MultAlin software (red—high consensus residues, blue—low consensus residues, at the given position). **(B)** Structural superposition of the modeled YbaB$_{Cc}$ protein with structural homologs. The modeled protein (colored green and shown in cartoon representation) is a homodimer, as seen in the structural homologs from *E. coli* (PDB id: 1PUG) colored light blue. Another structural homolog from *Haemophilus influenza* (PDB id:1J8B) colored magenta is in monomeric form. Image made using PyMOL. **(C)** Protein-DNA docking model of YbaB$_{Cc}$ (shown in cartoon and surface representation colored in secondary structure) is observed to bind to the DNA motif (5′-ATGTAACAGCTGAATGTAACAA-3′) as obtained from HADDOCK web server. The insets show the orientation of side chains of the interacting residues in stick representation (colored cyan). **(D)** Schematic representation of the *ybaB/ebfC* gene locus with adjacent genes in various bacteria. This arrangement of genes appears to be conserved in most bacteria that harbor YbaB/EbfC proteins. **(E)** YbaB homologs in alpha proteobacteria. Genomes displaying YbaB homologs with more than 40% sequence identity are labeled in red and proteins having identity below 40% are labeled in orange. Bacterial genomes with no YbaB homologs are labeled in yellow.

partial threads and gives the user five highly ranked models for a given sequence (Kim et al., 2004). For YbaB$_{Cc}$ protein sequence, we obtained five models that had a confidence score of 0.75. The confidence score is a qualitative indicator for the user to gauge the success of the modeling, where a confidence score of 0.7 and above for a model is considered as high-quality models that are

akin to those obtained by low-resolution X-ray crystallography or NMR experiments (Song et al., 2013). Thus, the confidence score of 0.75 for $YbaB_{Cc}$-modeled structure indicates a high-quality model and Model 1 was taken as the candidate structure for all further experimental analysis. The 3D structure of $YbaB_{Cc}$ was strikingly similar to crystalized YbaB/EbfC structures (Lim et al., 2002) showing characteristic tweezer-like conformation (**Figure 2B**). DNA-binding properties of Model 1 were probed further. The coordinates of Model 1 and DNA were given as input to HADDOCK and the resulting four top-ranked conformations showed that the protein has a preferential site of binding with the nucleic acid (**Figure 2C**). To explore further protein-DNA interactions were mapped using NUCPLOT, where for each conformation made both hydrogen and hydrophobic interactions were probed. $YbaB_{Cc}$ protein models displayed both hydrogen bonds and hydrophobic interactions with DNA (**Figure 2C**). Further, mapping the electrostatic surface of the modeled $YbaB_{Cc}$ protein showed a patch of positively charged surface (sequence range from Arg51 to Ile71, **Supplementary Figure 2**) that is most likely to interact with DNA. On mapping the electrostatic surface of the modeled $YbaB_{Cc}$ protein, a patch of positively charged region (sequence range from Arg51 to Ile71, **Supplementary Figure 1**) was found giving an idea about where the probable binding site is present. Further, the top scoring protein-DNA complex, obtained from HADDOCK web server, was given as input to the NUCPLOT program. Cut-offs of 3 and 3.35 Å were used for hydrogen bond and other non-bonded interactions, respectively. The interaction network constructed predicted 24 residues interact with the DNA forming both hydrophobic as well as hydrogen bonds. The network predicted a single hydrogen bond formed between Lys79 residue and the 31st nucleotide whereas other residues formed other kinds of non-bonded interactions. Multiple sequence alignment (MSA) of $YbaB_{Cc}$ was carried out using MultAlin software (Corpet, 1988), which generated the consensus residues (red color) also present in other YbaB/EbfC family proteins from *Borrelia burgdorferi* (EbfC), *E. coli* ($YbaB_{Ec}$), and *Deinococcus radiodurans* (DR_0199; **Figure 2A**). Gln-12 and Lys-80 residues are found to be crucial in DNA docking (**Figure 2C**) and are also common to all four YbaB/EbfC family proteins assessed by MSA (**Figure 2A**).

Localization of $YbaB_{Cc}$ Protein

To determine the cellular localization of the $YbaB_{Cc}$ in *Caulobacter* cells, $YbaB_{Cc}$ was fused with CFP at the N-terminal and expressed from *Pxyl* promoter. RP42 cells showed CFP signal throughout the cell (**Supplementary Figure 1**). In *C. crescentus* cells the chromosome spreads from pole to pole making it difficult to differentiate between cytoplasmic and NAPs colocalization with DNA (Arias-Cartin et al., 2017). To probe whether $YbaB_{Cc}$ protein colocalized with the nucleoid, the CFP-$YbaB_{Cc}$ fusion protein was expressed in a temperature sensitive *ftsI C. crescentus* mutant, which forms filaments at the restrictive temperature. In this mutant, $YbaB_{Cc}$ colocalized with the DAPI signal and was absent in few DNA-free regions (**Figure 3**), indicating that $YbaB_{Cc}$ is a nucleoid associated protein.

