Genetic and biochemical interactions between the DNA replication initiator and a chromosome architecture protein in *Bacillus subtilis*

by

Charlotte A. Seid

A.B., Chemical and Physical Biology
Harvard College, 2010

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology
at the Massachusetts Institute of Technology

September 2014

© 2014 Massachusetts Institute of Technology
All rights reserved

Signature of author: 

Charlotte A. Seid
Department of Biology
September 2, 2014

Certified by: 

Alan D. Grossman
Professor of Biology
Thesis Supervisor

Accepted by: 

Amy E. Keating
Associate Professor of Biology
Co-Chair, Biology Graduate Committee
Genetic and biochemical interactions between the DNA replication initiator and a chromosome architecture protein in *Bacillus subtilis*

By

Charlotte A. Seid

Submitted to the Department of Biology on September 2, 2014
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology
at the Massachusetts Institute of Technology

ABSTRACT

I described the co-association, *in vitro* interaction, and epistatic relationship between the DNA replication initiator DnaA and the nucleoid-associated protein Rok in *Bacillus subtilis*. Using ChIP-seq, I characterized the genome-wide association of DnaA and its regulator YabA. I found that DnaA and YabA associated with over 30 chromosomal regions that were bound by Rok, and association of DnaA and YabA with these regions was dependent on Rok. The DNA-binding domain of DnaA was dispensable for association of DnaA with these chromosomal regions. This indirect Rok-dependent association contrasted with the canonical sequence-specific binding of DnaA to eight chromosomal regions containing clusters of DnaA boxes. I found that association of YabA with Rok-dependent regions did not require DnaA, in contrast to the DnaA-dependent association of YabA with DnaA box cluster regions. Furthermore, I showed an *in vitro* interaction between DnaA and Rok using purified proteins in an electrophoretic mobility shift assay. DnaA depended on Rok for association with a DNA probe, recapitulating the dependence observed at this chromosomal region *in vivo*. DnaA and Rok are both transcription factors and regulate some of the same genes. I analyzed global gene expression to determine the genetic relationship between these two transcriptional regulators. In general, *rok* was epistatic to *dnaA*; that is, the gene expression effects of DnaA depended on Rok, consistent with the ChIP-seq dependence.

I investigated a potential role for Rok in regulating replication initiation, but I did not detect a replication phenotype of a *rok* null mutant under various conditions. I found that a *rok* null mutation did not detectably affect DnaA or YabA protein levels, and I showed that the stationary phase growth defect of this strain was dependent on *comK*, a downstream target of Rok. Additionally, I determined that a putative regulator of replication initiation, the pyruvate dehydrogenase enzyme PdhC, affected replication via its metabolic function but was not a direct regulator of replication initiation.

Thesis Supervisor: Alan D. Grossman
Title: Professor of Biology
Acknowledgments

I am extremely grateful for the scientific advising, support, and mentorship of Alan and the members of the Grossman lab. Alan is not only an incisive thinker but a dedicated mentor with a wise sense of perspective in science and in life. He is unpretentious, straightforward, and generous in sharing his wealth of expertise, especially his keen ability to prioritize experiments and focus on the big picture. As highly as Alan values microbiology research, his thoughtful approach to solving problems and navigating life is valuable in any discipline. Not only scientists but people from many walks of life could benefit from a thoughtful, laughter-filled conversation in his office.

Many past and present members of the Grossman lab contributed ideas, expertise, data, and research tools to my thesis work. Janet Smith performed the in vitro experiments that motivated my study of DnaA and Rok, and I am extremely grateful for her biochemistry expertise, custom scripts for ChIP-seq analysis, comprehensive lab notes, and clever and compassionate presence as my bench neighbor. C. Lee performed intricate feats of cloning and protocol optimization in support of this project and many others, and she enthusiastically taught me many fundamentals and tricks of Bacillus genetics. Houra Merrikh mentored me during my rotation project and performed the preliminary work on YabA and Rok with inspirational intensity, drive, and efficiency. Tracy Washington produced the strains and gene expression data that facilitated my work with DnaA and Rok, and I appreciated his help with microarrays and his perspective as a senior graduate student. I am also very grateful for the friendship, camaraderie, and support of fellow students Moni Avello, Tyler DeWitt, Eleina England, Luciane Schons da Fonseca, and Laurel Wright.

My thesis committee members have provided helpful perspectives, suggestions, and advice over the past years. Steve Bell brought great clarity of thought not only to my committee meetings but to our service on the MIT Graduate Student Housing Group. Mike Laub hosted me for an exceptionally productive rotation project, and he has the unusual honor of also having reviewed my undergraduate thesis. Ann Hochschild was kind enough to meet with me and discuss my project even many months in advance of my thesis defense. Jonathan King and Frank Solomon provided valuable career, personal, and moral advice, and I am very grateful for their informal mentorship. Stuart Levine and Shmulik Motola of the BioMicroCenter contributed essential sequencing services and expertise. Many members of the Bell, Baker, and Laub labs generously shared equipment and fun company as our lab neighbors. Lori Ling joined me for an intense and mutually encouraging two weeks of thesis writing bootcamp. Betsey Walsh, Janice Chang, Darlene McGurl, and Maggie Cabral dedicated many generous efforts to keeping the labs and graduate program running smoothly. My former labmates Rick Zuzow, Mike Rust, Joe Markson, Tim Schmidt, and Shawn Low and my former advisors Erin O’Shea and Randy Schekman were also instrumental in my scientific journey.

Finally, I am incredibly appreciative of the encouragement, personal support, and wise advice of my family and friends. Completing this Ph.D. thesis was an honor, but is not who I am, and it was not the entirety of my four-year MIT experience. Megan Warner has been a wonderful friend and source of clear-headed advice since the beginning of our first year, and our friendship and therapeutic gym sessions have been a defining part of graduate school for me. I am also thankful for Michael Napolitano’s generosity and companionship and for the endlessly amusing, supportive community of friends at the Biomansion. My parents have given me incredible educational opportunities and inspirational experiences traveling the world. They encouraged my every interest in the natural world, from starting my own aquarium to diving to an undersea research station. Most importantly, they have modeled and encouraged a life of authenticity, introspection, and self-discovery, independent of any popular definitions of success. I wish the same, and the best of luck, for you.
# Table of Contents

Abstract .................................................................................................................. 2  
Acknowledgments ................................................................................................. 3  
Table of Contents .................................................................................................. 4  
List of Tables ......................................................................................................... 5  
List of Figures ....................................................................................................... 6  
Chapter 1  Introduction .......................................................................................... 9  
Chapter 2  Genetic and biochemical interactions between the bacterial replication initiator DnaA and the nucleoid-associated protein Rok in *Bacillus subtilis* ............................................................................................................. 59  
Chapter 3  Recruitment of the bacterial replication regulator YabA to several chromosomal regions by the nucleoid-associated protein Rok in *Bacillus subtilis* ................................................................................................................. 103  
Chapter 4  Conclusions and Perspectives ............................................................ 151  
Appendix A  Growth and replication phenotypes of a *rok* null mutant ............... 159  
Appendix B  Conservation of DnaA, YabA, and Rok ............................................ 181  
Appendix C  The pyruvate dehydrogenase complex subunit PdhC is not a direct regulator of replication initiation .......................................................................................... 187  
References ........................................................................................................... 203
## List of Tables

### Chapter 2
- **Table 1.** Chromosomal regions associated with DnaA .......................... 65
- **Table 2.** Gene regulation by DnaA and Rok ................................. 88
- **Table 3.** *B. subtilis* strains used in this study ............................... 93

### Chapter 3
- **Table 1.** Association of YabA with DnaA box cluster regions .......... 110
- **Table 2.** Rok-bound chromosomal regions associated with YabA ...... 111
- **Table 3.** *B. subtilis* strains used in this study ............................... 138

### Appendix A
- **Table 1.** *B. subtilis* strains used in this study ............................... 176

### Appendix B
- **Table 1.** YabA proteins of *Bacillus* species with large or small Rok proteins ................................................................. 183

### Appendix C
- **Table 1.** *B. subtilis* strains used in this study ............................... 199
## List of Figures

| Chapter 1 | Figure 1. | Domain structure of DnaA in *Bacillus subtilis* | 15 |
| Chapter 1 | Figure 2. | Origin of replication (*oriC*) in *Bacillus subtilis* | 17 |
| Chapter 1 | Figure 3. | Domain structure of Rok in *Bacillus subtilis* | 47 |
| Chapter 2 | Figure 1. | Genome-wide binding of DnaA and Rok in wild-type cells | 64 |
| Chapter 2 | Figure 2. | Association of DnaA and Rok with DnaA box cluster chromosomal regions in wild-type cells | 66 |
| Chapter 2 | Figure 3. | Association of DnaA and Rok with Rok-bound chromosomal regions in wild-type cells | 68 |
| Chapter 2 | Figure 4. | Sequence logo of a potential Rok binding motif | 70 |
| Chapter 2 | Figure 5. | Relative enrichment of DnaA and Rok at co-associated regions in wild-type cells | 71 |
| Chapter 2 | Figure 6. | Genome-wide binding of DnaA in wild-type cells and a *rok* null mutant | 73 |
| Chapter 2 | Figure 7. | Genome-wide binding of DnaA and Rok in *dnaA*+ and *dnaA* null mutant cells | 75 |
| Chapter 2 | Figure 8. | ChIP-seq reads for selected regions bound by DnaA and Rok in wild-type cells | 77 |
| Chapter 2 | Figure 9. | Mutually independent association of DnaA and Rok with the *dnaA/dnaN* and *trmE/jag* chromosomal regions | 79 |
| Chapter 2 | Figure 10. | Genome-wide binding of DnaA and Rok in cells expressing wild type *dnaA* or a DNA-binding mutant | 81 |
| Chapter 2 | Figure 11. | Relative DnaA protein levels in mutant strains | 82 |
| Chapter 2 | Figure 12. | ChIP-seq reads for selected regions bound by DnaA and DnaAΔC | 83 |
| Chapter 2 | Figure 13. | DnaAΔC requires Rok for association with DNA *in vitro* | 85 |
Figure 1. Genome-wide binding of YabA and Rok in wild-type cells
Figure 2. Association of YabA and DnaA with DnaA box cluster chromosomal regions in wild-type cells
Figure 3. Association of YabA and Rok with Rok-bound chromosomal regions in wild-type cells
Figure 4. Comparison of chromosomal regions bound by YabA, DnaA, and Rok
Figure 5. Genome-wide binding of YabA in wild-type cells and a *rok* null mutant
Figure 6. Genome-wide binding of YabA and Rok in a *dnaA* null mutant
Figure 7. YabA does not require DnaA for association with some chromosomal regions
Figure 8. Genome-wide binding of Rok and DnaA in *yabA*+ and *yabA* null mutant cells
Figure 9. Relative enrichment of YabA, DnaA, and Rok at co-associated regions in wild-type cells
Figure 10. ChIP-seq reads for selected regions bound by YabA, DnaA, and Rok in wild-type cells
Figure 11. Viable cell density in *rok* null and *yabA* null mutants
Figure 12. Relative YabA, DnaA, and Rok protein levels in wild-type and mutant strains
Figure 13. Stability of Rok and YabA proteins
Figure 14. Relative YabA, DnaA, and Rok protein levels in exponential and stationary phase
Figure 15. Genome-wide binding of YabA and DnaA in stationary phase in wild-type and *rok* null mutant cells
Figure 16. Genome-wide binding of Rok in exponential and stationary phase in wild-type cells
Figure 17. Specificity of YabA immunoprecipitations
Figure 18. Verification of YabA association with chromosomal regions
Figure 19. Verification of Rok-dependent YabA association with chromosomal regions

---

Chapter 4
Figure 1. Model for DnaA association with different types of chromosomal regions
| Appendix A | Figure 1. | No growth phenotype of a *rok* null mutant during exponential growth in minimal glucose medium | 162 |
| Figure 2. | No growth phenotype of a *rok* null mutant during exponential growth in minimal arabinose or LB medium | 163 |
| Figure 3. | Growth defect of a *rok* null mutant upon entry to stationary phase | 164 |
| Figure 4. | No growth defect of a *rok* null mutant during recovery from stationary phase or in LB medium | 165 |
| Figure 5. | No growth defect of a *rok* null mutant in a *comK* null background | 167 |
| Figure 6. | No replication phenotype of a *rok* null mutant during exponential growth | 168 |
| Figure 7. | No replication phenotype of a *rok* null mutant in a *comK* null background | 169 |
| Figure 8. | No replication phenotype of a *rok* null mutant under DnaN perturbations | 172 |
| Figure 9. | Response of DnaN protein levels to variable xylose | 173 |
| Figure 10. | No replication phenotype of a *rok* null mutant in a *yabA* null background | 175 |

| Appendix B | Figure 1. | Sequence alignment of *B. subtilis* Rok protein and a hypothetical protein from *B. cereus* | 184 |
| Figure 2. | Sequence alignment of DnaA proteins from *B. subtilis* and *B. anthracis* | 185 |
| Figure 3. | Potential shared protein sequence motifs in DnaA and Rok | 186 |

| Appendix C | Figure 1. | Map of genes in the *pdh* operon in *B. subtilis* | 189 |
| Figure 2. | Growth of a *pdhC* null mutant in media with citric acid cycle intermediates | 191 |
| Figure 3. | Effect of medium concentration on growth of a *pdhC* null mutant | 193 |
| Figure 4. | The replication phenotype of PDH complex disruption is not specific to *pdhC* | 195 |
| Figure 5. | Overexpression of *pdhC*-myc does not complement the growth defect of a *pdhC* null mutant | 196 |
| Figure 6. | Overexpression of *pdhAB* complements the growth defect of a *pdhB* null mutant | 198 |
Chapter 1

Introduction
Overview: DNA replication and its regulation

Replicating genetic information is an essential feature of life. As cells reproduce, it is important to coordinate replication of the genome with growth, division, and differentiation. Regulating replication allows cells to take advantage of favorable conditions by maximizing their growth rate, and it allows cells to adjust to environmental cues such as starvation or DNA damage. Defects in replication and its regulation can lead to DNA damage, genomic instability, disease, and inviable cells (Magdalou et al., 2014; O’Donnell et al., 2013).

DNA replication includes three steps: initiation, elongation, and termination (O’Donnell et al., 2013; White, 2007). For initiation, a specialized initiator protein recognizes a chromosomal sequence that serves as an origin of replication. Bacteria typically have a circular chromosome with a single origin, although some species have multiple or linear chromosomes (Egan et al., 2005; Hinnebusch and Tilly, 1993). Archaea have one or more origins per circular chromosome. In eukaryotes, each linear chromosome uses multiple origins. Helicases are loaded onto the origin, and they unwind the DNA to create bidirectional replication forks. The replication priming complex (primosome) and the complete replication machinery (replisome) are then assembled. Key components of the replisome in both prokaryotes and eukaryotes include primase, helicase, DNA polymerase, the sliding processivity clamp, the clamp loader, and single-strand binding protein, although the interactions among these components are different for different organisms. During elongation, the chromosome is replicated. According to the factory model of bacterial replication, the replisome components stay in a fixed location and DNA moves through the replisome (Lemon and Grossman, 1998). Termination involves the resolution of replication forks, resulting in two complete copies of the chromosome. In bacteria, the forks meet at a terminus region opposite the origin of replication on the chromosome.
Specific Ter sequences within the terminus region each allow only unidirectional replication, ensuring that the forks will meet within a specific, limited chromosomal region (O’Donnell et al., 2013; White, 2007).

The initiation step of replication is highly regulated and will be a major focus of this work. In some organisms, such as the human gut bacterium *Escherichia coli* and the soil bacterium *Bacillus subtilis*, replication can re-initiate before elongation of the previous round is complete, allowing parallel duplications of the chromosome during fast growth. This multifork replication mechanism allows cells to grow and divide in less time than it takes to replicate the entire genome. With multifork replication, each newly divided cell receives an actively replicating genome with multiple origins. In other organisms, or in the same organisms during slow growth, initiation occurs only after the previous round of replication has completed. The aquatic bacterium *Caulobacter crescentus* uses this mode of replication, perhaps because it grows in nutrient-poor conditions and may not be optimized for fast growth. Optimal regulation of initiation, then, is different for different organisms and can depend on environmental conditions (Marczynski, 1999; Skarstad and Katayama, 2013).