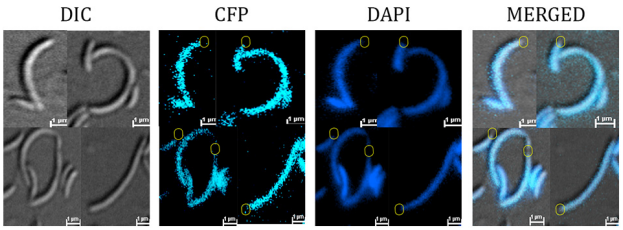


FIGURE 3 | $YbaB_{Cc}$ colocalizes with nucleoid *in vivo*. RP44 cells (CB15N *cc2570*:pTC67 *ftsI*(Ts) *xylX*:pXCFPN-5-$ybab_{Cc}$) were cultured at 30°C for 1 h and then shifted to the restrictive temperature (37°C) for 4.5 h in PYE medium, followed by DAPI staining and imaging on agarose-padded glass slides. Fluorescence imaging shows $YbaB_{Cc}$-CFP colocalizing with DAPI-stained nucleoid in a temperature-sensitive PBP3 mutant *Caulobacter crescentus*. DNA-free regions devoid of both DAPI and CFP signal are depicted in circles (yellow).

$YbaB_{Cc}$ Interaction With DNA Is Not Sequence-Dependent

In order to probe the DNA-binding activity of *C. crescentus* YbaB, we purified $YbaB_{Cc}$ protein, tagged with 6X His at the N-terminal. DNA binding of $YbaB_{Cc}$ was first tested using biotin labeled DNA probe corresponding to operator sequences of *erpAB* in *B. burgdorferi* (b-WT; Riley et al., 2009). This DNA sequence was selected as EbfC protein from *B. burgdorferi* preferentially binds to sequences within this region. Oligonucleotide b-WT (124 bp) from *Borrelia burgdorferi* strain B31 (Riley et al., 2009) was incubated with increasing concentrations (0, 4.7, 14, 23, 32, 42, 45, 47, and 94 μM) of $YbaB_{Cc}$ protein for electrophoretic mobility shift assay. $YbaB_{Cc}$ protein bound to the duplex DNA probes and formed stable protein-DNA complexes (**Figure 4A**). EbfC preferentially binds to GTnAC sequence found in *erpAB* operator 2 region (Riley et al., 2009). To investigate the specificity of $YbaB_{Cc}$ binding, we performed EMSA with increasing concentrations of poly (dI-dC) probe as a competitor for non-specific DNA-binding activities. As evident from **Figure 4B**, concentrations above 2.5 μg of poly (dI-dC) probes (∼60-fold higher or above) were able to abolish $YbaB_{Cc}$-DNA complexes, indicating the $YbaB_{Cc}$ in *C. crescentus* might be a non-specific DNA-binding protein. To confirm non-specific DNA-binding activity of $YbaB_{Cc}$, a 200-bp DNA sequence (cc-WT) of $ybab_{Cc}$ gene from *C. crescentus* was amplified and this biotin-labeled DNA duplex was used as probe in EMSA. Even very low concentrations (14 μM) of $YbaB_{Cc}$ protein caused super shift of cc-WT sequence probe (**Figure 4C**). Taken together, our data suggests that $YbaB_{Cc}$ interaction with DNA is not sequence dependent.

B. burgdorferi EbfC variants carrying mutations in α-helix 1 or 3 are defective in DNA binding (Riley et al., 2009). *C. crescentus* $YbaB_{Cc}$ protein, DNA-binding domains are present at both the N- and C-terminus. To investigate if both DNA-binding domains are essential for $YbaB_{Cc}$ DNA-binding activities, we created a truncated version—$YbaB_{Cc}$ (Tn), deleted for C-terminal DNA-binding domain and consisting of 1–200 bp of $ybab_{Cc}$ gene. $YbaB_{Cc}$ (Tn) protein was purified and EMSA was performed

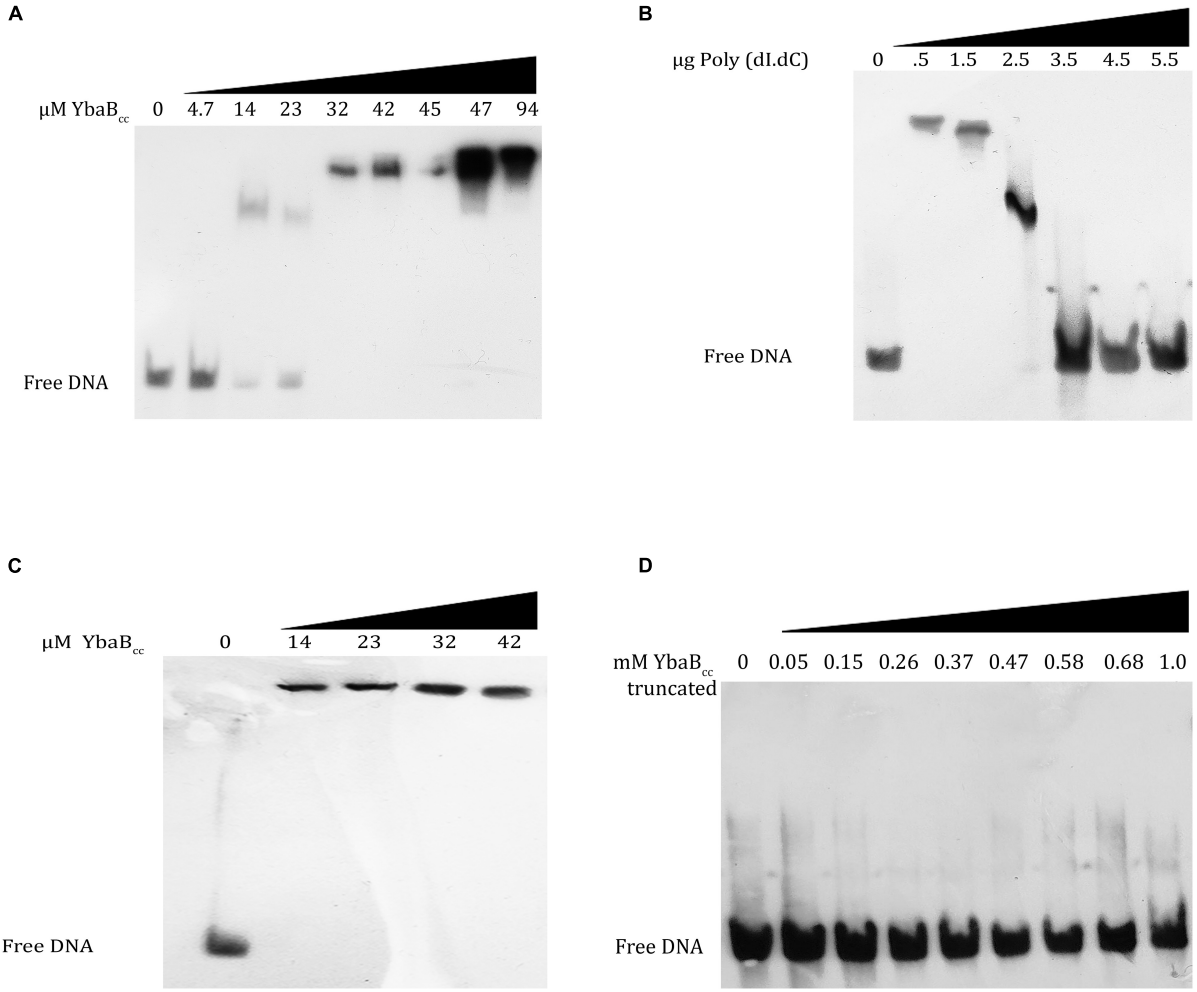


FIGURE 4 | YbaB$_{Cc}$ binds to DNA in a sequence-independent manner *in vitro*. YbaB$_{Cc}$ was incubated with 8 nM each of biotin-labeled probes. **(A)** b-WT (124 bp) from *Borrelia burgdorferi* strain B31 was incubated with increasing concentrations (0, 4.7, 14, 23, 32, 42, 45, 47, and 94 μM) of YbaB$_{Cc}$ protein. **(C)** cc-WT (200 bp) *Caulobacter* probe was incubated with increasing concentrations (0, 14, 23, 32, and 42 μM) of YbaB$_{Cc}$ protein. Reactions were electrophoresed on 6% DNA retardation gels and visualized by chemiluminescent detection. YbaB$_{Cc}$ forms DNA–protein complexes with both probes **(A–C)** causing shifts of protein–DNA complexes. **(B)** Competitive binding was also studied by adding varied concentrations (0, 0.5, 1.5, 2.5, 3.5, 4.5, and 5.5 μg) of Poly (dI.dC) to YbaB$_{Cc}$ (32 μM). The presence of free DNA with increasing Poly (dI.dC) indicates sequence independent DNA-binding activity of YbaB$_{Cc}$ by exchange of b-WT (124 bp) from *Borrelia burgdorferi* strainB31 probe with Poly (dI.dC). **(D)** YbaB$_{Cc}$ (Tn) protein (containing only N-terminal DNA-binding domain) was also tested for its ability to bind b-WT *Borrelia* probe. However, there was no shift observed even at very high protein concentrations rendering this truncated version of YbaB$_{Cc}$ to be ineffective in DNA-binding suggesting that both DNA-binding domains are essential for YbaB$_{Cc}$ protein to bind target DNA.

using b-WT *Borrelia* oligonucleotide. YbaB$_{Cc}$ (Tn) protein was unable to form stable protein-DNA complexes and no shift was observed (**Figure 4D**). These results suggest that C-terminal DNA-binding domain of YbaB$_{Cc}$ is important for DNA-binding activity.