Replication initiation is mediated by initiator proteins that recognize an origin of replication. In bacteria, the initiator protein is DnaA, a AAA+ protein (ATPase associated with various cellular activities). DnaA binds to the chromosomal origin of replication oriC. In eukaryotes, the initiator is the six-subunit origin recognition complex (ORC). ORC binds to replication origins known as autonomously replicating sequences (ARS), which occur as multiple copies per chromosome. Of the ORC subunits, Orc1-5 have AAA+ or AAA+-like domains with distant homology to DnaA (Kawakami and Katayama, 2010; Scholefield et al., 2011). In archaea, the initiator is also a AAA+ protein that is homologous to both Orc1 and the
eukaryotic helicase loader protein Cdc6. A given archaeal species may contain several different orc1/cdc6 initiator genes, each with a cognate origin sequence. These initiator-origin pairs likely originated from extrachromosomal elements (Wu et al., 2014). Studying the function and regulation of DnaA, as a widely conserved replication protein, is therefore a valuable approach for uncovering both general principles and specific adaptations in the regulation of DNA replication.

The Gram-positive soil bacterium Bacillus subtilis is a tractable and informative model for studying the regulation of replication. B. subtilis is capable of undergoing different developmental states in response to environmental conditions. It can grow rapidly, with a doubling time of 20 minutes (Yoshikawa et al., 1964), and some strains can also form multicellular aggregates called biofilms (Earl et al., 2008). Under nutrient depletion, as B. subtilis enters stationary phase, it can respond in several possible ways: induction of motility and chemotaxis systems to seek more favorable conditions, secretion of degradative enzymes to scavenge nutrients, secretion of antibiotics to kill competitor cells, natural competence development to obtain potentially advantageous external DNA, or sporulation into a dormant differentiated state (Claverys and Håvarstein, 2007; Msadek, 1999; White, 2007). The complex developmental program of sporulation is considered a last resort in response to extreme starvation, conditions under which growth and DNA replication are not favorable. The DNA replication and sporulation pathways are therefore mutually inhibitory (White, 2007).

These adaptive and developmental pathways are beneficial for a benign, even edible (Cutting, 2011; Inatsu et al., 2006; Kiers et al., 2000) soil bacterium that experiences drastic environmental fluctuations, nutrient limitation, and competition with other organisms (Earl et al., 2008). In B. subtilis’ pathogenic relatives, these abilities to replicate under variable conditions,
form resistant spores, and uptake DNA for antibiotic resistance are clinical concerns. *B. subtilis* is thus a useful model system for understanding fundamental processes relevant to *B. anthracis*, *Clostridium*, *Listeria*, *Staphylococcus*, *Streptococcus*, and other medically important Gram-positive bacteria. Since these environmental responses and differentiated states have different requirements for DNA replication, *B. subtilis* is a promising system to study the coordination of DNA replication with other cellular processes and environmental cues.

**DnaA**

**Structure, properties, and conservation**

The replication initiator DnaA belongs to the AAA+ superfamily, a diverse group of ATPases that includes the ClpX protein unfoldase, the homologous recombination protein RuvB, the eukaryotic motor protein dynein, and other molecular remodeling proteins. Within the AAA+ group, DnaA belongs to the helicases and clamp loaders clade (Snider *et al.*, 2008). In *B. subtilis*, DnaA has 446 amino acids and a molecular weight of 50.9 kDa.

DnaA can bind and hydrolyze ATP. The ATP-bound form is the active form for replication (Sekimizu *et al.*, 1987). DnaA has equal affinity for both ATP and ADP (Sekimizu *et al.*, 1987), so the nucleotide-bound state is thought to reflect the intracellular ratio of ATP/ADP, which shows a ten-fold higher concentration of ATP (Bochner and Ames, 1982). Both DnaA-ATP and DnaA-ADP can bind oriC, but DnaA-ADP cannot initiate DNA duplex opening or primosome assembly (Sekimizu *et al.*, 1987). The role of DnaA-ATP in replication initiation is discussed in more detail below. In *E. coli*, the ratio of DnaA-ATP/DnaA-ADP peaks before replication initiation, so DnaA-ATP levels are thought to be an important determinant of the frequency of initiation (Kurokawa *et al.*, 1999).
DnaA contains four domains, defined by sequence homology and mutational analysis (Fig. 1). The N-terminus (Domain I) is used in oligomerization and cooperative DNA binding (Kawakami and Katayama, 2010; Leonard and Grimwade, 2010; Messer, 2002). In *E. coli*, this domain interacts with the replicative helicase; in *B. subtilis*, DnaA does not interact with the helicase directly (Leonard and Grimwade, 2010; Messer, 2002). The linker region (Domain II) is poorly conserved, variable in size across bacterial species, and not essential for DnaA function in *E. coli*. The AAA+ domain (III) is highly conserved (Leonard and Grimwade, 2011). This domain of DnaA not only shares distant sequence homology with ORC but has a similar structural conformation (Erzberger *et al.*, 2002). Within Domain III, the N-terminal section (Domain IIIa) contains a Walker A motif that binds ATP or ADP, and the C-terminal section (Domain IIIb) contains a Walker B motif that interacts with ATP and is used in ATP hydrolysis (Kawakami and Katayama, 2010; Messer, 2002). The AAA+ domain also interacts with single-stranded DNA as DnaA melts the DNA duplex for replication initiation. Each DnaA molecule uses two pairs of helices in the AAA+ region to contact three nucleotides via the phosphodiester backbone (Duderstadt *et al.*, 2011). The C-terminal domain (IV), also highly conserved, contains a helix-turn-helix motif and is necessary and sufficient for binding to double-stranded DNA (Krause *et al.*, 1997; Leonard and Grimwade, 2011; Roth and Messer, 1995).
Figure 1. Domain structure of DnaA in *Bacillus subtilis*.
Amino acid numbers are shown above the cartoon of conserved domains and their functions. Adapted from (Cho et al., 2008).

DnaA recognizes a sequence of nine base pairs known as a DnaA box. The same consensus sequence, 5’-TTATNCACA-3’, is used by *E. coli* and *B. subtilis*, a remarkable degree of conservation (Fuller et al., 1984; Messer, 2002; Moriya et al., 1988; Schaper and Messer, 1995). DnaA still recognizes sequences with mismatches from the consensus sequence, although with weaker interaction (Fujikawa *et al*., 2003; Leonard and Grimwade, 2011). Other DnaA proteins, such as those from *Helicobacter pylori* and *Streptomyces*, recognize the same or slightly deviated consensus sequences (Messer, 2002). Binding of DnaA introduces a 40-degree bend at each DnaA box (Schaper and Messer, 1995). The arginine finger of Domain IV interacts with the first three bases, 5’-TTA-3’, via hydrogen bonding in the minor groove. The helix-turn-helix motif of Domain IV recognizes the rest of the sequence by base-specific interactions with the major groove (Erzberger *et al*., 2002; Fujikawa *et al*., 2003). DnaA boxes are found at the origin of replication and elsewhere on the chromosome. The binding and multimerization of DnaA are important for DnaA function during DNA replication initiation and gene regulation.

DnaA autoregulates its own transcription via DnaA boxes at the *dnaA* promoter in *B. subtilis*, *E. coli*, and other organisms (Atlung *et al*., 1985; Messer and Weigel, 1997; Ogura *et al*., 2003).
2001; Zakrzewska-Czerwinska et al., 2007). In some organisms, such as *Thermus thermophilus*, *Pseudomonas putida*, and *Synechocystis*, dnaA is not autoregulated (Messer, 2002; Messer and Weigel, 1997). DnaA represses the dnaAN promoter in *B. subtilis* (Goranov et al., 2005; Ogura et al., 2001) and represses the dnaA promoter in *E. coli* (Atlung et al., 1985; Braun et al., 1985; Messer and Weigel, 1997). In *E. coli*, DnaA-ATP represses transcription more strongly than DnaA-ADP because additional binding sites are occupied (Speck et al., 1999).

DnaA is conserved in virtually all bacteria. The only species known to lack dnaA are obligate endosymbionts that live in the cytosol of insect cells: *Wigglesworthia glossinidia*, *Blochmannia floridanus*, *B. pennsylvanicus*, and *Baumannia cicadellinicola* (Zakrzewska-Czerwinska et al., 2007). dnaA is essential in nearly all bacteria that have the gene. One known exception is the cyanobacterium *Synechocystis*, which is presumed to have another cryptic origin-initiator pair (Richter et al., 1998a).

**Replication initiation by DnaA**

DnaA initiates replication by binding to the origin of replication, melting the DNA duplex, and recruiting components of the replication machinery. In *B. subtilis*, the chromosomal origin of replication, oriC, is a 2.2-kb region containing two clusters of DnaA boxes separated by the dnaA gene ([Fig. 2](#fig2)) (Kadoya et al., 2002; Moriya et al., 1992, 1994). The DNA-unwinding element, or DUE, is an AT-rich region of 27 bp, downstream of the second DnaA box cluster and followed by dnaN, encoding the beta (processivity) clamp of DNA polymerase III (Moriya et al., 1994). DUE regions are melted by DnaA after it binds to upstream DnaA boxes. The thermodynamic instability of the AT-rich region, rather than the specific sequence, is important for melting (Kowalski and Eddy, 1989; Speck and Messer, 2001).
Figure 2. Origin of replication (oriC) in Bacillus subtilis.
The left portion of oriC includes a cluster of DnaA boxes (red triangles) upstream of the dnaA gene. The right portion includes a second cluster of DnaA boxes and the A+T-rich DNA Unwinding Element (DUE). dnaA and dnaN are expressed from the dnaA promoter. Chromosome position is indicated in base pairs.

The relative positions of dnaA and oriC vary in different organisms, and the bipartite arrangement of oriC in B. subtilis appears unusual. In Pseudomonas putida, oriC is upstream of dnaA and dnaN but not divided by the dnaA gene as it is in B. subtilis. A one-part oriC is located between dnaA and dnaN in organisms including Thermus thermophilus, Streptomyces coelicolor, Spiroplasma citri, and Mycoplasma capricolum. dnaAN is located far upstream of oriC in Haemophilus influenzae and far downstream of oriC in E. coli and Vibrio harveyi. In Caulobacter crescentus, dnaA follows oriC, but both are far from dnaN. In Prochlorococcus marinus and Synechocystis, dnaA is separate from both dnaN and oriC (Messer, 2002; Ogasawara et al., 1985; Richter et al., 1998b). Different genomic arrangements could allow for different mechanisms that couple or uncouple replication from the copy number and expression of the dnaA and dnaN genes.

Replication initiation begins with DnaA binding to DnaA boxes at oriC. Studies from E. coli show that high-affinity boxes are bound first, followed by cooperative binding to lower-affinity boxes. Binding to lower affinity sites requires oligomerization (Leonard and Grimwade,
2011; Messer, 2002; Miller et al., 2009; Rozgaja et al., 2011; Speck and Messer, 2001). Similar observations have been made for *B. subtilis* DnaA complex assembly, and the ATP-bound form of DnaA binds more cooperatively than the ADP-bound form to oriC (Krause et al., 1997; Merrikh and Grossman, 2011). As multiple DnaA molecules bind to oriC, they form a nucleoprotein complex in the form of a right-handed helix around the DNA, and this oligomerization requires ATP binding (Crooke et al., 1993; Duderstadt et al., 2011; Erzberger et al., 2006; Fuller et al., 1984; Funnell et al., 1987; McGarry et al., 2004; Messer, 2002). The AAA+ domain is proposed to switch between a closed configuration in the ADP-bound form and an open configuration in the ATP-bound form. The open configuration allows interaction between the arginine finger of one DnaA molecule and the gamma-phosphate of ATP bound to an adjacent DnaA molecule. These interactions on successive protomers lead to helical filament assembly (Erzberger et al., 2006; Mott and Berger, 2007). As DnaA melts the DNA duplex at the DUE, the exposed single-stranded DNA is bound and extended by the AAA+ domains of the DnaA helical assembly (Bramhill and Kornberg, 1988a; Duderstadt et al., 2011).

The replicative helicases are then recruited. In *B. subtilis*, the helicase loader proteins DnaD, DnaB, and DnaI load the helicase, DnaC, at oriC (Bruand et al., 1995, 2001, 2005; Smits et al., 2010; Velten et al., 2003). These proteins are essential and conserved in low-G+C Gram-positive bacteria, so the *B. subtilis* mechanism is likely to be used in related organisms (Bruck and O’Donnell, 2000; Li et al., 2004, 2007). The helicase loader proteins associate with DnaA at oriC in a defined order: DnaA interacts with and recruits DnaD, which then recruits DnaB; DnaB then associates with the helicase (DnaC) and DnaI (Ishigo-Oka et al., 2001; Smits et al., 2010). DnaB and DnaD weakly interact (Bruand et al., 2005), and they can bind double- and single-stranded DNA (Marsin et al., 2001). DnaB can laterally compact supercoiled DNA, whereas
DnaD can untwist supercoiled DNA, which might facilitate DnaA binding to low affinity sites (Zhang et al., 2005, 2006). DnaD forms complexes on a plasmid and counteracts plasmid compaction, and this function has been loosely compared to that of the nucleoid-associated protein HU in *E. coli* (Turner et al., 2004). DnaB is membrane-associated, as is the process of replication initiation in *B. subtilis* (Hoshino et al., 1987; Rokop et al., 2004; Sueoka, 1998). The recruitment of DnaD to the membrane by DnaB is thought to be important in temporally and spatially regulating replication (Rokop et al., 2004). DnaB and DnaI are thought to facilitate helicase hexamerization around single-stranded DNA (Velten et al., 2003).

Helicase loading is regulated by different mechanisms in other organisms. In *E. coli*, there is one helicase loader protein (DnaC) rather than three, and there are no known homologues of DnaB, DnaD, or DnaI from *B. subtilis*. (The protein names are not consistent between *E. coli* and *B. subtilis.*) The *E. coli* helicase (DnaB) is loaded by direct interaction with oriC-bound DnaA, since the helicase does not interact with single-stranded DNA that is covered by single-strand binding protein (Messer, 2002; Skarstad and Katayama, 2013). In eukaryotes, loading of the Mcm helicases, also as a double hexamer, depends on the loaders Cdc6 and Cdt1 and is highly regulated during the cell cycle (Bell and Dutta, 2002; Kawakami and Katayama, 2010).

Proper positioning of DnaA at oriC is important. Although the *B. subtilis* and *E. coli* DnaA proteins are 51% identical and share a consensus binding sequence, they cannot fully substitute for each other in replication initiation. Purified DnaA proteins can bind to the oriC of the other species but cannot initiate DNA melting (Krause et al., 1997). From studies of a hybrid origin containing the DnaA boxes of *E. coli* and the DUE of *B. subtilis*, the DnaA box region determines the specificity of the initiation. Furthermore, *E. coli* DnaA can load the *E. coli* helicase loader and helicase onto the hybrid origin, even though replication does not occur (Seitz
et al., 2001). Positioning of the DnaA boxes or interaction with other components such as DNA-remodeling proteins could be responsible for this specificity.

In *E. coli* but not *B. subtilis*, several nucleoid-binding and DNA-bending proteins contribute to replication initiation. *E. coli oriC* contains binding sites for IHF (integration host factor) and Fis (factor for inversion stimulation) (Gille et al., 1991; Katayama et al., 2010; Messer, 2002; Polaczek, 1990; Ryan et al., 2004; Schaeffer et al., 2005). IHF binds near the DUE, bends DNA in a manner that enhances the binding of DnaA, and promotes DNA melting (Polaczek, 1990; Ryan et al., 2004; Schaeffer et al., 2005). IHF helps recruit DnaA to lower-affinity sites, whereas Fis inhibits IHF or DnaA binding at low Fis concentrations (Ryan et al., 2004). This role for IHF in promoting replication initiation contrasts with its activation of *datA*, a negative regulator of initiation discussed below. The overall replication phenotype of an IHF mutant shows rapid replication, indicating a net inhibitory role; IHF mutants overinitiate in a DnaA-dependent manner (Von Freiesleben et al., 2000). Another nucleoid-binding protein, HU (heat-unstable), interacts with DnaA (Chodavarapu et al., 2008) and promotes open complex formation, although HU does not bind sequence-specifically to DNA (Bramhill and Kornberg, 1988b; Dixon and Kornberg, 1984; van der Ende et al., 1985; Ogawa et al., 1985). These divergent mechanisms and interactions used in *B. subtilis* and *E. coli* illustrate important functional differences despite the close sequence conservation and similar biochemical properties of the *B. subtilis* and *E. coli* DnaA proteins.
Regulation of replication initiation in *B. subtilis*

In *B. subtilis*, all known regulators of replication initiation inhibit the cooperative binding of DnaA to oriC. These regulators are SirA, Soj, DnaD, and YabA.