YbaB$_{Cc}$ Protects DNA From Degradation

Most NAPs in bacteria are able to protect DNA from degradation. We investigated YbaB$_{Cc}$ DNA protection activity by performing DNase I enzymatic degradation assay. Supercoiled pBAD18 plasmid was incubated with increasing YbaB$_{Cc}$ protein concentrations and then treated with DNaseI enzyme (**Figure 5**). As seen in **Figure 5**, supercoiled plasmid DNA was protected from enzymatic degradation in presence of YbaB$_{Cc}$. In contrast, the control sample incubated with BSA was degraded upon treatment with DNaseI (**Figure 5**). Our results indicate that YbaB$_{Cc}$ protein may protect DNA against damage and degradation.

Overexpression of YbaB$_{Cc}$ in Nutrient Limiting Conditions Leads to Morphological Aberrations

Abundance of NAPs is highly regulated and fluctuations in protein levels cause adverse effects on cells (Ali Azam et al., 1999). Constitutive overexpression of GapR at high levels is lethal for *C. crescentus* (Ricci et al., 2016). To investigate

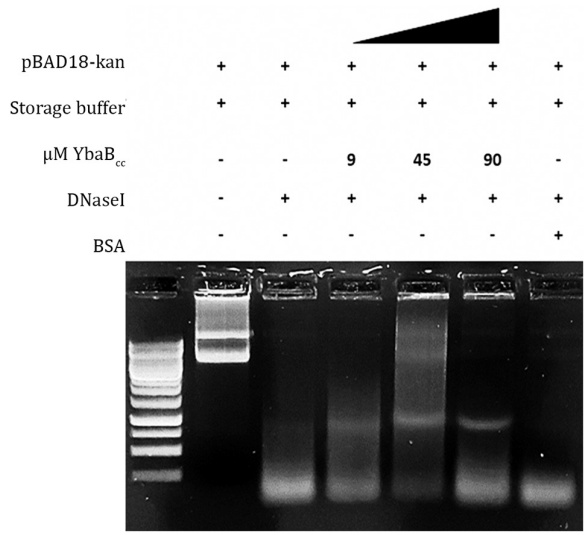


FIGURE 5 | YbaB$_{Cc}$ protects DNA from enzyme degradation. Different concentrations of YbaB$_{Cc}$ protein were incubated with supercoiled pBAD18 plasmid DNA, followed by treatment with DNase I restriction enzyme. Control reactions were incubated with BSA in place of YbaB$_{Cc}$ protein. YbaB$_{Cc}$ prevents complete plasmid DNA degradation by the enzyme as opposed to the BSA control reaction, where DNase I is able to degrade the plasmid completely. Thus indicating that YbaB$_{Cc}$ protein-bound plasmid DNA was protected by the degradative action of DNase I restriction enzyme and BSA could not protect the plasmid DNA from being degraded as it does not have DNA-binding activity.

whether YbaB$_{Cc}$ overexpression leads to abnormalities, we expressed YbaB$_{Cc}$ from pJS14 plasmid under the control of xylose promoter. Over expression of YbaB$_{Cc}$ caused morphological defects in *Caulobacter* cells grown in nutrient limiting conditions. Specifically, we observed filamentous cells and conjoined daughter cells with cell separation defects (**Figure 6A**). Cell length was increased by 44% compared to control cell grown in glucose (**Figures 6A,B**). It should be noted that overexpression of YbaB$_{Cc}$ in cells growing in nutrient rich PYE medium displayed no growth and morphological defects (data not shown). Collectively, our data suggests that YbaB$_{Cc}$ concentration in regulated in *C. crescentus* and probably plays a critical role in stress response.

DISCUSSION

Bacterial chromosomes are condensed and stabilized by the action of NAPs. NAPs are also involved in global gene expression by making alterations in DNA structure or by interactions with transcriptional machinery. In this study, we have characterized the DNA-binding properties of YbaB$_{Cc}$ protein in *C. crescentus*. YbaB$_{Cc}$ belongs to EbfC/YbaB class of DNA-binding proteins, first discovered in *B. burgdorferi* (Babb et al., 2006). EbfC homodimers have unique tweezer-like structure, where extending α-helices form the arms of the "tweezer" that bind DNA (Lim et al., 2002; Riley et al., 2009). *B. burgdorferi* EbfC protein is a sequence specific NAP and acts as a global regulator of gene expression. However, YbaB$_{Ec}$ and YbaB$_{Hi}$ do not preferentially bind *B. burgdorferi* palindromic DNA sequence (Cooley et al., 2009) EbfC homolog from *D. radiodurans* also acts as a non-specific DNA-binding protein (Wang et al., 2012). Our results indicate that YbaB$_{Cc}$ might act as a non-specific DNA-binding protein. Differences in DNA binding are attributed to amino-acid differences in the putative DNA-binding domain (Cooley et al., 2009). *C. crescentus* YbaB$_{Cc}$ exhibits characteristic tweezer-like conformation of EbfC/YbaB family of NAPs and probably functions as a homodimer. Both N and C-terminals of YbaB$_{Cc}$ protein have DNA-binding domains and deletion of the C-terminus DNA-binding domain is sufficient to abolish DNA-duplex binding activity (**Figure 4D**).

Similar to *B. burgdorferi* and *D. radiodurans*, YbaB$_{Cc}$ is associated with the nucleoid in *C. crescentus* (Jutras et al., 2012; Wang et al., 2012; **Figure 3**). Ectopic expression of YbaB$_{Cc}$ in *E. coli* condensed the nucleoid suggesting involvement of YbaB$_{Cc}$ in nucleoid organization and DNA structuring in cells. Collectively our data indicates that YbaB$_{Cc}$ is a homolog of EbfC/YbaB and has histone-like activity in *C. crescentus*.

NAPs are involved in global gene expression by making alterations in DNA structure or by interactions with transcriptional machinery. HU is one of the most abundant NAP in bacteria involved in nucleoid structuring (Shindo et al., 1992; Ghosh and Grove, 2004; Nguyen et al., 2009; Oberto et al., 2009; Berger et al., 2010). Surprisingly, *C. crescentus* hu deletion mutants deletion mutants display no adverse effect on cell growth, fitness and chromosome architecture (Christen et al., 2011; Lee et al., 2011). Moreover, loss of other NAPs such as IHF and DPS has no fitness cost on *C. crescentus* indicating functional redundancy among NAPs (Christen et al., 2011). Thus, it is not surprising that YbaB$_{Cc}$ is not essential for growth and viability of *C. crescentus* under standard laboratory conditions (Christen et al., 2011). While *ebfC* is an essential gene in *B. burgdorferi* (Riley et al., 2009; Jutras et al., 2012), deletion of *efbC* homolog from *D. radiodurans* does not affect cell viability (Wang et al., 2012). However, *D. radiodurans dr0199* mutants have increased sensitivity to UV radiation and oxidative stress (Wang et al., 2012).

Role of NAPs in DNA protection is well established. HU from *Thermotoga maritima* and *Helicobactor pylori* shield DNA from endolytic cleavage by DNase I and hydroxyl radical-mediated damage (Mukherjee et al., 2008; Almarza et al., 2015). Lsr2 from *Mycobacterium tuberculosis* acts against reactive oxygen intermediates by directly binding to DNA and shielding the DNA from damage (Colangeli et al., 2009). EbfC homolog from *D. radiodurans* protects double-stranded DNA from digestion and damage from reactive oxygen species (Wang et al., 2012). Similarly, our results reveal that YbaB$_{Cc}$ protects double-stranded DNA from enzymatic degradation, probably by direct protein-DNA binding (**Figure 5**). YbaB$_{Cc}$ may play a crucial role in maintaining genomic integrity by restricting DNA damage in *C. crescentus*.

Most sequenced eubacterial genomes contain adjacent *dnaX* and *ebfC* genes (Flower and Mchenry, 1990; Flower and McHenry, 1991; Chen et al., 1993). In *B. burgdorferi* and *E. coli*, *dnaX* and *ebfC* are arranged in an operon and are co-transcribed