**SirA.** The sporulation inhibitor of replication, SirA, inhibits DNA replication during the transition from growth to sporulation. In the sporulation pathway, the final round of replication produces two chromosomes, with one transported to the forespore and the other remaining in the mother cell (Setlow et al., 1991; Wu and Errington, 1994). For proper sporulation, the number of chromosomes must be reduced to two. SirA expression is induced by the sporulation master regulator Spo0A. When artificially overexpressed during growth, SirA causes anucleate cells or guillotined chromosomes leading to lethality (Rahn-Lee et al., 2009; Wagner et al., 2009). SirA interacts with DnaA in a yeast two-hybrid assay (Wagner et al., 2009). SirA inhibits replication initiation by reducing DnaA association with oriC, and SirA depends on DnaA for localization to oriC (Rahn-Lee et al., 2009; Wagner et al., 2009). SirA is conserved in spore-forming *Bacillus* species but not in non-sporulating relatives such as *Listeria* (Rahn-Lee et al., 2009; Wagner et al., 2009). *sirA* mutants are impaired in their ability to reduce chromosome number during sporulation, but another chromosome reduction mechanism likely exists because replication is still reduced to some degree in *sirA* mutants (Rahn-Lee et al., 2009). Nutrient depletion, which typically accompanies the transition to sporulation, is also thought to contribute to replication inhibition before sporulation (Wagner et al., 2009).

**Soj.** Soj is a chromosome partitioning protein that inhibits DnaA helix formation. *soj* was identified as a suppressor of *spo0J*, which is required for sporulation. Soj and Spo0J are homologues of the chromosome segregation proteins ParA and ParB, which bind to parS regions of the chromosome, near oriC, to promote origin separation after replication (Ireton et al., 1994;
Lin and Grossman, 1998). Par proteins are found in nearly all bacteria, although some species, including *E. coli*, do not have homologs, likely through evolutionary loss (Livny et al., 2007). Null mutations of *soj*, *spo0J*, or both lead to overinitiation, indicating a role in inhibiting initiation (Lee and Grossman, 2006; Murray and Errington, 2008). Single or double mutants also have origin-specific defects in separating replicated DNA (Lee and Grossman, 2006). As a ParA homolog, Soj is a Walker-type ATPase (Leonard et al., 2005). Soj and DnaA form a complex *in vivo*, and they directly interact by SPR and in a bacterial two-hybrid assay (Murray and Errington, 2008; Scholefield et al., 2012).

Soj can activate or inhibit replication initiation via DnaA, depending on the ratio of Soj/Spo0J (Murray and Errington, 2008; Ogura et al., 2003). Monomeric Soj inhibits DnaA helix formation, likely by preventing the conformational change required for helix assembly (Scholefield et al., 2012). Soj can undergo ATP-dependent dimerization, and dimeric Soj binds DNA (*parS*) (Leonard et al., 2005). According to the model for Soj regulation of DnaA, Soj is predominantly in the monomeric, unbound state that inhibits initiation. Switching to the dimeric, DNA-bound state, in response to some cellular cue, is thought to relieve this inhibition (Murray and Errington, 2008). Spo0J stimulates Soj ATPase activity, which promotes release of Soj from DNA (Leonard et al., 2005). The action of Soj and Spo0J on DnaA may couple replication initiation to the partitioning of replicated chromosomes.

**DnaD.** The helicase loader DnaD is recruited by DnaA at oriC to promote primosome assembly, but DnaD also inhibits cooperative binding and helix formation of DnaA at oriC (Bonilla and Grossman, 2012; Scholefield and Murray, 2013). DnaA-ATP binds the oriC region *in vitro* in a highly cooperative fashion (Merrikh and Grossman, 2011; Miller et al., 2009). Addition of DnaD makes this binding noncooperative (Bonilla and Grossman, 2012). DnaD has
the same effect on cooperative binding of DnaA to another DnaA-box-containing sequence, the chromosomal region yydA, suggesting that DnaD could affect DnaA function at chromosomal regions beyond oriC (Bonilla and Grossman, 2012). In another in vitro assay, DnaD inhibits the formation of the DnaA-ATP multimeric helix at oriC but does not disrupt previously formed helices, consistent with a role in affecting initial DnaA-DNA binding (Scholefield and Murray, 2013). Although DnaD promotes initiation in its capacity as a helicase loader protein, its regulation of DnaA likely inhibits initiation, given the importance of cooperative DnaA binding. Inhibition by DnaD has been suggested to keep DnaA inactive at oriC until the availability of DnaD changes during the replication cycle, possibly with DnaB contributing to regulation (Bonilla and Grossman, 2012). It is difficult to test the effects of DnaD in vivo, however, since it is an essential protein.

**YabA.** Another negative regulator of *B. subtilis* replication initiation, YabA, is discussed further in Chapter 3. YabA was identified in a yeast two-hybrid screen for interactors with DnaA, and it also interacts with DnaN, the (beta) processivity clamp of DNA polymerase (Noirot-Gros et al., 2002, 2006). Cells overexpressing yabA under-initiate replication, and cells lacking yabA over-initiate, indicating a negative regulatory function (Hayashi et al., 2005; Noirot-Gros et al., 2002). yabA null cells also have asynchronous replication, and yabA overexpression delays replication, leading to a higher cell mass at initiation (Hayashi et al., 2005; Noirot-Gros et al., 2006).

In *B. subtilis*, YabA is a 14-kDa protein with 119 amino acids. It has no known homology to other proteins, but it is found in other low-GC Gram-positive bacteria including *Lactococcus, Listeria, Staphylococcus*, and *Streptococcus* (Hayashi et al., 2005; Noirot-Gros et al., 2002). YabA contains two predicted domains. The N-terminal domain includes a predicted
leucine zipper, a structure known to mediate specific protein-protein interactions. The C-terminal zinc cluster domain has three cysteines that could coordinate two zinc ions. These cysteines are required for interaction with DnaA and DnaN (Noirot-Gros et al., 2006). YabA forms tetramers and higher-order oligomers (Noirot-Gros et al., 2006; Soufo et al., 2008).

In vivo, YabA inhibits the association of DnaA with oriC, and YabA depends on DnaA for association with oriC and other regions of direct DnaA binding (Merrikh and Grossman, 2011). In fluorescence localization studies of a yabA null mutant, DnaA shows enhanced colocalization with oriC, consistent with the role for YabA in regulating DnaA binding (Soufo et al., 2008). In vitro, YabA inhibits the cooperativity of DnaA binding to oriC (Merrikh and Grossman, 2011) and inhibits helix formation (Scholefield and Murray, 2013). Similarly to DnaD, YabA inhibits the formation of DnaA higher-order structures (Scholefield and Murray, 2013; Soufo et al., 2008) but does not disassemble pre-existing structures (Scholefield and Murray, 2013).

DnaN is thought to act as a positive regulator of replication initiation by antagonizing the function of YabA. Perturbations in DnaN levels affect replication in a YabA-dependent manner; that is, YabA mediates the effect of DnaN on replication initiation. Overproduction of DnaN increases replication initiation and decreased DnaN levels decrease replication initiation. This phenotype depends on YabA (Goranov et al., 2009). Some YabA is associated with the replisome, typically near the end of the replication cycle (Hayashi et al., 2005). YabA forms replication-dependent foci even in strains that replicate in the absence of DnaA and oriC (Goranov et al., 2009), and these foci likely represent association of YabA with the replisome via DnaN.
The YabA-DnaA-DnaN interactions likely function in coordinating replication initiation with the completion of previous rounds of replication. DnaN levels are thought to increase after completion of a round of replication, as the replisome components are released. DnaN is then thought to titrate free YabA away from DnaA bound to oriC, restoring cooperative binding of DnaA and allowing replication initiation (Merrikh and Grossman, 2011). DnaN localizes with replication forks, as a DnaN dimer is associated with every priming event on the lagging strand (Bruck and O’Donnell, 2001). In response to replication stress, both DnaN and YabA foci dissipate and DnaN is released from the replication complex (Goranov et al., 2009). Consistent with the model of DnaN titrating YabA away from DnaA at oriC, both replication stress and DnaN overexpression decrease YabA association and increase DnaA association with oriC and other chromosomal regions of direct DnaA binding (Merrikh and Grossman, 2011). YabA inhibits the association of DnaA with these additional DnaA-box regions similarly to oriC (Merrikh and Grossman, 2011), so the effects of YabA on DnaA may not be limited to replication initiation. YabA also interacts with the methyl-accepting chemotaxis proteins TlpA and McpA, suggesting a possible connection to chemotaxis and signaling (Noirot-Gros et al., 2002).

Interactions of regulators with DnaA. Although the four known B. subtilis DnaA regulators all affect cooperative binding of DnaA to oriC, the regulators are thought to interact with distinct regions of DnaA. SirA interacts with Domain I of DnaA, which is also an important regulatory site in other organisms. Three particular DnaA residues, predicted to be located on the same patch of the protein surface, are required for the interaction (Rahn-Lee et al., 2011). Residue A50 of DnaA is an important determinant of the SirA-DnaA interaction. Bacterial species with SirA homologs have DnaA proteins with alanine at this position, and
species without SirA have valine or threonine. Changing the corresponding *E. coli* residue from valine to alanine was sufficient to confer interaction with SirA (Rahn-Lee et al., 2011). In Gram-negative bacteria, this region of the DnaA Domain I instead interacts with the regulators DiaA of *E. coli* (Keyamura et al., 2007) or HobA of *H. pylori* (Natrajan et al., 2009), both of which affect DnaA binding to oriC (see below). These regulators have structural but not sequence homology to each other (Natrajan et al., 2007) and no known homology or predicted structural similarity to SirA (Rahn-Lee et al., 2011). Despite their divergent origins, these three regulators affect the same function of DnaA via interaction with the same region of the protein.

Soj, YabA, and DnaD are thought to interact with the AAA+ region of DnaA. Soj likely interacts with Domain IIIB, based on suppressor mutations in DnaA (Murray and Errington, 2008; Scholefield et al., 2012). YabA is suspected to interact with Domain IIIA of DnaA, on the opposite surface from the ATP-binding pocket, based on a yeast two-hybrid assay with DnaA fragments (Cho et al., 2008). DnaA mutants with reduced YabA interaction also showed reduced interaction with DnaD, suggesting that DnaD interacts via a similar region (Cho et al., 2008). This surface of DnaA is predicted to be buried in the full-length protein, however (Cho et al., 2008; Erzberger et al., 2002). These residues are unlikely to be directly accessible to regulators, so the interpretation of these DnaA mutants is unclear; perhaps their effects on interaction are due to larger-scale conformational changes in DnaA.

It has been suggested that regulation of DnaA helix assembly, by keeping DnaA-ATP bound and poised at oriC, could allow cells to respond rapidly to an environmental cue that triggers initiation. Soj, for example, might respond to such a cue (e.g., nutrient availability, developmental signals, or cell-cycle information) by transitioning from an inhibitory monomer to
the permissive dimer form, with the action of Spo0J (Scholefield et al., 2012). How such cues could work is still an open question.

**Regulation of replication initiation in other organisms**

**Regulation of DnaA protein levels.** One mechanism used by other bacteria to regulate replication initiation is the regulation of DnaA protein levels, in addition to dnaA transcriptional regulation described above. In Caulobacter crescentus, unlike E. coli and B. subtilis, DnaA is unstable and degraded after replication initiation (Gorbatyuk and Marczynski, 2005). DnaA degradation is also used to arrest growth in response to proteotoxic stress or carbon starvation. In the proteotoxic stress response, depletion of the chaperone DnaK induces synthesis of the Lon protease, which degrades DnaA (Jonas et al., 2013). Under carbon starvation, DnaA degradation depends on the stringent response control protein SpoT (Lesley and Shapiro, 2008).

**Sequestration of oriC.** A second category of mechanisms involves sequestration of oriC to prevent DnaA binding. In E. coli, the SeqA protein binds to newly replicated and therefore hemimethylated sequences in oriC to occlude DnaA binding. The E. coli origin of replication contains eleven copies of the Dam methylase recognition sequence (5’-GATC-3’), and hemimethylated sites are preferentially bound by SeqA. SeqA inhibits DnaA binding by masking low-affinity DnaA binding sites (Lu et al., 1994; Mott and Berger, 2007; Nievera et al., 2006). SeqA binding is highly cooperative, and it forms higher-order complexes in a filament, reminiscent of DnaA helical oligomerization (Brendler and Austin, 1999; Guarné et al., 2005; Skarstad et al., 2000). SeqA also inhibits dnaA transcription by sequestering the dnaA gene (Campbell and Kleckner, 1990). Another methyltransferase, CcrM, is found in alpha-
proteobacteria and is also implicated in affecting replication synchrony and dnaA transcription (Donczew et al., 2014). *B. subtilis* lacks seqA and dam, so this specific mechanism is not used.

In *Caulobacter*, binding of the cell-cycle transcriptional regulator CtrA to oriC prevents replication initiation in a cell-type-specific manner. CtrA is present in swarmer cells, which do not initiate DNA replication, but absent from stalked cells, which are competent for replication. CtrA prevents initiation by occluding an essential DnaA box in oriC (Quon et al., 1998). CtrA is degraded by ClpP to allow differentiation into stalked cells that are competent for replication (Gorbatyuk and Marczynski, 2005; McGrath et al., 2006).

In *Streptomyces coelicolor*, a bacterium that forms a sporulating fungus-resembling structure called an aerial mycelium, the mycelium formation master regulator AdpA plays a similar role in occluding oriC. AdpA specifically binds the 5’ region of oriC at sequences that partially overlap DnaA boxes. AdpA and DnaA compete for binding using an ATP-dependent mechanism. In contrast to DnaA, AdpA has lower affinity for AdpA boxes when the protein is bound to ATP or ADP. Consequently, under conditions of high ATP, DnaA binding is favored, whereas low ATP favors AdpA binding. The inhibition of replication initiation by AdpA affects the timing of aerial hyphae emergence and likely helps coordinate different growth rates for different parts of the structure (Wolanski et al., 2012). These examples from aerial mycelium formation and swarmer-cell differentiation, along with sporulation in *B. subtilis*, illustrate the different replication needs of differentiated states specific to the lifestyle of each organism.

Several other proteins in various organisms are known to bind oriC, in many cases as inhibitors of initiation, but their connections to the cell cycle or signaling are not well understood. In *E. coli*, the inhibitor of chromosomal initiation IciA binds oriC and blocks initiation *in vitro* (Thöny et al., 1991). IciA is also known as ArgP for its role as an arginine-
responsive transcription factor in the LysR family, and it counteracts DnaA autorepression of dnaA transcription (Lee et al., 1996). Mycobacterium tuberculosis has a homolog of IciA that binds the DUE and prevents open complex formation by DnaA in vitro (Kumar et al., 2009). ArcA, a response regulator involved in E. coli anaerobic growth signaling, binds the left region of oriC and also inhibits open complex formation (Lee et al., 2001). Rob (right origin binding) is another E. coli oriC-binding protein, but it does not show a replication phenotype and its function is unclear (Skarstad et al., 1993).

**Sequestration of DnaA.** A third mechanism to regulate replication uses the opposite approach of sequestering DnaA. In E. coli, the DnaA-titrating locus datA locus titrates free DnaA away from oriC. datA is a 1-kb region that contains only five DnaA boxes but in immunoassays binds eight times as many DnaA molecules as oriC (Kitagawa et al., 1996). It is located near oriC, approximately 250 kb away, and may serve as a sink for DnaA as datA is replicated soon after initiation (Kitagawa et al., 1996, 1998). Loss of datA can be suppressed by DnaA-titrating plasmids (Kitagawa et al., 1998). The titrating ability of datA depends on the DNA-bending protein integration host factor (IHF). datA contains an IHF binding site between two DnaA boxes, and IHF binds in vitro and in vivo to enhance the number of DnaA molecules bound (Nozaki et al., 2009). These datA-DnaA complexes might interact with oriC-DnaA complexes directly, in addition to titrating free DnaA (Nozaki et al., 2009). The mechanism of datA is consistent with the initiator titration model, in which initiation occurs when free DnaA-ATP levels exceed a threshold. Below this threshold, DnaA primarily occupies non-oriC sites, and above the threshold, newly synthesized DnaA can bind to oriC and promote initiation (Hansen et al., 1991).
Other chromosomal regions in different species contain DnaA boxes but may not function like \textit{datA}. The \textit{E. coli} chromosome contains five sites, other than \textit{oriC}, that bind DnaA with high affinity via one or two strong DnaA boxes. These sites are not known to have a regulatory function and appear randomly distributed along the chromosome (Roth and Messer, 1998). Similarly, the \textit{B. subtilis} chromosome contains six clusters of DnaA boxes outside \textit{oriC}, and their effects on replication initiation are modest. Deletion of all six clusters leads to early initiation of replication, but deletion of any single cluster has little or no effect. The phenotype is likely due to differences in copy number, since insertion of DnaA box clusters can rescue the overinitiation phenotype, with insertions closer to \textit{oriC} being more effective and those closer to \textit{terC} requiring more copies to rescue the same phenotype. A plasmid containing one of the clusters can bind DnaA but not complement overinitiation, indicating that the DnaA box clusters in \textit{B. subtilis} function differently from the more straightforward titration mechanism in \textit{E. coli} (Okumura et al., 2012). DnaA availability in \textit{B. subtilis} might also be affected by the regulators YabA and DnaD, as well as the helicase loader DnaB. These three proteins are associated with DnaA at regions other than \textit{oriC}, and YabA and DnaD affect DnaA box cluster regions similarly to \textit{oriC} (Merrikh and Grossman, 2011; Smits et al., 2011). Titration-like mechanisms are suspected to be relevant to \textit{B. subtilis} as well as \textit{E. coli} because \textit{B. subtilis} replication is more sensitive than \textit{E. coli} replication to a multicopy plasmid containing DnaA boxes (Katayama et al., 2010; Moriya et al., 1988).