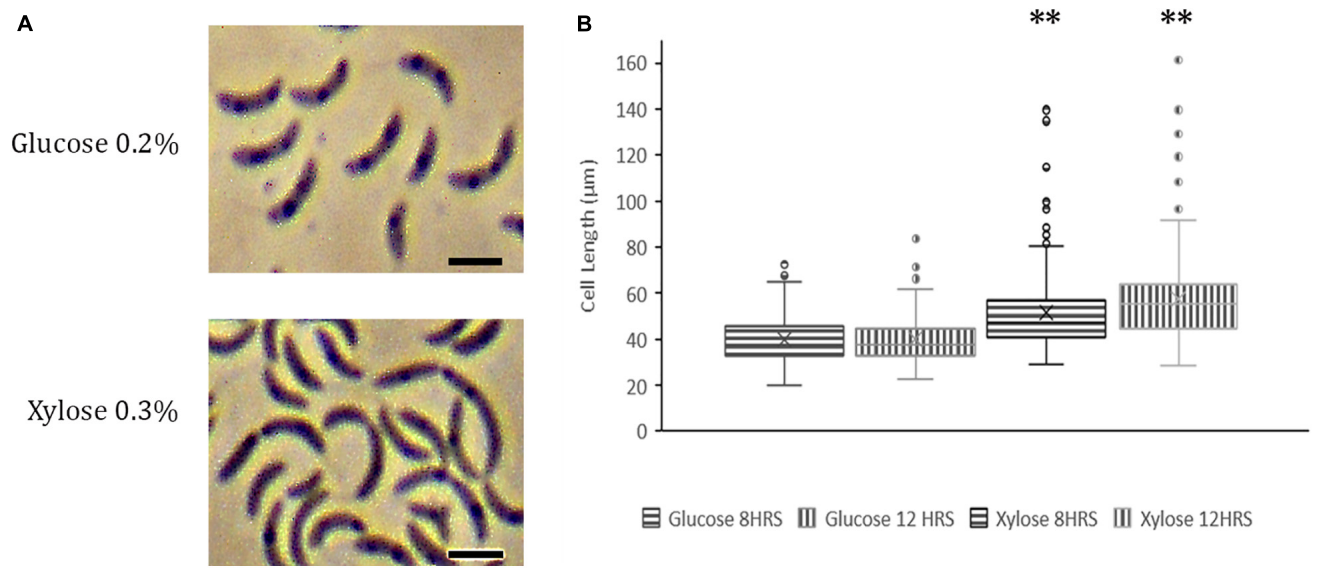


FIGURE 6 | YbaB$_{Cc}$ overexpression in wild-type *C. crescentus*. **(A)** Overexpression of YbaB$_{cc}$ leads to morphological defects in *C. crescentus*. Expression of YbAB$_{Cc}$ from plasmid pJS14 was induced by the addition of 0.3% xylose at an OD of 0.2 and cells were visualized at the indicated time points by Phase Contrast Microscopy (left panel). All images were taken at 100X magnification and optical zoom 1.5X (Scale bar: 2 μm). **(B)** Distribution of cell length in populations of YbaB$_{Cc}$ overexpressing cells. Micrographs of the strains described in **(A)** were subjected to automated image analysis. Cell lengths were determined by Oufti Software. Box and whisker plots ($n = \pm 210$ per time point and strain) were generated for statistical analysis of imaging data with the whiskers including all data points within 1.5*IQR. Cross (x) represents mean values. $^{**}p < 0.0001$ by one-way ANOVA compared to the glucose-induced control cells.

together (Riley et al., 2009). *C. crescentus* genome also has *dnaX* gene adjacent to YbaB$_{Cc}$ (**Figure 2D**). The *dnaX* gene encodes the tau and gamma subunits of DNA polymerase and YbaB is a NAP with ability to compact and protect DNA (Flower and McHenry, 1991; Chen et al., 1993). Transcriptional linkage of *dnaX* and *ybaB/ebfC* may allow bacterial cells to respond to DNA damage in a timely and regulated manner. Our preliminary experiments to detect transcriptional linkage between *dnaX* and YbaB$_{Cc}$ did not yield positive results (data not shown). This could be due to low levels of transcripts, as the expression of these genes is growth phase-dependent, or *dnaX* and YbaB$_{Cc}$ maybe transcribed independently in *C. crescentus*. EbfC levels are highest in exponential phase and rapidly decline in stationary phase in *B. burgdorferi* (Jutras et al., 2012). RecR protein plays crucial role in DNA recombination and repair is also located on the same locus (**Figure 2D**; Mahdi and Lloyd, 1989; Yeung' et al., 1990). Similar to YbaB$_{Cc}$, RecR too is expendable for growth under normal laboratory conditions in *C. crescentus* (Christen et al., 2011). It is interesting to speculate that both YbaB and RecR may play a role in DNA repair and protection during stress conditions. YbaB homologs are present in most alpha proteobacteria, suggesting an important physiological role of this NAP. Further studies are necessary to elucidate the physiological role of YbaB$_{Cc}$ in *C. crescentus*.

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/s.

AUTHOR CONTRIBUTIONS

RP and PP designed and conceptualized the study. PP and MM performed the experiments. SR and RY performed the bioinformatics analysis. All authors contributed toward the writing of the manuscript.

FUNDING

PP was supported by UGC-NET-JRF and MM was supported by Inspire-Doctoral fellowship. RP lab was supported by CSIR-EMR grant and start-up funds from Shiv Nadar University. Authors also thank SNU for the use of DST-FIST funded Confocal Imaging facility. SR and RY acknowledge SASTRA Deemed to be University for the infrastructural support and research facilities.

ACKNOWLEDGMENTS

We thank Dr. Karthik Krishnan and the members of the RP lab for their helpful comments and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.733344/full#supplementary-material

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.

Almarza, O., Nunez, D., and Toledo, H. (2015). The DNA-binding protein HU has a regulatory role in the acid stress response mechanism in *Helicobacter pylori*. *Helicobacter* 20, 29–40. doi: 10.1111/hel.12171

Arias-Cartin, R., Dobihal, G. S., Campos, M., Surovtsev, I. V., Parry, B., and Jacobs-Wagner, C. (2017). Replication fork passage drives asymmetric dynamics of a critical nucleoid-associated protein in *Caulobacter*. *EMBO J.* 36, 301–318. doi: 10.15252/embj.201695513

Babb, K., Bykowski, T., Riley, S. P., Miller, M. C., Demoll, E., and Stevenson, B. (2006). *Borrelia burgdorferi* EbfC, a novel, chromosomally encoded protein, binds specific DNA sequences adjacent to erp loci on the spirochete's resident cp32 prophages. *J. Bacteriol.* 188, 4331–4339. doi: 10.1128/JB.00005-06

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

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

Bhat, A. I., and Rao, G. P. (2020). "Dot-blot hybridization technique," in *Characterization of Plant Viruses: Methods and Protocols*, eds A. I. Bhat, and G. P. Rao (New York, NY: Springer US), 303–321. doi: 10.1007/978-1-0716-0334-5_34

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Chen, K.-S., Saxena, P., and Walker, J. R. (1993). Expression of the *Escherichia coli* dnaX Gene. *J. Bacteriol.* 175, 6663–6670. doi: 10.1128/jb.175.20.6663-6670.1993

Christen, B., Abeliuk, E., Collier, J. M., Kalogeraki, V. S., Passarelli, B., Coller, J. A., et al. (2011). The essential genome of a bacterium. *Mol. Syst. Biol.* 7:528. doi: 10.1038/msb.2011.58

Colangeli, R., Haq, A., Arcus, V. L., Summers, E., Magliozzo, R. S., McBride, A., et al. (2009). The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4414–4418. doi: 10.1073/pnas.0810126106

Cooley, A. E., Riley, S. P., Kral, K., Miller, M. C., DeMoll, E., Fried, M. G., et al. (2009). DNA-binding by *Haemophilus influenzae* and *Escherichia coli* YbaB, members of a widely-distributed bacterial protein family. *BMC Microbiol.* 9:137. doi: 10.1186/1471-2180-9-137

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881–10890. doi: 10.1093/nar/16.22.10881

Costa, T., Priyadarshini, R., and Jacobs-Wagner, C. (2008). Localization of PBP3 in *Caulobacter crescentus* is highly dynamic and largely relies on its functional transpeptidase domain. *Mol. Microbiol.* 70, 634–651. doi: 10.1111/j.1365-2958.2008.06432.x

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

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

Dorman, C. J. (2009). Nucleoid-associated proteins and bacterial physiology. *Adv. Appl. Microbiol.* 67, 47–64. doi: 10.1016/S0065-2164(08)01002-2

Dorman, C. J. (2014). Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. *J. Mol. Microbiol. Biotechnol.* 24, 316–331. doi: 10.1159/000368850

Dubey, A., and Priyadarshini, R. (2018). Amidase activity is essential for medial localization of AmiC in *Caulobacter crescentus*. *Curr. Genet.* 64, 661–675. doi: 10.1007/s00294-017-0781-9

Evinger, M., and Agabian, N. (1977). Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. *J. Bacteriol.* 132, 294–301. doi: 10.1128/jb.132.1.294-301.1977

Flower, A. M., and Mchenry, C. S. (1990). The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting (DNA replication/dnaX/dnaZ/translation/LacZ fusions). *Proc. Natl. Acad. Sci. U.S.A.* 87, 3713–3717.

Flower, A. M., and McHenry, C. S. (1991). Transcriptional organization of the *Escherichia coli* dnaX gene. *J. Mol. Biol.* 220, 649–658. doi: 10.1016/0022-2836(91)90107-H

Ghosh, S., and Grove, A. (2004). Histone-like protein HU from *Deinococcus radiodurans* binds preferentially to four-way DNA junctions. *J. Mol. Biol.* 337, 561–571. doi: 10.1016/j.jmb.2004.02.010

Ghosh, S., Indi, S. S., and Nagaraja, V. (2013). Regulation of lipid biosynthesis, sliding motility, and biofilm formation by a membrane-anchored nucleoid-associated protein of *Mycobacterium tuberculosis*. *J. Bacteriol.* 195, 1769–1778. doi: 10.1128/JB.02081-12

Gondil, V. S., Dube, T., Panda, J. J., Yennamalli, R. M., Harjai, K., and Chhibber, S. (2020). Comprehensive evaluation of chitosan nanoparticle based phage lysin delivery system; a novel approach to counter *S. pneumoniae* infections. *Int. J. Pharm.* 573:118850. doi: 10.1016/j.ijpharm.2019.118850

Guzman, L.-M., Belin, D., Carson, M. J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* 177, 4121–4130. doi: 10.1128/jb.177.14.4121-4130.1995

Heté Nyi, C., and van der Spoel, D. (2002). Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Sci.* 11, 1729–1737. doi: 10.1110/ps.0202302

Jutras, B. L., Bowman, A., Brissette, C. A., Adams, C. A., Verma, A., Chenail, A. M., et al. (2012). EbfC (YbaB) is a new type of bacterial nucleoid-associated protein and a global regulator of gene expression in the Lyme disease spirochete. *J. Bacteriol.* 194, 3395–3406. doi: 10.1128/JB.00252-12

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

Kim, D. E., Chivian, D., and Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 32, W526–W531. doi: 10.1093/nar/gkh468

Le, T. B., and Laub, M. T. (2016). Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. *EMBO J.* 35, 1582–1595. doi: 10.15252/embj.201593561