\textbf{Regulation of DnaA nucleotide binding.} A fourth category of mechanisms is regulation of DnaA nucleotide-bound state. Different mechanisms promote ATP hydrolysis or recharge DnaA to the ATP-bound form. In \textit{E. coli}, the regulator Hda (homologous to DnaA) forms a complex with DnaN (the processivity clamp of DNA polymerase) to promote ATP hydrolysis by
DnaA. DnaA-ATP increases before initiation, and ATP is hydrolyzed after initiation (Kurokawa et al., 1999). Since ADP-DnaA is inactive for replication initiation, this inhibitory mechanism is known as regulatory inactivation of DnaA (RIDA). Hda is a chaperone-like AAA+ protein and a parologue of DnaA. The overall structure of Hda is homologous to the AAA+ domain of DnaA (Kato and Katayama, 2001). Hda is found in *Haemophilus influenzae*, *Xylella fastidiosa*, *Neisseria meningitidis* and *Rickettsia* and likely other organisms (Kato and Katayama, 2001). Hda promotes ATP hydrolysis by DnaA in vivo and in vitro (Kato and Katayama, 2001). Hda itself has a high affinity for ADP, not ATP, and the active form is monomeric ADP-Hda (Su’etsugu et al., 2008). The interaction between Hda and DnaA likely involves the interface of AAA+ domains (Su’etsugu et al., 2005). Hda-ADP monomers are thought to allow interaction between the arginine finger of Hda and the ATP bound to the DnaA AAA+ domain (Su’etsugu et al., 2008). It has been proposed that the free C-terminal end of a DnaA-ATP helical filament could provide access to Hda (Mott and Berger, 2007). Hda is estimated at 100 molecules per cell (Su’etsugu et al., 2005), so this mechanism of inhibition would allow relatively few Hda molecules to inhibit a large DnaA complex.

Regulatory inactivation of DnaA requires DnaN to be bound to DNA and to Hda (Nishida et al., 2002; Su’etsugu et al., 2005). Hda directly interacts with DnaN via a hexapeptide beta-binding motif at the N-terminus of Hda (Kurz et al., 2004). This motif is a variant of the widely distributed, versatile pentapeptide motif that enables a variety of bacterial polymerases and mismatch repair proteins to interact with a hydrophobic pocket of the sliding clamp (Dalrymple et al., 2001). Since clamps associate with the replisome, the Hda-DnaN interaction is thought to couple active replication with the inactivation of DnaA, to prevent immediate re-initiation and limit the number of initiations per cell cycle (Kato and Katayama, 2001; Kurokawa et al., 1999).
YabA, as another clamp-associated negative regulator, has been compared to Hda in the timing of regulation, although YabA inhibits DnaA cooperative binding rather than ATP hydrolysis. Thus, different organisms can use the processivity clamp to regulate initiation via different associated regulators and mechanisms for inhibiting DnaA. *Caulobacter* has an Hda homologue, HdaA, which is required for replication regulation, colocalizes with the replisome, and interacts with DnaN. It is also transcriptionally activated by DnaA, allowing Hda levels to adjust to DnaA levels (Collier and Shapiro, 2009; Fernandez-Fernandez et al., 2013). The interaction with DnaN suggests a similar role as *E. coli* Hda, although a different mode of regulation is entirely possible, given the case of *B. subtilis* DnaN and YabA. In eukaryotes, the processivity clamp (proliferating-cell nuclear antigen, PCNA) is also used to inactivate initiation (Arias and Walter, 2006; Katayama et al., 2010). It appears that bacterial, eukaryotic, and archaeal sliding clamps all use loosely similar binding motifs (Dalrymple et al., 2001). The conserved but versatile binding region of DnaN may then allow a variety of organism-specific regulators to use clamp association as a temporal control mechanism.

Conversely, there are mechanisms to regenerate DnaA-ATP from DnaA-ADP. In *E. coli*, two chromosomal regions called DnaA-reactivating sequences (DARS1 and DARS2) bind DnaA-ADP and promote nucleotide release, likely by the formation of higher-order DnaA complexes. DARS1 is located slightly more than halfway between oriC and terC on the right side of the chromosome, between *bioD* and *uvrB*. Its minimal activating sequence is approximately 100 bp. DARS2 is located slightly less than halfway between oriC and terC on the left side of the chromosome, between *ygpD* and *mutH*, and its minimal activating sequence is approximately 500 bp. DARS1 and DARS2 contain at least one consensus DnaA box and two or more DnaA boxes with one or more mismatches from the consensus (Fujimitsu et al., 2009).
The originally identified DARS belonged to the ColE1 plasmid origin and contained one DnaA box and two sequences with one or two mismatches (Fujimitsu and Katayama, 2004). DARS1 and DARS2 promote replication initiation \textit{in vivo}, and they increase levels of DnaA-ATP additively, independently of protein synthesis. Deletion of either DARS rescues the lethality of an \textit{hda} null mutation and suppresses the overinitiation in \textit{datA} or \textit{seqA} mutants. \textit{In vitro}, both DARSs promote dissociation of ADP from DnaA, including DnaA-ADP that was inactivated by a reconstituted RIDA system. Both DARS generate DnaA-ATP that is capable of replication in a minichromosome system. The activity of DARS2 is stimulated by crude cell extract, suggesting the existence of a positive regulator, and DARS2 likely has higher activity \textit{in vivo}, given the more severe phenotype of DARS2 mutants relative to DARS1 mutants. DnaA-ADP forms higher-order complexes on DARS1 \textit{in vitro}, but apo-DnaA, with no nucleotide bound, does not stably associate with DARS, consistent with a mechanism of dissociation upon ADP release. In this way, DARS is thought to serve a catalytic function in recycling DnaA. Other proteobacteria have DnaA boxes in the same locations and orientations as those in DARS1 and DARS2, and this mechanism might extend to other bacteria with the appropriately positioned DnaA boxes elsewhere on the chromosome (Fujimitsu et al., 2009).

Acidic phospholipids such as cardiolipin and phosphatidylglycerol are also thought to promote nucleotide exchange of DnaA in \textit{E. coli}. Cardiolipin (diphosphatidylglycerol), a cell membrane phospholipid, can displace ADP from DnaA and can bind nucleotide-free DnaA to prevent ATP loading (Sekimizu and Kornberg, 1988). Interaction of DnaA with \textit{oriC} was required for this phospholipid-based reactivation (Crooke et al., 1992). Depletion of acidic phospholipids can inhibit plasmid replication from \textit{oriC} (Xia and Dowhan, 1995), but it is
unknown how acidic phospholipids coordinate DnaA activity with environmental cues or the cell cycle (Skarstad and Katayama, 2013).

None of the known *B. subtilis* regulators are known to affect the nucleotide-bound state of DnaA. YabA (Merrikh and Grossman, 2011; Scholefield and Murray, 2013), DnaD (Bonilla and Grossman, 2012), and monomeric Soj (Scholefield et al., 2012) do not affect ATP hydrolysis or nucleotide exchange. The nucleotide exchange rates in DnaA proteins from different organisms likely explain why ATPase activity is not known to be regulated in *B. subtilis*. In *E. coli*, nucleotide exchange is slow (half-life of 45 min) relative to hydrolysis (half-life of 15 min), so ATPase activity can be regulated (Sekimizu et al., 1987). In several Gram-positive organisms, the nucleotide exchange rate is much faster: 1 min in *Staphylococcus aureus*, 12 min for *Streptococcus pyogenes*, and 5 min for *B. subtilis* (Bonilla and Grossman, 2012; Kurokawa et al., 2009). This timescale is not expected to allow efficient regulation of ATPase activity.

**Regulation of cooperative DnaA binding to oriC.** Regulation of DnaA cooperative binding to *oriC*, the predominant mechanism in *B. subtilis*, is also used by other organisms. In *E. coli*, DiaA (DnaA initiator-associating factor) promotes cooperative binding of DnaA to *oriC*, in contrast to the negative regulators of cooperative binding in *B. subtilis*. DiaA is required for the proper timing of replication initiation, and it stimulates initiation *in vitro* (Ishida et al., 2004). It forms tetramers, as pairs of dimers, and each subunit simultaneously binds DnaA (Ishida et al., 2004; Keyamura et al., 2007). DiaA has similar affinity for the ATP- and ADP-bound forms of DnaA (Ishida et al., 2004). There are about 280 DiaA dimers per cell (Ishida et al., 2004). DiaA tetramerization is required for stimulating initiation, as is a putative phosphosugar binding motif in DiaA (Keyamura et al., 2007). By stimulating assembly of DnaA at *oriC*, DiaA also promotes conformational changes in the DnaA-ATP complex and *oriC* unwinding (Keyamura et al., 2007).
DiaA can also stimulate assembly of DnaA-ADP-oriC complexes, although these are not active for replication (Keyamura et al., 2007). The N-terminus of DnaA is important for binding to both DiaA and the E. coli helicase (DnaB). DiaA can inhibit DnaB loading, suggesting competition for a shared region of interaction (Keyamura et al., 2009). A functional analogue of DiaA, called HobA, is found in H. pylori and conserved only among epsilon-proteobacteria (Zawilak-Pawlik et al., 2007). HobA and DiaA show structural as well as functional similarity, but not sequence similarity. The N-terminal Domain I is sufficient to allow cross-species interaction of HobA or DiaA with the heterologous DnaA (Zawilak-Pawlik et al., 2011). HobA interacts with DnaA in the oriC-DnaA complex, and like DiaA has a positive role. HobA depletion reduces initiation, although there is no growth phenotype of HobA overexpression (Zawilak-Pawlik et al., 2007). They also differ in their temporal effects; HobA accelerates DnaA binding to oriC whereas DiaA slows DnaA binding (Zawilak-Pawlik et al., 2011).

**Effects of transcription near oriC on replication initiation.** Finally, a somewhat controversial mode of regulation involves transcription in the oriC region. Transcription of genes near oriC is thought to interfere with replication initiation by preventing replisome assembly or introducing transcriptional structures that disrupt DNA topology in the oriC region (Donczew et al., 2014; Messer and Weigel, 1997). Since DUE unwinding is favored under negative supercoiling, positive supercoiling in front of the transcriptional machinery is expected to inhibit initiation, whereas negative supercoiling behind the transcription bubble is expected to promote initiation. *In vitro* studies with minichromosomes have supported the role of DNA topology. Transcription of the mioC gene, upstream of oriC, inhibits initiation, whereas transcription of gidA, downstream of oriC, enhances initiation (Baker and Kornberg, 1988; Donczew et al., 2014; Tanaka and Hiraga, 1985). DnaA represses mioC transcription (Messer
and Weigel, 1997), and transcription of these promoters in vivo is coordinated with the cell cycle and consistent with the effects demonstrated in vitro. gidA transcription is high before initiation and mioC transcription is high after initiation (Ogawa and Okazaki, 1994). Several Mycobacterium species have also been suggested to regulate initiation by transcription near oriC (Donczew et al., 2014). Chromosomal replication is almost certainly regulated differently than from minichromosomes, however, and some in vivo studies do not support the role of transcription in regulating chromosomal initiation (Bates et al., 1997). Transcriptional regulation of initiation is complex and likely redundant, since simple deletion of mioC or the mioC DnaA box does not result in an initiation phenotype (Løbner-Olesen and Boye, 1992). Additionally, transcriptional effects are likely to be condition-sensitive and difficult to isolate from other regulatory mechanisms (Donczew et al., 2014).

Together, these examples illustrate the diversity of mechanisms and interactions used to regulate replication initiation by DnaA. Regulators with no known sequence similarity, such as SeqA, CtrA, and AdpA, can perform similar functions in different organisms with distinct developmental pathways. Similar aspects of regulation, such as association with DnaN, can be paired with different mechanisms of inhibition, as in Hda from E. coli and YabA from B. subtilis. Regulation of DnaA as a transcription factor is not as well described as regulation of initiation, but the regulators that affect DnaA multimerization and DNA binding at oriC might also be candidates for regulating DnaA-DNA binding in a transcription factor capacity.
Gene regulation by DnaA

DnaA also functions as a transcription factor, although the regulation of this activity is not well understood. In addition to autoregulation, DnaA can activate or repress gene expression in *E. coli, B. subtilis*, and likely other organisms (Burkholder et al., 2001; Ishikawa et al., 2007; Messer and Weigel, 1997; Ogura et al., 2001). In *B. subtilis*, DnaA regulates several genes directly and many genes indirectly. DnaA binds to six chromosomal regions containing clusters of DnaA boxes, in addition to *oriC* and the *dnaA* promoter (Breier and Grossman, 2009; Ishikawa et al., 2007). These regions include *sda, ywlC, ywcI, yydA, gcp*, and *trmE*. They each include at least four DnaA boxes, and all except *yydA* contain at least one consensus DnaA box (Ishikawa et al., 2007). Five of the six regions (excluding *trmE*) show transcriptional regulation by DnaA ((Goranov et al., 2005; Ishikawa et al., 2007); T.A. Washington, J.L. Smith, and ADG, in preparation). The DnaA box clusters are located directly upstream of genes, except for *gcp*, at which the boxes are downstream of the coding sequence. DnaA may regulate transcriptional termination rather than initiation at this site. At *ywcI, ywlC*, and *yydA*, DnaA represses transcription, and DnaA activates transcription at *sda* and *gcp* ((Ishikawa et al., 2007); T.A. Washington, J.L. Smith, and ADG, in preparation). The orientation, spacing, and affinity of DnaA boxes may contribute to these differences, or additional regulators may be involved.

An important role for DnaA is the transcriptional response to replication stress. Replication can stall due to DNA damage, defects in replication proteins, chemical inhibition, or even under normal growth conditions (Cox et al., 2000). Genes affected by replication stress include those involved in nucleotide synthesis, cell division, DNA repair, oxidative stress response, and other processes. Many bacteria use the well-characterized global DNA damage (SOS) response, as well as other pathways, to respond to replication stress (Kreuzer, 2013).
DnaA is required for the transcriptional response of many genes to replication stress, via direct and indirect regulation (Goranov et al., 2005; T.A. Washington, J.L. Smith, and ADG, in preparation).


The indirect transcriptional regulation by DnaA is largely mediated by Sda. Sda, named for its role as a suppressor of DnaA, serves as a checkpoint by inhibiting sporulation in response to DNA damage (Burkholder et al., 2001). Sda is activated by DnaA (Goranov et al., 2005; Ishikawa et al., 2007; T.A. Washington, J.L. Smith, and ADG, in preparation) and represses sporulation via the sporulation master regulator Spo0A. Sda inhibits KinA, an upstream kinase in the phosphotransfer pathway that leads to phosphorylation and activation of Spo0A (Burbulys et al., 1991; Burkholder et al., 2001). Sda is a highly unstable protein, so transcriptional activation by DnaA maintains Sda levels necessary for repressing sporulation. Sda has a half-life of 1-4 minutes and is degraded by the ClpXP proteasome. This proteolysis is required for sporulation initiation and recovery from replication stress (Ruvolo et al., 2006). Sda expression cycles with cell growth, with a burst of expression correlating with replication (Veening et al.,
This coordination with the cell cycle helps separate the processes of replication and sporulation. Sda is responsible for many of the indirect effects of DnaA on gene expression. Global gene expression in an sda null mutant is similar to that of a dnaA null mutant. Likewise, overexpression of sda or deletion of spo0A, the target inhibited by Sda, overcomes many of the gene expression effects of a dnaA null mutant (T.A. Washington, J.L., Smith, and ADG, in preparation). Spo0A affects several other regulators, including AbrB, SinR, and PhoP, and expression of these regulons is affected by DnaA. Thus, through Sda and Spo0A, DnaA connects replication to the expression of genes involved in many other processes (T.A. Washington, J.L., Smith, and ADG, in preparation).

_E. coli_ offers additional examples of DnaA transcriptional regulation. DnaA is known to repress several _E. coli_ genes: mioC (near oriC and implicated in transcriptional regulation of initiation), rpoH, uvrB, and proS. In most cases, one or more DnaA boxes are located 20 to 80 bp of the transcriptional start site, although for proS the bound DnaA box is 220 bp upstream. DnaA boxes are found in different orientations and may include one or more mismatches (Messer and Weigel, 1997). DnaA is known to activate glpD and fliC. Both of these promoters have two DnaA boxes mostly mismatches 20-70 bp upstream of the transcriptional start site. The mechanism of activation is unknown, but DnaA is structurally similar to the NtrC family of enhancers, suggesting that the long-range loop-forming activation mechanism of NtrC might be possible for DnaA (Messer and Weigel, 1997). _E. coli_ DnaA interacts with RNA polymerase in an enzyme-linked immunosorbent assay (ELISA), but it is not clear which subunit is involved or precisely how the interaction contributes to transcriptional regulation (Flåtten et al., 2009). Genetic studies in _E. coli_ suggest that the beta subunit is a likely interaction site (Bagdasarian et al., 1977; Szalewska-Pałasz et al., 1998). DnaA can also terminate transcription by binding.
DnaA boxes in the coding sequence; examples include \textit{asnC} (Schaefer and Messer, 1988) and likely \textit{guaB} (Messer and Weigel, 1997). Binding two weak DnaA boxes can loop DNA in a way that promotes termination (Konopa et al., 1999).

In \textit{E. coli}, DnaA regulates the \textit{nrd} (ribonucleotide reductase) operon to adjust the available pool of deoxyribonucleotides according to replication status. The ribonucleotide reductase (RNR) enzyme, encoded by the \textit{nrdAB} operon, catalyzes the production of dNDPs from NDPs, which is the limiting step in dNTP production. The rate and fidelity of replication depend on RNR and dNTP supply. DnaA can activate or repress \textit{nrdAB} expression depending on the nucleotide-bound state of DnaA, which is itself regulated (Herrick and Sclavi, 2007; Olliver et al., 2010). According to the current model, when dNTP levels are high, DnaA predominantly binds ATP and represses its own transcription and \textit{nrdAB}. When dNTP levels are low or DnaA-ATP is inactivated by the regulator Hda, DnaA is primarily bound to ADP, and both \textit{dnaA} and \textit{nrdAB} expression are enhanced. Replication stress induces \textit{dnaA} and \textit{nrdAB} expression in \textit{E. coli} (Herrick and Sclavi, 2007). The \textit{nrd} genes are also upregulated in response to replication stress in \textit{B. subtilis} ((Goranov et al., 2005); T.A. Washington, J.L., Smith, and ADG, in preparation). Although variations in the nucleotide-bound state of DnaA are suspected to affect transcription and the response to replication stress in \textit{B. subtilis} (Goranov et al., 2005), there is not the same feedback mechanism as in \textit{E. coli}, since nucleotide hydrolysis and exchange are not known to be regulated in \textit{B. subtilis}.

Since DnaA associates with chromosomal regions in addition to \textit{oriC}, and certain regulators of DnaA binding also associate with DnaA at these regions, it is possible that replication regulators may be influencing the activity of DnaA at these regions. In particular, YabA and DnaD associate with DnaA at certain chromosomal regions containing clusters of
DnaA boxes, and they depend on DnaA for this association (Merrikh and Grossman, 2011; Smits et al., 2011). Although YabA is not required for the transcriptional response to replication stress (Goranov et al., 2009), YabA associates with DnaA at certain replication-stress-responsive genes (e.g., ywlC, yydA) and might be involved in a more nuanced effect on transcription. In Chapters 2 and 3, I discuss the genome-wide association of DnaA and YabA and their connection to the nucleoid-associated protein Rok.

**Rok**

**Function and Regulation**

Rok is a nucleoid-associated protein and a transcriptional repressor. It was identified by and named for its function as a repressor of *comK*, the master regulator of the competence pathway (Hoa et al., 2002). ComK is a transcription factor that is necessary and sufficient for competence development (Hahn et al., 1996; Van Sinderen et al., 1994, 1995). It regulates more than 100 genes, primarily as an activator (Berka et al., 2002; Hamoen et al., 2002; Ogura et al., 2002; Van Sinderen et al., 1995), and positively regulates its own transcription (van Sinderen and Venema, 1994). Genes upregulated by ComK include operons involved in DNA binding (*comC, comG*), translocation (*comE, comF*), and recombination (*addAB, recA*), as well as other functions associated with a growth-arrested physiology known as the K-state (Berka et al., 2002; Haijema et al., 1995; Hamoen et al., 2002; Ogura et al., 2002; Van Sinderen et al., 1995).

Competence is stimulated by two secreted peptides that signal high cell density. Competence stimulating factor (CSF) and ComX stimulate the ComP-ComA two-component signaling system. Upon entering the cell, CSF inhibits the phosphatase RapC, increasing phosphorylation of the response regulator ComA. Similarly, extracellular ComX activates the
membrane-bound histidine kinase ComP, which stimulates phosphorylation of ComA. Phosphorylated ComA activates transcription of *comS*. In the absence of competence-stimulating peptides, ComK is degraded by the ClpP protease, which forms a complex with ClpC, MecA, and ComK. In the presence of the peptides, ComS displaces ComK from the inhibitory complex and allows ComK to activate competence genes (Lazazzera and Grossman, 1997; Magnuson et al., 1994; White, 2007).

Deletion of *rok* increases the proportion of cells that develop competence. Competence development occurs in a minority of cells during late exponential growth, under specific nutrient conditions at high cell density (Hamoen et al., 2003). Expression of *comK* is bistable due to positive autoregulation, and ComK levels must exceed a threshold to initiate competence development (Maamar and Dubnau, 2005). In competent cells, which must have exceeded this threshold, the concentration of ComK tetramers is estimated to be at least 100 µM (Hoa et al., 2002). In wild-type cells, only about 10% develop competence (Cahn and Fox, 1968), whereas over 60% of cells develop competence in a *rok* null mutant. Consequently, transformation efficiency in a *rok* null mutant increases five-fold (Hoa et al., 2002). Rok overexpression was reported to be lethal (Hoa et al., 2002), although the threshold protein level and effects of strain background are unclear. I have overexpressed a myc-tagged version of Rok from an IPTG-inducible promoter without any obvious adverse effects, although the tag may have compromised Rok function.

ComK activates expression at its own promoter by antagonizing repression by Rok and another repressor, CodY. ComK binds to the minor groove of DNA (Hamoen et al., 1998) whereas Rok binds to the major groove (Smits et al., 2007). Since ComK activation occurs without displacing the repressors from the *comK* promoter, ComK and Rok likely bind to
different surfaces of the DNA helix (Smits et al., 2007). At the comG promoter, representative of the competence genes that are regulated by ComK but not by other known regulators, ComK activates transcription *in vitro* but does not require the alpha subunit of RNA polymerase (Susanna et al., 2004). ComK has instead been proposed to regulate transcription by affecting DNA topology (Susanna et al., 2004). By extension, Rok has also been proposed to modulate DNA topology at ComK-regulated genes, perhaps by causing DNA to adopt a conformation unfavorable for RNA polymerase binding (Smits et al., 2007).

Rok is regulated by several other proteins involved in stationary phase physiology, and one of these regulators, SinR, belongs to a genetic network involving DnaA. Rok and ComK both directly repress *rok* transcription by binding to the promoter (Hoa et al., 2002). The transcriptional regulators AbrB and SinR also repress *rok* expression (Hoa et al., 2002). AbrB is a highly pleiotropic DNA-binding protein involved in expression of genes in the transition from exponential growth to stationary phase (Chumsakul et al., 2011; Perego et al., 1988; Robertson et al., 1989; Strauch et al., 1989; Zuber and Losick, 1987). SinR, the master regulator of biofilm formation, affects genes responsible for the transition from single-celled motile forms to multicellular aggregations (Kearns et al., 2005). The sporulation regulator Spo0A represses expression of *abrB* (Hahn et al., 1995) and inhibits the DNA-binding activity of SinR via SinI (Bai et al., 1993). Since DnaA indirectly regulates Spo0A, Rok and DnaA belong to a shared genetic network of regulators of sporulation and competence development. The relationship between DnaA and Rok in this network is not straightforward, however, since many of the regulators are pleiotropic or involved in feedback loops, and temporal regulation is not well understood for all interactions. I discuss the direct interaction between Rok and DnaA in Chapter 2.
Gene regulation by Rok

Independently of ComK, Rok regulates the expression of approximately 40 genes, many of which are involved in extracellular functions (Albano et al., 2005). Rok directly represses at least seven operons during growth and/or in stationary phase, and Rok binds to many of these promoters in vitro (Albano et al., 2005). Several Rok-regulated operons are involved in producing peptide antibiotics. Subtilosin A, a small cyclic peptide encoded by the sbo operon, has bactericidal activity against Listeria monocytogenes and other species, including Gram-positive and Gram-negative pathogens (Babasaki et al., 1985; Noll et al., 2011; Shelburne et al., 2007; Zheng et al., 1999). Subtilosin interacts with phospholipid bilayers to disrupt the conformation of the head groups, disrupts the transmembrane pH gradient and electrical potential, and is thought to permeabilize cell membranes (van Kuijk et al., 2012; Noll et al., 2011; Thennarasu et al., 2005). Sublancin, an antibiotic encoded by the sun operon in the prophage SPβ (Paik et al., 1998), has activity against several Gram-positive bacteria. It may target or be imported by the mechanosensitive channel MscL (Kouwen et al., 2009). The derepression of antibiotic synthesis genes, such as the sbo and sun operons, has been postulated to help competent cells by inducing the lysis of nearby cells and release of their DNA (Claverys and Håvarstein, 2007).

Other Rok-regulated secreted products include the extracellular signaling molecule SdpC, which delays sporulation and contributes to cannibalism in stationary phase. Like subtilosin, SdpC might permeabilize membranes using a predicted transmembrane segment (Ellermeier et al., 2006). Rok indirectly promotes expression of the biofilm matrix gene bslA (yuaB) (Kovács and Kuipers, 2011), although Rok is not known to activate transcription directly. BslA is a small secreted protein that confers hydrophobic properties to biofilms and is involved in complex
colony architecture (Hobley et al., 2013; Kovács and Kuipers, 2011). Extracellular functions are a common theme for the Rok regulon, as many Rok-regulated genes with unknown function are predicted to encode transmembrane proteins (yybN operon, yxaJ, yjcN), transporters (yydI), or secreted proteins (yxaL) (Albano et al., 2005; Tjalsma et al., 2004).

**Nucleoid binding by Rok**

Rok preferentially binds to regions of the chromosome with high A+T content and is thus considered a nucleoid-associated protein (Smits and Grossman, 2010). Nucleoid-associated proteins are a diverse class of proteins that are functionally reminiscent of eukaryotic histones (Dillon and Dorman, 2010). In *E. coli*, twelve types of nucleoid-associated proteins have been described, and they vary in their sequence specificity, expression levels, and growth phase dependence (Ali Azam and Ishihama, 1999; Ali Azam et al., 1999). Fluorescent imaging of Rok fused to Yellow Fluorescent Protein shows that Rok associates with the nucleoid, although with preferences for certain regions. This preferential binding, characteristic of certain nucleoid-associated proteins discussed in more detail below, distinguishes Rok from other nucleoid-associated proteins that bind the chromosome uniformly. One such uniformly binding protein is HBsu, the *B. subtilis* homologue of HU (heat-unstable nucleoid protein) from Gram-negative bacteria. *B. subtilis* lacks homologues of the other major nucleoid-associated proteins found in *E. coli* (Ohniwa et al., 2011). In contrast to the well-defined binding sequence of DnaA, Rok is not known to have a specific recognition sequence beyond its general preference for A+T-rich DNA (Smits and Grossman, 2010).

Many chromosomal regions have distinctive A+T content because they were horizontally acquired. Examples of *B. subtilis* horizontally acquired elements bound by Rok include the *sdp*
operon, \textit{yybN} operon, \textit{yefC} region, and the integrative and conjugative element ICE\textit{Bs1} (Smits and Grossman, 2010). These regions are more A+T-rich than the \textit{B. subtilis} genome overall, which already has a relatively low average G+C content of 44\% (Kunst \textit{et al.}, 1997). Some Rok-bound horizontally acquired regions, such as prophages 4, 5, and 6, have G+C content as low as 30\% (Smits and Grossman, 2010). Rok does not bind to horizontally acquired elements with average G+C content, such as the prophage PBSX, and Rok preferentially binds to A+T-rich DNA in a heterologous organism, \textit{E. coli}, indicating that Rok binding is determined primarily by base composition rather than horizontal origin per se (Smits and Grossman, 2010). Rok itself was most likely horizontally acquired. Rok is found only in certain \textit{Bacillus} species, and it is not known to share significant sequence homology with other proteins (Albano \textit{et al.}, 2005).

In addition to repressing expression of horizontally acquired genes, Rok inhibits the excision of the mobile genetic element ICE\textit{Bs1} and likely others. ICE\textit{Bs1} is a 20-kb element that integrates at a tRNA gene in \textit{B. subtilis} and is capable of excision and transfer to a recipient cell (Auchtung \textit{et al.}, 2005). ICE\textit{Bs1} gene expression is repressed during exponential growth, but derepression and ICE\textit{Bs1} excision from the chromosome can be stimulated by production of the signaling regulator RapI or by the SOS response (Auchtung \textit{et al.}, 2007). Rok associates with several genes at the ends of ICE\textit{Bs1}, and loss of \textit{rok} enhances ICE\textit{Bs1} excision four-fold, indicating that Rok represses excision of ICE\textit{Bs1} (Smits and Grossman, 2010). The inhibition of conjugation represents another way in which Rok represses gene exchange, in addition to its role in inhibiting external DNA uptake via repression of \textit{comK}.

In \textit{B. subtilis}, Rok is a 22-kDa protein of 191 amino acids. It is highly basic, with a predicted isoelectric point of 9.3. Based on sequence conservation, the Rok protein has three distinct regions (\textbf{Fig. 3}). The N-terminus (Domain I) is highly conserved, but the central region
(Domain II) is less well conserved. The well conserved C-terminus (Domain III) is necessary and sufficient for DNA binding, and truncation mutants containing Domain III maintain a preference for A+T-rich DNA. Domain III shows some predicted similarity to winged helix DNA binding domains, but the actual structure has not been observed. Mutants lacking Domain I and/or Domain II do not bind DNA as well as the full-length protein, indicating that these domains still contribute to DNA binding, perhaps by enhancing Rok oligomerization (Smits and Grossman, 2010).

A recently discovered small Rok protein lacks most of Domain II and part of Domain III but still binds DNA. This rok gene is found on the conjugative plasmid pLS20 from Bacillus natto, and it is responsible for the inhibitory effect of pLS20 on competence in the host cell. Rok-LS20, at 132 amino acids, shares over 30% identity with chromosomal Rok. Ectopic expression of Rok-LS20 increases transformation efficiency and represses transcription of comK. The chromosomal and plasmid rok genes have additive effects on competence. Rok-LS20 associates with the chromosome and preferentially associates with the comK promoter in vitro, indicating that it acts by a similar mechanism as chromosomal Rok. Several Bacillus species
containing a full-length chromosomal \textit{rok} gene also have a second, smaller chromosomal \textit{rok} gene similar to \textit{rok-LS20}. The sizes and genomic positions of the small \textit{rok} genes were less well conserved than for the full-length genes. The small \textit{rok} genes were likely the result of a second horizontal acquisition event, separate from the large \textit{rok} genes. The presence of two \textit{rok} genes may account for the apparent lack of transformability in some of these \textit{Bacillus} strains. It has been suggested that the \textit{rok-LS20}, by repressing competence in the host, could minimize the fitness cost to the host by reducing possible interference between the competence and conjugation machinery or restricting recombination between pLS20 and exogenous DNA (Singh \textit{et al.}, 2012).

\textbf{Rok analogues in other organisms}

The ability to bind A+T-rich DNA, repress gene expression, and silence horizontally acquired elements makes Rok analogous to a class of nucleoid-associated proteins represented by H-NS (heat-stable nucleoid-structuring) from Gram-negative bacteria. The small size, prevalence of basic residues, and use of a C-terminal DNA binding domain in Rok are also characteristics shared by H-NS-like proteins, although Rok is not known to share evolutionary homology with them.