Le, T. B. K., Imakaev, M. V., Mirny, L. A., and Laub, M. T. (2013). High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* 342, 731–734. doi: 10.1126/science.1242059

Lee, S. F., Thompson, M. A., Schwartz, M. A., Shapiro, L., and Moerner, W. E. (2011). Super-resolution imaging of the nucleoid-associated protein HU in *Caulobacter crescentus*. *Biophys. J.* 100, L31–L33. doi: 10.1016/j.bpj.2011.02.022

Lim, K., Tempczyk, A., Parsons, J. F., Bonander, N., Toedt, J., Kelman, Z., et al. (2002). Structure note crystal structure of YbaB from *Haemophilus influenzae* (HI0442), a protein of unknown function coexpressed with the recombinational DNA repair protein RecR. *Proteins* 50, 375–379.

Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (1997). NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions. *Nucleic Acids Res.* 25, 4940–4945. doi: 10.1093/nar/25.24.4940

Mahdi, A. A., and Lloyd, R. G. (1989). The recR locus of *Escherichia coli* K-12: molecular cloning, DNA sequencing and identification of the gene product: thI eRlcso shrci oiK1:mlclr lnn,DAsqecn n dniiaino h. *Nucleic Acids Res.* 17, 6781–9794.

Mukherjee, A., Sokunbi, A. O., and Grove, A. (2008). DNA protection by histone-like protein HU from the hyperthermophilic eubacterium *Thermotoga maritima*. *Nucleic Acids Res.* 36, 3956–3968. doi: 10.1093/nar/gkn348

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

Oberto, J., Nabti, S., Jooste, V., Mignot, H., and Rouviere-Yaniv, J. (2009). The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS One* 4:e4367. doi: 10.1371/journal.pone.0004367

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

Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., et al. (2015). Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol. Microbiol.* 99, 767–777. doi: 10.1111/mmi.13264

Ricci, D. P., Melfi, M. D., Lasker, K., Dill, D. L., McAdams, H. H., and Shapiro, L. (2016). Cell cycle progression in *Caulobacter* requires a nucleoid-associated protein with high AT sequence recognition. *Proc. Natl. Acad. Sci. U.S.A.* 113, E5952–E5961. doi: 10.1073/pnas.1612579113

Riley, S. P., Bykowski, T., Cooley, A. E., Burns, L. H., Babb, K., Brissette, C. A., et al. (2009). *Borrelia burgdorferi* EbfC defines a newly-identified, widespread family of bacterial DNA-binding proteins. *Nucleic Acids Res.* 37, 1973–1983. doi: 10.1093/nar/gkp027

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Shindo, H., Furubayashi, A., Shimizu, M., Miyake, M., and Imamoto, F. (1992). Preferential binding of *E.coli* histone-like protein HU alpha to negatively supercoiled DNA. *Nucleic Acids Res.* 20, 1553–1558. doi: 10.1093/nar/20.7.1553

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.* 1, 784–791. doi: 10.1038/nbt1183-784

Song, Y., DiMaio, F., Wang, R. Y.-R., Kim, D., Miles, C., Brunette, T., et al. (2013). High-resolution comparative modeling with RosettaCM. *Structure* 21, 1735–1742. doi: 10.1016/j.str.2013.08.005

Stove Poindexter', J. (1964). Biological properties and classification of the *Caulobacter* group. *Bacteriol. Rev.* 28, 231–295.

Taylor, J. A., Panis, G., Viollier, P. H., and Marczynski, G. T. (2017). A novel nucleoid-associated protein coordinates chromosome replication and chromosome partition. *Nucleic Acids Res.* 45, 8916–8929. doi: 10.1093/nar/gkx596

Teotia, D., Gaid, M., Saini, S. S., Verma, A., Yennamalli, R. M., Khare, S. P., et al. (2019). Cinnamate-CoA ligase is involved in biosynthesis of benzoate-derived biphenyl phytoalexin in Malus 3 domestica "Golden Delicious" cell cultures. *Plant J.* 100, 1176–1192. doi: 10.1111/tpj.14506

Thanbichler, M., Iniesta, A. A., and Shapiro, L. (2007). A comprehensive set of plasmids for vanillate-and xylose-inducible gene expression in *Caulobacter crescentus*. *Nucleic Acids Res.* 35:e137. doi: 10.1093/nar/gkm818

van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., et al. (2016). The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* 428, 720–725. doi: 10.1016/j.jmb.2015.09.014

Wang, H., Wang, F., Hua, X., Ma, T., Chen, J., Xu, X., et al. (2012). Genetic and biochemical characteristics of the histone-like protein DR0199 in *Deinococcus radiodurans*. *Microbiology* 158(Pt 4), 936–943. doi: 10.1099/mic.0.053702-0

Yeung', T., Mullin, D. A., Chen, K.-S., Craig, E. A., Bardwell, J. C. A., and Walker', J. R. (1990). Sequence and expression of the *Escherichia coli* recR locus. *J. Bacteriol.* 172, 6042–6047.

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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