\textbf{H-NS.} H-NS is perhaps the best-known analogue of Rok. H-NS is found in Gram-negative organisms including \textit{E. coli}, \textit{Salmonella}, \textit{Shigella}, \textit{Vibrio}, and other representatives of the alpha-, beta-, and gamma-proteobacteria (Dorman, 2004, 2007; Tendeng and Bertin, 2003). In \textit{E. coli}, H-NS is a 15-kDa protein with high abundance, at approximately 20,000 copies per genome (Dorman \textit{et al.}, 1999; Hulton \textit{et al.}, 1990). Similar to Rok, the N-terminal dimerization domain is connected to the C-terminal DNA-binding domain by a linker region (Dorman \textit{et al.},
The C-terminal domain contains a core DNA-binding motif (TWTG-GR-P) that is highly conserved among proteobacteria. Two arginine residues in the dimerization domain also contribute to DNA binding (Bloch et al., 2003; Dorman, 2004). Biochemical studies indicate that H-NS binds to the major groove of DNA (Tippner and Wagner, 1995; Tippner et al., 1994), although an NMR structure of the monomeric C-terminal domain shows interactions via the minor groove (Sette et al., 2009).

Like Rok, H-NS preferentially associates with A+T-rich DNA and thereby selectively silences horizontally acquired genes. In Salmonella typhimurium, for instance, H-NS represses pathogenicity islands, virulence islets, and plasmid genes. A+T content rather than chromosomal location is sufficient to determine the binding of H-NS (Navarre et al., 2006). Similar results regarding horizontal silencing have been found in E. coli and other species (Dorman, 2007; Kahramanoglu et al., 2011; Oshima et al., 2006). H-NS and Rok likely show different degrees of A+T preference that reflect their host genomes. Regions repressed by H-NS in Salmonella have an average G+C content of 47%, which is low compared to the average G+C content of 52% for the Salmonella genome but higher than the average G+C content of the B. subtilis genome (Navarre et al., 2006). Silencing of horizontally acquired elements by H-NS confers a fitness advantage (Lucchini et al., 2006), and H-NS is thought to act as a “genome sentinel” by minimizing the deleterious effects of newly acquired genetic elements (Dorman, 2007).

H-NS also affects the expression of host genes. In E. coli, over 1,000 genes are regulated by H-NS (Kahramanoglu et al., 2011). Similar to Rok, H-NS almost exclusively acts as a repressor, and H-NS directly represses its own transcription (Free and Dorman, 1995; Ueguchi et al., 1993). Several instances of positive regulation by H-NS are likely indirect effects of repression, although there are rare examples of direct activation (Johansson et al., 1998;
H-NS can prevent transcriptional elongation by forming a protein bridge between the promoter and upstream DNA, trapping RNA polymerase in a DNA loop (Afflerbach et al., 1999; Dame et al., 2001, 2002; Dorman, 2004, 2007; Schröder and Wagner, 2000). At a model promoter in enteropathogenic E. coli, H-NS can form inhibitory contacts with the alpha subunit of RNA polymerase (Shin et al., 2012). H-NS has been shown to co-localize with RNA polymerase in E. coli (Grainger et al., 2010; Oshima et al., 2006) but not in Salmonella typhimurium (Lucchini et al., 2006). There may be species-specific differences in H-NS behavior or multiple ways for H-NS to regulate transcription.

H-NS mutants are highly pleiotropic (Higgins et al., 1988). One notable phenotype is the loss of motility, due to transcriptional effects on adhesion-promoting curli biogenesis genes (Weber et al., 2006), flagellar regulators (Bertin et al., 1994), and other motility genes (Kahramanoglou et al., 2011). Other phenotypes include increased resistance to low pH and high osmolarity, and many H-NS target genes are sensitive to environmental factors (Atlung and Ingmer, 1997; Dorman, 2004; Higgins et al., 1988).

An H-NS consensus binding sequence has been identified from in vitro footprinting: 5’-TCGATAAATT-3’ (Lang et al., 2007), although a “consensus structure” determined by DNA curvature and overall A+T-richness is still thought to be an important determinant of H-NS binding (Dorman, 2004). Consistent with the known oligomerization of HNS, a similar short motif, 5’-AATAAT-3’, is found in 96% of the chromosomal regions bound by H-NS, and the motif occurs on average 20 times in each of these regions (Kahramanoglou et al., 2011). From a ChIP-seq analysis, H-NS binds to 17% of the E. coli genome, in tracts of approximately 1-2 kb (Kahramanoglou et al., 2011).
Furthermore, H-NS contributes to global chromosome architecture. H-NS overproduction is lethal and causes nucleoid compaction, as well as inhibition of transcription and translation (Hulton et al., 1990; Spurio et al., 1992). Super-resolution imaging of H-NS in live *E. coli* shows that it localizes to clusters on the nucleoid, similar to the non-uniform chromosomal distribution of Rok. Formation of these clusters depends on the oligomerization activity of H-NS. Typically, each cell has two H-NS clusters per copy of the chromosome, and the clusters are equally spaced along the long axis of the cell. Chromosome conformation capture shows that H-NS regulated operons are also located in spatial clusters, indicating that H-NS mediates long-range genomic interactions (Wang et al., 2011). I suspect that the clustered localization of Rok on the *B. subtilis* nucleoid may contribute to similar long-range interactions and chromosome structuring.

The DNA-binding properties of H-NS affect its functions in transcriptional repression and chromosome architecture. The sequence of a DNA molecule imparts intrinsic curvature, and H-NS preferentially binds strongly curved DNA (Dame et al., 2001; Yamada et al., 1990). H-NS constrains negative DNA supercoils *in vitro* (Tupper et al., 1994), and it is thought to stabilize compact DNA loops *in vivo* (Dame et al., 2000; Noom et al., 2007). H-NS can bridge separate DNA molecules by forming oligomers, most likely dimers, that each interact with a DNA molecule. Each dimer is predicted to extend over one turn of the DNA helix. There are two models, parallel and antiparallel, for the relative orientation of the hns dimerization domains (Dame et al., 2006). In another model of bridging, the linker domains of H-NS also contribute to long-range interactions (Dorman, 2004). H-NS can bind DNA at high density, and the DNA bridging activity shows cooperative behavior (Dame et al., 2000). H-NS has also been described to stiffen DNA by polymerizing over long tracts of sequence, with approximately one H-NS
dimer per 15-20 bp. The stiffening mode is directly susceptible to changes in temperature, pH, and osmolarity in vitro, suggesting a direct mechanism for these environmental parameters to influence chromosome organization and gene regulation in vivo (Amit et al., 2003; Liu et al., 2010). H-NS can interconvert between stiffening and bridging modes in response to calcium or magnesium concentrations in vitro. Both modes are likely to be physiologically relevant, suggesting that different modes could correspond to different functions of H-NS and that switching between these modes could be regulated (Liu et al., 2010). The biophysics of Rok-DNA binding have not yet been characterized, but at least some of the properties of H-NS may also apply to Rok.

H-NS-like proteins are found in diverse bacteria. Some bacteria have multiple copies of hns-like genes (Tendeng and Bertin, 2003), such as the hns paralogs stpA and hfp in E. coli (Cusick and Belfort, 1998; Dorman, 2004; Muller et al., 2010). Although hns is not essential, at least in the selected organisms studied, the hns gene is retained even in the obligate endosymbiont bacteria that have highly reduced genomes and lack dnaA. There is a single hns gene in the drastically reduced genomes of the intracellular symbionts B. aphidicola and W. glossinidia, suggesting an important and perhaps highly organism-specific role for H-NS (Tendeng and Bertin, 2003).

Sfh. The Shigella flexneri protein Sfh is an H-NS paralog encoded by the plasmid pSf-R27 (Dorman, 2004). Sfh is 59% identical to H-NS, with a similar domain structure, and it shares the preference for A+T-rich curved DNA and the ability to repress virulence genes (Beloin et al., 2003; Dorman, 2004). Sfh is considered a stealth protein because it is thought to allow plasmid entry without perturbing the chromosomal distribution of H-NS. Sfh thereby minimizes the impact of plasmid entry on the host. Introduction of a plasmid lacking sfh leads to
reduced fitness and altered global gene expression, due to titration of chromosomally encoded H-NS. The *sfh* null phenotype is similar to that of *hns* mutants (reduced motility and increased virulence), and it can be complemented by H-NS (Doyle et al., 2007). Given the likely horizontal origin of Rok and its occurrence on a conjugative plasmid, perhaps Rok originally served a similar stealth purpose by minimizing H-NS titration in the host.

**Lsr2.** The first H-NS-like protein to be identified in Gram-positive bacteria was Lsr2 from *Mycobacterium*, a member of the high-G+C actinomycete group. Lsr2 was initially identified as an antigen in leprosy patients infected with *M. leprae* (Laal et al., 1991), and homologues are found in other actinomycetes including *Streptomyces* and *Rhodococcus* (Gordon et al., 2008). Lsr2 is highly conserved in *Mycobacteria* and highly basic like many nucleoid-associated proteins, but it does not have a classical DNA-binding domain (Chen et al., 2006). Although Lsr2 has less than 20% identity with H-NS, Lsr2 is also a dimeric DNA-bridging protein that cooperatively binds DNA to form a rigid DNA-compacting complex (Chen et al., 2006, 2008; Qu et al., 2013). The C-terminal DNA-binding domain of Lsr2 contains an unusual Arg-Gly-Arg structure that inserts into the minor groove of DNA (Gordon et al., 2010). DNA binding by Lsr2 is thought to provide physical protection that may account for the protective effect of Lsr2 against reactive oxygen species (Colangeli et al., 2009).

Lsr2 binds A+T-rich sequences, including virulence genes, in *M. tuberculosis* and *M. smegmatis*. Although the genomic G+C content of *Mycobacterium* species is approximately 66%, Lsr2 prefers regions with G+C content of 47% or lower (Gordon et al., 2010). Despite the evolutionary distance between Lsr2 and H-NS, Lsr2 complements multiple H-NS phenotypes, binds H-NS-regulated genes, and can be negatively regulated by the H-NS antagonist SlyA in *E. coli*, and H-NS complements Lsr2 phenotypes in *M. smegmatis* (Gordon et al., 2008). These
interchangeable functions, despite the negligible homology between H-NS and Lsr2, highlight the role of A+T-richness rather than sequence in determining the DNA-binding activity of H-NS-like proteins. The discovery of Lsr2 suggests that additional cryptic H-NS analogues could be found in other organisms, as illustrated by the characterization of Rok as a nucleoid-associated protein (Smits and Grossman, 2010).

**Diversity of H-NS-like proteins.** Numerous other proteins contain an H-NS-like C-terminal DNA-binding domain fused to a divergent N-terminus. Some examples include BpH3 from *Bordetella pertussis*, KorB from an *E. coli* plasmid, SPB in *Rhodobacter sphaeroides*, XrvA in *Xanthomonas oryzae*, and VicH in *Vibrio cholerae* (Bertin et al., 1999; Dorman et al., 1999; Goyard and Bertin, 1997). Strikingly, the mouse protein Btcd, which binds curved DNA, complements the H-NS motility phenotype in *E. coli* and represses transcription initiation at a model *E. coli* promoter (Timchenko et al., 1996). The diversity, promiscuity, and likely ubiquity of H-NS-like proteins appear to make them versatile regulators of gene expression and gene exchange across bacterial species.
Regulation and interactions of Rok analogues

Not much is known about proteins that directly interact with Rok. Rok regulates promoters that are also bound by other proteins, such as SinR, AbrB, CodY, and ComK at the \textit{comK} promoter (Hahn et al., 1996; Hoa et al., 2002; Smits et al., 2007), so incidental or functionally relevant interactions at these regions might be possible. Numerous plasmids, phages, pathogenicity islands, and mobile genetic elements encode anti-repressor proteins that displace H-NS from promoters (Dorman, 2007; Navarre et al., 2006), so perhaps Rok is also antagonized by horizontally acquired regulators. In \textbf{Chapter 2}, I show that DnaA and Rok interact, and I suggest that DnaA, being widely conserved, may interact with H-NS and other Rok-like analogues in other organisms.

DNA replication and chromosome architecture

The genetic and biochemical interactions between DnaA and Rok suggest a connection between DNA replication and chromosome architecture. There is no known precedent for a direct connection between these processes in \textit{B. subtilis}. There are examples in \textit{E. coli}, in addition to the previously discussed roles of nucleoid-associated proteins in replication initiation.

H-NS has an indirect positive effect on replication initiation in \textit{E. coli}. \textit{hns} mutants have fewer origins per cell during fast growth (one or two origins) compared to wild-type cells (four or eight origins). In wild-type cells, replication re-initiates before termination of the previous round of replication, such that a newly divided cell contains four origins, but in \textit{hns} mutants, the newly divided cell contains two complete chromosomes that each initiate and complete replication only once before division (Atlung and Hansen, 2002). Synchrony is normal in \textit{hns} mutants, and the effect of H-NS is not due to decreased \textit{dnaA} transcription or DnaA protein
levels. In fact, DnaA protein levels are slightly higher, 20%, in an \textit{hns} mutant, and it seems that comparatively more DnaA is needed for initiation in the mutant (Atlung and Hansen, 2002). The effect of \textit{hns} is likely indirect because H-NS does not preferentially bind to the \textit{oriC} region, and because a burst of H-NS synthesis restores initiation after a lag longer than one generation time (Atlung and Hansen, 2002). Overexpressing H-NS restores initiation to wild-type levels but not higher (Atlung and Hansen, 2002), suggesting that the effects of H-NS are subject to additional levels of regulation. This positive role for H-NS in replication initiation is consistent with the observation that lack of \textit{hns} suppresses overreplication by a cold-sensitive \textit{dnaA} mutant (Katayama et al., 1996). It has been suggested that H-NS could affect chromosome segregation or transcription of cell division genes, or that its DNA-binding and remodeling activities could affect replication fork progression (Atlung and Hansen, 2002).

H-NS levels are also modulated in response to DNA replication. Autoregulation of \textit{hns} transcription couples H-NS levels to DNA synthesis, likely using H-NS titration to maintain a relatively constant ratio of H-NS to DNA. Blocking replication initiation or elongation downregulates \textit{hns} transcription. There is a more rapid response for an elongation block than for an initiation block, and the observed time lags are consistent with a mechanism of genome-wide H-NS distribution. It is not timing of the cell cycle but more likely changes in DNA levels that produce this response (Free and Dorman, 1995).

Additionally, H-NS regulates genes that are also targets of DnaA in \textit{E. coli}. H-NS represses the DnaA-regulated \textit{nrdAB} operon and other \textit{rnr} genes, likely accounting for the downregulation of \textit{nrdAB} during anaerobic growth. H-NS regulation is independent of other \textit{nrdAB} regulators, namely Fis and IciA, as well as StpA. H-NS binds to two A+T-rich regions of the \textit{nrdA} promoter and coding sequence. The upstream sequence contains two H-NS consensus
sequences and overlaps with several IciA binding sites. The second sequence, extending into the coding sequence, overlaps with DnaA boxes, suggesting possible joint regulation with DnaA (Cendra et al., 2013). Given the range of genes regulated by H-NS, there are likely other regions of regulatory overlap with DnaA.

**Thesis summary**

I have described a role for Rok in affecting the chromosomal localization of DnaA and its regulator YabA. Using ChIP-seq, I characterized the genome-wide association of DnaA. I found that DnaA associates with two types of regions: (i) those with direct binding of DnaA to DnaA boxes and (ii) those with indirect, Rok-dependent association of DnaA with chromosomal regions bound by Rok. I found that YabA also depends on Rok to associate with Rok-bound chromosomal regions. At these regions, DnaA and YabA do not require each other for association with Rok. At regions of direct DnaA binding, however, association of YabA depends on DnaA. Using purified proteins, I found that DnaA depends on Rok for association with a DNA probe representing one of the Rok-dependent chromosomal regions. This *in vitro* result recapitulates the Rok-dependent association of DnaA observed *in vivo*.

DnaA and Rok regulate some of the same genes, and I analyzed global gene expression to determine the genetic relationship between these two transcriptional regulators. For several regions that showed Rok-dependent DnaA association by ChIP-seq, both DnaA and Rok contributed to repression of gene expression. The effect of *rok* was generally epistatic. That is, Rok most likely represses these genes directly, whereas DnaA depends on Rok and likely represses expression indirectly, by enhancing the activity of Rok. DnaA appears positioned to influence Rok transcriptional activity via protein-protein interaction, as the proteins interact *in*
vitro and are associated with the same sequences in vivo. The complete picture of gene regulation at these regions appears complex, however, as previous work indicates that indirect regulation by Sda, downstream of DnaA, also contributes to gene expression at these regions.

The genetic and biochemical connections between DnaA and Rok in B. subtilis could potentially be representative of interactions between DnaA and the numerous Rok analogues in other organisms. Such interactions might represent a regulatory connection between DNA replication and chromosome architecture. I have demonstrated an unexpected role for the highly conserved replication initiator protein. Further understanding of factors that regulate DnaA-mediated replication initiation and gene expression is likely to reveal additional relationships between DNA replication and other cellular processes.
Chapter 2

Genetic and biochemical interactions between the bacterial replication initiator DnaA and the nucleoid-associated protein Rok in *Bacillus subtilis*

Charlotte A. Seid and Alan D. Grossman

This chapter is being prepared for publication.
Abstract

We identified genetic and biochemical interactions between the DNA replication initiator DnaA and the nucleoid-associated protein Rok of *Bacillus subtilis*. Using ChIP-seq, we found that DnaA was associated with many of the same chromosomal regions as Rok. Association of DnaA with the Rok-bound chromosomal regions was dependent on Rok. In contrast, the binding of DnaA to chromosomal regions with clusters of DnaA boxes, such as the origin of replication, was independent of Rok. The DNA binding domain of DnaA was dispensable for association of DnaA with Rok-bound chromosomal regions. Furthermore, DnaA and Rok interact directly in an *in vitro* electrophoretic mobility shift assay. We found that several genes, including the sublancin antibiotic gene cluster, were co-regulated by DnaA and Rok. The effects of DnaA on expression of these genes were dependent on Rok; that is, *rok* was epistatic to *dnaA*. The interaction between these transcriptional regulators may represent a regulatory connection between DNA replication and chromosome architecture.

Introduction

DnaA is the conserved DNA replication initiator in bacteria and a transcription factor (Katayama et al., 2010; Leonard and Grimwade, 2010, 2011; Messer, 2002; Messer and Weigel, 1997; Mott and Berger, 2007). DnaA is a AAA+ ATPase (Erzberger et al., 2002; Kawakami and Katayama, 2010; Snider et al., 2008) that binds ATP or ADP (Erzberger et al., 2006; Kurokawa et al., 1999; Nishida et al., 2002; Sekimizu et al., 1987; Speck et al., 1999). It is active for replication in the ATP-bound form (Crooke et al., 1993; Duderstadt et al., 2011; Erzberger et al., 2006; Fuller et al., 1984; Funnell et al., 1987; McGarry et al., 2004; Sekimizu et al., 1987). The N-terminal domain and AAA+ domain of DnaA contribute to oligomerization (Erzberger et al.,
2006; Kawakami and Katayama, 2010; Leonard and Grimwade, 2010; Messer, 2002). The C-terminal domain is necessary and sufficient for DNA binding (Krause et al., 1997; Roth and Messer, 1995).

DnaA binds to a 9-bp motif, the DnaA box, which occurs multiple times at the origin of replication, and elsewhere around the chromosome. In *Bacillus subtilis*, DnaA regulates the expression of many genes, likely by binding to promoter regions (Burkholder et al., 2001; Goranov et al., 2005; Ogura et al., 2001). DnaA also binds to several chromosomal regions other than the origin of replication (Breier and Grossman, 2009; Goranov et al., 2005; Ishikawa et al., 2007). DnaA is the target of diverse regulatory systems that modulate replication initiation in response to various cellular conditions (Katayama et al., 2010; Leonard and Grimwade, 2011; Mott and Berger, 2007; Skarstad and Katayama, 2013). Regulation of the transcriptional activity and chromosomal localization of DnaA have been less extensively studied.

We recently characterized the binding of DnaA to chromosomal regions *in vitro* (J. L. Smith, ADG, submitted). This work led to the recognition that there were some genomic regions associated with DnaA *in vivo* that were not bound *in vitro*, indicating that other proteins might influence DnaA DNA binding *in vivo*. One of these candidates was the nucleoid-associated protein Rok (J. L. Smith, ADG, submitted).

To define the genomic regions associated with DnaA *in vivo*, we used chromatin immunoprecipitation and deep sequencing (ChIP-seq). We found that, in addition to associating with previously identified chromosomal regions, DnaA associated with many additional regions known to be bound by the nucleoid-associated protein Rok (Albano et al., 2005; Smits and Grossman, 2010). Rok was originally identified as a repressor of *comK* and consequently a negative regulator of competence development in *B. subtilis* (Hoa et al., 2002). Rok also
regulates expression of other genes (Albano et al., 2005; Kovács and Kuipers, 2011; Marciniak et al., 2012) and binds directly to several promoters \textit{in vivo} and \textit{in vitro} (Albano et al., 2005; Smits and Grossman, 2010). In contrast to DnaA, Rok is not known to have a well-defined binding site (Albano et al., 2005; Smits and Grossman, 2010). Rok is a nucleoid-associated protein that binds A+T-rich regions throughout the chromosome and has a role in silencing some regions of horizontally acquired DNA (Smits and Grossman, 2010). In these ways, Rok is analogous, but not homologous, to H-NS from \textit{E. coli} and other gamma-proteobacteria (Dorman, 2007, 2010; Navarre et al., 2006) and to Lsr2 from \textit{Mycobacteria} (Gordon et al., 2008, 2010).

We further investigated the relationship between DnaA and Rok and found that DnaA required Rok for association with several chromosomal regions. Furthermore, the DNA-binding domain of DnaA was not necessary for DnaA association with the Rok-bound regions, in contrast to regions where DnaA binds directly to DnaA boxes. We found a direct interaction between DnaA and Rok \textit{in vitro} using an electrophoretic mobility shift assay. We also identified several genes that were indirectly regulated by DnaA and likely directly regulated by Rok. The interaction between DnaA and Rok may contribute to the co-regulation of gene expression. Our results indicate that the association of DnaA with chromosomal regions extends beyond those regions bound directly by DnaA, and that the regulation and function of DnaA and Rok may be more complex than previously thought.
Results and Discussion

Overview of ChIP-seq analysis of DnaA and Rok

We examined the chromosome-wide association of DnaA and Rok at high resolution using ChIP-seq of \textit{B. subtilis} cells growing exponentially in defined minimal glucose medium. We used the peak-calling algorithm SISSRs (Jothi et al., 2008) and a five-fold enrichment cutoff (Experimental Procedures) to define chromosomal regions that were associated with DnaA or Rok. Antibody specificity in the immunoprecipitations was verified by performing analogous ChIP-seq experiments with anti-DnaA and anti-Rok antibodies in \textit{dnaA} and \textit{rok} null mutants (Experimental Procedures). The anti-DnaA and anti-Rok antibodies were highly specific. That is, there was little or no specific DNA in the immunoprecipitates from the relevant null mutants.

ChIP-seq analysis of genome-wide binding by DnaA

DnaA was detectably associated, either directly or indirectly, with 44 chromosomal regions in the ChIP-seq experiments (Fig. 1A; Table 1). These regions included the eight DnaA binding regions that were previously identified by ChIP-chip (Breier and Grossman, 2009) and ChAP-chip (Ishikawa et al., 2007) (Fig. 2). These regions all contain clusters of DnaA boxes ("DnaA box clusters" or "DBCs" in (Ishikawa et al., 2007)) and are located in the intergenic regions between: \textit{rpmH/dnaA} (the \textit{dnaA} promoter region), \textit{dnaA/dnaN} (directly upstream of the DNA unwinding element in \textit{oriC}), \textit{gcp/ydiF}, \textit{yqeG/sda}, \textit{ywlc/ywlB}, \textit{ywcl/vpr}, \textit{yydA/yycS}, and \textit{trmE/jag}. Each of these regions contains at least four DnaA binding motifs. These are the strongest binding regions identified by \textit{in vitro} DNA affinity purification of DnaA (J. L. Smith, ADG, submitted).
Figure 1. Genome-wide binding of DnaA and Rok in wild-type cells.
Wild-type cells (strain AG174) were grown to mid-exponential phase, and DnaA (A) and Rok (B) were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC is set as 0. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments. Asterisks and black labels in (A) indicate previously identified DnaA binding sites (Breier and Grossman, 2009; Ishikawa et al., 2007). Dashed lines and red labels in (B) indicate selected previously identified Rok binding sites (Albano et al., 2005; Smits and Grossman, 2010).
<table>
<thead>
<tr>
<th>Nearest gene(s)</th>
<th>Annotated function(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dnaA *</td>
<td>chromosome replication initiator</td>
</tr>
<tr>
<td>dnaN *</td>
<td>DNA polymerase III subunit beta</td>
</tr>
<tr>
<td>yceB/C</td>
<td>monooxygenase; stress adaptation protein</td>
</tr>
<tr>
<td>dtpT/yclG</td>
<td>di-tripeptide-proton ABC symporter; uronase</td>
</tr>
<tr>
<td>yddM</td>
<td>helicase in ICEBs1</td>
</tr>
<tr>
<td>gcp/ydiF *</td>
<td>DNA-binding/iron metalloprotein/AP endonuclease; ABC transporter ATP binding protein</td>
</tr>
<tr>
<td>pspA</td>
<td>phage shock protein A</td>
</tr>
<tr>
<td>ssuB/katA</td>
<td>aliphatic sulfonate ABC transporter ATP-binding protein, catalase</td>
</tr>
<tr>
<td>comK</td>
<td>competence transcription factor</td>
</tr>
<tr>
<td>appD</td>
<td>oligopeptide ABC transporter ATP-binding protein</td>
</tr>
<tr>
<td>yjcM/N</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>rok</td>
<td>repressor of comK</td>
</tr>
<tr>
<td>ppsA</td>
<td>plipastatin synthetase (antifungal peptide)</td>
</tr>
<tr>
<td>yobI cds</td>
<td>NTPase with transmembrane helices</td>
</tr>
<tr>
<td>yobI cds</td>
<td>NTPase with transmembrane helices</td>
</tr>
<tr>
<td>yosX</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yonX</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>sunT cds</td>
<td>lantibiotic precursor peptide</td>
</tr>
<tr>
<td>sunA</td>
<td>lantibiotic precursor peptide</td>
</tr>
<tr>
<td>yqgA/B</td>
<td>hypothetical protein; factor involved in motility</td>
</tr>
<tr>
<td>sda *</td>
<td>sporulation checkpoint protein</td>
</tr>
<tr>
<td>yqxI</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yraN/M</td>
<td>LysR transcriptional regulator, hypothetical protein</td>
</tr>
<tr>
<td>araA</td>
<td>L-arabinose isomerase</td>
</tr>
<tr>
<td>argG</td>
<td>arginosuccinate synthase</td>
</tr>
<tr>
<td>iscS/braB</td>
<td>cysteine desulfurase; branched chain amino acid/Na+ symporter</td>
</tr>
<tr>
<td>nupN</td>
<td>lipoprotein</td>
</tr>
<tr>
<td>yuzB/yutJ</td>
<td>hypothetical protein, NADH dehydrogenase</td>
</tr>
<tr>
<td>lytA/tagU</td>
<td>membrane-bound lipoprotein; membrane-bound transcriptional regulator</td>
</tr>
<tr>
<td>ggaB cds</td>
<td>poly(glucosyl NAG 1-phosphate) glucosyltransferase</td>
</tr>
<tr>
<td>ggaA</td>
<td>poly(glucosyl NAG 1-phosphate) glucosyltransferase</td>
</tr>
<tr>
<td>glyA</td>
<td>serine hydroxymethyltransferase</td>
</tr>
<tr>
<td>ywlC *</td>
<td>ribosome maturation factor; RNA binding protein</td>
</tr>
<tr>
<td>sboA</td>
<td>subtilosin (and associated operon including transporter)</td>
</tr>
<tr>
<td>ywcI/vpr *</td>
<td>hypothetical protein; serine protease</td>
</tr>
<tr>
<td>yxkC/D</td>
<td>hypothetical protein, efflux transporter</td>
</tr>
<tr>
<td>yxaI/J</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>ahpC</td>
<td>alkyl hydroperoxide reductase</td>
</tr>
<tr>
<td>yydH cds</td>
<td>hypothetical protein</td>
</tr>
</tbody>
</table>
yydD cds  hypothetical protein
yydA *  rRNA large subunit methyltransferase
yybN  hypothetical protein
yybM cds  hypothetical protein
trmE *  tRNA modification GTPase

<table>
<thead>
<tr>
<th>Table 1. Chromosomal regions associated with DnaA.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions are listed in order of chromosomal position. Association within a coding sequence is indicated by “cds.” Asterisks indicate DnaA box cluster regions previously known to be bound by DnaA (Breier and Grossman, 2009; Ishikawa et al., 2007).</td>
</tr>
</tbody>
</table>

**Figure 2. Association of DnaA and Rok with DnaA box cluster chromosomal regions in wild-type cells.**

DnaA and Rok were immunoprecipitated from wild-type cells (strain AG174). Each panel shows a magnified view of the data from Fig. 1. The eight DnaA box cluster regions include (A) rpmH/dnaA, (B) dnaA/dnaN, (C) gcp/ydiF, (D) yqeG/sda, (E) ywlC/ywlB, (F) ywcl/vpr, (G) yydA/yycS, and (H) trmE/jag (Breier and Grossman, 2009; Ishikawa et al., 2007). ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. DnaA is plotted on a larger y-axis scale in panels (A) and (B). The position of oriC is set as 0. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions. DnaA boxes were defined by a position-specific site matrix (J. L. Smith, ADG, submitted). DnaA boxes on the plus and minus strands are plotted as right-facing and left-facing arrows, respectively.
In addition to the eight DnaA box cluster regions, we found that DnaA was associated with 36 other chromosomal regions (Table 1; Fig. 1A). Many of these regions were previously found to be associated with Rok in ChIP-chip experiments (Smits and Grossman, 2010), and all were found to be associated with Rok in ChIP-seq experiments (see below). DnaA had not been previously detected at these 36 regions, except for sunA, most likely due to limited sensitivity.

In a previous genome-wide study of DnaA localization, DnaA was difficult to detect by ChIP-chip at regions other than the DnaA box clusters (Breier and Grossman, 2009). Use of ChIP-PCR enabled detection of DnaA association at additional regions (Breier and Grossman, 2009), including sunA, indicating that more sensitive assays could reveal additional DnaA-bound chromosomal regions.

The enhanced sensitivity from the PCR amplification step of ChIP-seq likely enabled us to detect DnaA at more chromosomal regions than detected by ChIP-chip. Association of DnaA with these additional regions is unlikely to be an artifact of ChIP-seq sample amplification because these regions were not enriched in a control non-immunoprecipitated sample or in samples from a dnaA null mutant or a rok null mutant (see below). Additionally, there were several regions where DnaA was previously detected by ChIP-PCR (Breier and Grossman, 2009; Goranov et al., 2005) but that were not detected in the ChIP-seq results presented here. The strains and growth conditions were virtually identical among these experiments and the ChIP-seq experiments, so the differences in detection may be due to the use of different anti-DnaA polyclonal antibodies (chicken vs. rabbit), the use of a secondary antibody in the previous studies, and/or the greater sensitivity of ChIP-PCR.

Some of the 36 regions that were associated with both Rok and DnaA contained one or more recognizable DnaA box(es) (Fig. 3), indicating that DnaA could be binding directly to
these sites. For example, in the sunA, sboA, and yxaI regions, binding by DnaA was roughly centered on a DnaA box (Fig. 3A-C), similar to binding of DnaA to the eight DnaA box cluster regions (Fig. 2). In contrast, association of DnaA with several of the other regions (e.g., yuzB/yutJ, yxkC/yxkD, and lytA) was not centered on a putative DnaA box (Fig. 3D-H), indicating that association of DnaA with these regions might be dependent on another DNA-binding protein (see below).

Figure 3. Association of DnaA and Rok with Rok-bound chromosomal regions in wild-type cells.
DnaA and Rok were immunoprecipitated from wild-type cells (strain AG174). Each panel shows a magnified view of the data from Fig. 1. Eight selected chromosomal regions include (A) sunA, (B) sboA, (C) yxaI/yxaJ, (D) lytA/tagU, (E) yuzB/yutJ, (F) yxkC/D, (G) yybN, and (H) yydD. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. DnaA and Rok are plotted on a larger y-axis scale in panels (A) and (B). The position of oriC is set as 0. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions.
**ChIP-seq analysis of genome-wide binding by Rok**

As a nucleoid-associated protein, Rok is known to associate *in vivo* with many places on the *B. subtilis* genome, particularly A+T-rich regions (Smits and Grossman, 2010). In ChIP-seq experiments, we identified 264 chromosomal regions bound by Rok (**Fig. 1B**). This ChIP-seq profile agreed well with previous ChIP-chip results, including the strong correlation between Rok binding and A+T content (Smits and Grossman, 2010). Many examples of Rok binding occurred as a single peak upstream of a gene (**Fig. 3A-F**), although at several regions (e.g., *yybN*, *yydD*), there were multiple overlapping Rok peaks within a coding sequence (**Fig. 3G-H**). Many of the regions bound by Rok were in or upstream of operons that are regulated by Rok, including *sboA*, *sunA*, *yybN*, *yydH*, and *yxaJ* (this study and (Albano et al., 2005)).

Some Rok-bound regions were also associated with DnaA. In addition to the 36 regions mentioned above, we detected Rok at two regions that are known to be associated with DnaA and contain multiple DnaA binding motifs: the *trmE/jag* intergenic region and the *dnaA/dnaN* intergenic region of oriC (**Fig. 2B, H**). Given the significance of these DnaA-binding regions in replication (Ishikawa et al., 2007), we investigated, but did not detect any effect of a *rok* null mutation on DNA replication (see below and **Appendix A**).

We used the Rok-associated regions to examine the sequence specificity of Rok binding. Previous studies, including transcriptional profiling and *in vitro* DNA binding (Albano et al., 2005; Hoa et al., 2002), and ChIP-chip analyses (Smits and Grossman, 2010), did not indicate a strong binding motif, possibly due to limited datasets, although a few A+T-rich sequences appeared overrepresented in Rok-bound regions (Smits and Grossman, 2010). We extracted sequences corresponding to the centers of Rok ChIP-seq binding peaks (Experimental Procedures) and searched for a motif using the ChIP-seq motif-finding tool DREME (Bailey,
2011). We identified an A+T-rich potential binding motif (Fig. 4), but since the sequence consists almost entirely of As, the motif may simply reflect the general preference of Rok for A+T-rich regions. This motif closely resembles the center of a previously identified 19-bp A+T-rich motif, which was derived from only four Rok-bound sequences and two negative control regions (Smits and Grossman, 2010). The only other motif-finding result was the Shine-Dalgarno sequence, which likely simply reflects the general tendency of Rok to bind upstream of coding sequences.

![Sequence logo of a potential Rok binding motif](LOGO.png)

**Figure 4. Sequence logo of a potential Rok binding motif.**
This motif, generated by the online motif-finding tool DREME (http://meme.nbcr.net/meme/cgi-bin/meme.cgi), had a P-value of 2.1e-8 and an E-value of 3.6e-4. The motif was found in 200 of 264 input sequences.

At each chromosomal region bound by Rok and DnaA, the relative amount of DNA that was isolated in the Rok immunoprecipitates was closely correlated with that from the DnaA immunoprecipitates (Fig. 5). That is, the regions with the strongest Rok enrichment were those with the strongest DnaA enrichment, whereas the regions with weaker Rok enrichment were those with weaker or no DnaA enrichment. The higher enrichment of Rok relative to DnaA at each region is consistent with our estimates that Rok is approximately ten-fold more abundant than DnaA in the cell (Chapter 3).
Figure 5. Relative enrichment of DnaA and Rok at co-associated regions in wild-type cells. Each data point represents one chromosomal region. Enrichment values were determined using the peak-calling algorithm SISSRs. Selected regions are labeled. The minimum enrichment for inclusion in this dataset was three-fold.

We observed many regions bound by Rok but not DnaA. DnaA might associate with only a subset of Rok-bound regions. Alternatively, DnaA might be present at all regions bound by Rok, but below the limit of detection in these experiments. Based on the close correlation of DnaA and Rok enrichment levels (Fig. 5) and the relatively low Rok enrichment of regions without detectable DnaA (median 5.4-fold enrichment, compared to median 30.5-fold enrichment for regions with detectable DnaA), we suspect that DnaA is present at most if not all regions bound by Rok, but below the limit of detection at some regions.

We considered several possibilities to explain the association of DnaA and Rok with several of the same chromosomal regions. 1) Since the ChIP-seq analyses were done on a
population of cells, different subpopulations could have one or another of the proteins associated with a given region. In this way, each protein could be associated with the same region, but not necessarily in the same cells. 2) Since Rok is associated with many chromosomal regions, and there are many potential DnaA boxes throughout the genome, the association of these proteins could simply be a coincidence of independent binding to the same chromosomal regions. 3) One of these proteins could depend on the other for association. Using rok or dnaA null mutants, we found that Rok was required for the association of DnaA with Rok-associated regions but not for association of DnaA with DnaA box regions, supporting the latter model.

**ChIP-seq analysis of genome-wide binding in a rok null mutant**

We found that Rok was required for association of DnaA with some chromosomal regions, but not with others. We analyzed the genome-wide binding of DnaA in a rok null mutant using ChIP-seq. There was little or no effect of rok at the eight DnaA box cluster regions (Fig. 6A, B). That is, association of DnaA with these regions was not significantly altered in the rok null mutant compared to the rok+ strain. This result was expected since Rok was not detectably associated with most of these regions.

In contrast, Rok was required for association of DnaA at the other chromosomal regions. In the rok null mutant, we did not detect association of DnaA at the 36 regions that showed co-association with Rok in wild-type cells (Fig. 6A, C). No additional DnaA-bound regions were detected in the rok null mutant. We confirmed that DnaA protein levels were similar between the wild-type and rok null strains (Chapter 3), indicating that the changes in binding did not reflect changes in the amount of DnaA.
Figure 6. Genome-wide binding of DnaA in wild-type cells and a rok null mutant.
Wild-type cells (AG174) and a rok null mutant (WKS1038) were grown to mid-exponential phase, and DnaA was immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC is set as 0. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments. (A) Genome-wide binding of DnaA in wild-type cells (upper y-axis, black) and a rok null mutant (lower y-axis, red). Asterisks indicate the eight previously identified DnaA binding sites. Association of DnaA with the eight DnaA box cluster regions (B) and selected Rok-bound regions (C) in wild-type cells (solid lines) and a rok null mutant (dashed lines). Each panel shows a magnified and superimposed view of the data from (A). The selected Rok-bound chromosomal regions include sunA, sboA, yuzB/yutJ, and yybN. DnaA boxes (upper row) and gene annotations (lower row) are shown below the corresponding chromosomal positions.
The Rok-dependent binding of DnaA to certain chromosomal regions indicated that both proteins were associated with the target regions in the same cells and not in separate subpopulations of cells. The chromosomal regions with Rok-dependent DnaA binding demonstrate that DnaA association is not limited to chromosomal regions containing DnaA-binding sites. The co-association of these proteins with the same chromosomal regions suggested a direct interaction between Rok and DnaA, which we subsequently demonstrated in vitro (see below).

**ChIP-seq analysis of genome-wide binding in a dnaA null mutant**

We found that loss of DnaA did not detectably alter the genome-wide association of Rok. We analyzed the genome-wide binding of Rok, using ChIP-seq, in a dnaA null mutant. dnaA is usually essential because of its role in replication initiation. To study gene expression in a dnaA null mutant, we used a strain that initiates replication from a heterologous origin (oriN) using its cognate initiator (RepN, encoded within oriN). In this strain, dnaA is no longer essential (Hassan et al., 1997; Moriya et al., 1997). We used a strain in which oriN was inserted near oriC, and the oriC region, including dnaA and dnaN, was removed. dnaN encodes the processivity clamp of DNA polymerase and is essential for growth. dnaN was expressed from a xylose-inducible promoter (Pxyl-dnaN), integrated at another region of the chromosome, as described in (Merrikh and Grossman, 2011). We compared this dnaA null mutant to an isogenic strain expressing dnaA from an IPTG-inducible promoter (Pspank-dnaA) integrated into the chromosome at an ectopic site (Experimental Procedures). The genome-wide binding of DnaA was similar between the dnaA+ control strain and the wild-type strain, except at the oriC region, which was deleted (Fig. 7A).
Figure 7. Genome-wide binding of DnaA and Rok in dnaA+ and dnaA null mutant cells. dnaA+ cells (strain TAW5) and a dnaA null mutant (strain AIG200) were grown to mid-exponential phase, and DnaA (A) and Rok (B) were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC in the wild-type strain is set as 0, but oriC was deleted from strains TAW5 and AIG200. These strains contained a deletion of dnaA and dnaN. These strains replicated from the heterologous origin oriN and expressed dnaN from Pxy1-dnaN. In strain TAW5, dnaA was expressed from Pspank-dnaA with 0.1 mM IPTG. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.

(A) Genome-wide binding of DnaA in the wild-type oriC+ strain (AG174, upper y-axis, black) and the oriN+ dnaA+ strain (TAW5, lower y-axis, red).

(B) Genome-wide binding of Rok in the dnaA+ strain (upper y-axis, black) and the dnaA null mutant (lower y-axis, red).
Loss of DnaA had little or no effect on association of Rok with chromosomal regions. In both the dnaA+ and dnaA null mutant strains, Rok was still associated with the same chromosomal regions as in the wild-type strain (Fig. 7B). No additional Rok-binding regions were identified in the dnaA null mutant.

**Comparison of ChIP-seq binding profiles of DnaA and Rok**

We identified two categories of chromosomal regions bound by DnaA: (1) regions with DnaA bound to clusters of DnaA binding sites and (2) regions bound by DnaA and Rok, at which association of DnaA depended on Rok. At the 36 chromosomal regions where DnaA and Rok co-associated in a Rok-dependent manner, the binding peaks were closely coincident (representative regions shown in Fig. 8). These observations, together with the dependence results and the known *in vitro* binding of Rok to some of these regions (Albano et al., 2005), indicated that Rok was directly bound to DNA at these regions, whereas DnaA likely associated with these regions indirectly, via Rok.
Figure 8. ChIP-seq reads for selected regions bound by DnaA and Rok in wild-type cells. DnaA and Rok were immunoprecipitated from wild-type cells (strain AG174). Each panel shows a magnified view of the data from Fig. 1. The selected chromosomal regions include (A) sunA, (B) sboA, and (C) yuzB/yutJ. Numbers of reads are shown on the y-axis. Reads mapping to the forward strand are shown on the upper y-axis in blue, and reads mapping to the reverse strand are shown on the lower y-axis in green. Chromosomal position is shown on the x-axis. The ChIP-seq point of symmetry indicates the inferred center of the binding site (Belitsky and Sonenshein, 2013). The center of each DnaA binding peak, as determined by SISSRs, is indicated by a bold vertical line. Positions of DnaA boxes are indicated by dashed red vertical lines and red points on each x-axis. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions.

Rok and DnaA associated with two DnaA box cluster regions, dnaA/dnaN and trmE/jag, independently of each other. In contrast to the Rok-dependent DnaA association with other chromosomal regions, association of DnaA did not depend on Rok (Fig. 9A), nor did association of Rok depend on DnaA (Fig. 9B) at these two regions. This mutual independence of DnaA and Rok binding allowed for the possibility that different subpopulations of cells could have either DnaA or Rok associated with these regions.
At the dnaA/dnaN intergenic region, the DnaA and Rok binding peaks were separated by 40 nucleotides, indicating that the two proteins bound to different sequences within this region (possibly also in different subpopulations) (Fig. 9C). The DnaA and Rok binding peaks were not detectably separated at the trmE/jag region (Fig. 9D). The separate binding peaks at dnaA/dnaN and the lack of an observed effect of Rok on replication (see below) indicated that the presence of Rok at the DUE reflected the protein’s general preference for A+T-rich regions, rather than a specific replication-related function.

Both DnaA (Erzberger et al., 2006; Fujikawa et al., 2003) and Rok (Smits et al., 2007) bind to the major groove of DNA. It is therefore unlikely that DnaA and Rok can bind to the same sequence via opposite faces of the DNA helix, as is the case for Rok and the minor-groove binding protein ComK (Smits et al., 2007). Consequently, we interpreted the coincidence between DnaA and Rok ChIP-seq peaks as evidence that at most one protein (i.e., Rok) directly interacted with DNA. Using a DnaA mutant, we found that the ability of DnaA to bind DNA was, in fact, dispensable for association with Rok-bound chromosomal regions (see below).
Figure 9. Mutually independent association of DnaA and Rok with the dnaA/dnaN and trmE/jag chromosomal regions.

DnaA and Rok were immunoprecipitated from the indicated strains. Each panel contains a magnified view of the wild-type data from Fig. 1 at the dnaA/dnaN region or the trmE/jag region. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC is set as 0. DnaA boxes (upper row) and gene annotations (lower row) are shown below the corresponding chromosomal positions. Data represent single biological experiments.

(A) DnaA does not depend on Rok at dnaA/dnaN or trmE/jag. DnaA was immunoprecipitated from wild-type cells (AG174, solid lines) and a rok null mutant (WKS1038, dashed lines). Data for the rok null mutant represent a magnified view of the data in Fig. 6.

(B) Rok does not depend on DnaA at trmE/jag. Rok was immunoprecipitated from wild-type cells (AG174, solid lines) and a dnaA null mutant (AIG200, dashed line). Strain AIG200 contains a deletion of oriC, including dnaA and dnaN, so Rok association with the dnaA/dnaN chromosomal region could not be measured in this strain.

ChIP-seq reads for association of DnaA and Rok with the dnaA/dnaN (C) and trmE/jag (D) chromosomal regions in wild-type cells. Numbers of reads are shown on the y-axis. Reads mapping to the forward strand are shown on the upper y-axis in blue, and reads mapping to the reverse strand are shown on the lower y-axis in green. The ChIP-seq point of symmetry indicates the inferred center of the binding site (Belitsky and Sonenshein, 2013). The center of the DnaA binding peak, as determined by SISSRs, is indicated by a solid vertical line in the DnaA plots and a dashed vertical line in the Rok plots. The center of the Rok binding peak is indicated by a solid vertical line in the Rok plots and a dashed vertical line in the DnaA plots.
ChIP-seq analysis of genome-wide binding in a dnaA DNA-binding mutant

To determine if DnaA was bound directly to DNA at the Rok-associated chromosomal regions, we used a DnaA mutant that was missing the DNA binding domain (DnaAΔC). We analyzed the genome-wide binding of the mutant DnaA and Rok using ChIP-seq. We deleted the C-terminal 91 amino acids of DnaA, corresponding to the conserved Domain IV region that is necessary and sufficient for DNA binding in E. coli (Roth and Messer, 1995). The mutant dnaA (dnaAΔC) was expressed from an IPTG-inducible promoter (Pspank) integrated into the chromosome at an ectopic site. A congenic control strain expressed full-length dnaA instead. We used an oriN+ strain in which the oriC region, including the native dnaA, was removed, and dnaN was constitutively expressed at its native locus from a derivative of Ppen (Experimental Procedures).

We found that the DNA-binding domain of DnaA was not required for association of DnaA with Rok-bound chromosomal regions. DnaAΔC was associated with many (32) of the regions bound by Rok (Fig. 10A, C). In contrast, DnaAΔC was not associated with any DnaA box cluster regions (Fig. 10B), as expected for regions of direct DnaA binding. The genome-wide binding profile of Rok was similar between the dnaA+ and dnaAΔC strains (data not shown), consistent with the dnaA null results.
Figure 10. Genome-wide binding of DnaA and Rok in cells expressing wild type dnaA or a DNA-binding mutant.

Cells expressing wild-type dnaA (CAS221) or mutant dnaAΔC (CAS231) were grown to mid-exponential phase, and DnaA and Rok were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC in the wild-type strain is set as 0, but oriC is deleted from these strains. Strains contained a deletion of dnaA from its native locus. These strains replicated from the heterologous origin oriN and expressed dnaN from Ppen2028-dnaN. In both strains, dnaA or dnaAΔC was ectopically expressed from Pspank with 0.1 mM IPTG. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.

(A) Genome-wide binding of DnaAΔC (upper y-axis, black) and Rok (lower y-axis, red) in cells expressing dnaAΔC.

Association of wild-type DnaA (solid lines) and DnaAΔC (dashed lines) with DnaA box cluster regions (B) and selected Rok-bound regions (C). Each panel shows a magnified and superimposed view of the data from (A). Two DnaA box cluster regions (rpmH/dnaA and dnaA/dnaN) are deleted from the strains used in these experiments. The selected Rok-bound regions in (C) are the same as those shown in Fig. 6C. DnaA boxes (upper row) and gene annotations (lower row) are shown below the corresponding chromosomal positions.
There appeared to be an increase in the amount of DNA associated with DnaAΔC compared to full-length DnaA at each of the Rok-bound regions (Fig. 10C). This increase in signal was not due to higher DnaA protein levels in the dnaAΔC mutant (Fig. 11). Additionally, DnaAΔC was detectably associated with a greater number of Rok-bound regions than full-length DnaA expressed in the same strain background (32 regions vs. 13 regions, using a cutoff of five-fold enrichment). The reduced number of regions associated with full-length DnaA in this oriN, IPTG-inducible strain (13 regions), relative to the wild-type strain (36 regions), may be due to the reduction in full-length DnaA protein levels in the oriN background (Fig. 11). The increased signal for DnaAΔC relative to full-length DnaA at Rok-bound regions might reflect differences in antibody recognition, or there might actually be higher enrichment of DnaAΔC. If there is higher enrichment of DnaAΔC relative to full-length DnaA, it might be due to a global redistribution of DnaA that would otherwise associate with the DnaA box-cluster regions. The observed enrichment of DnaA at certain Rok-bound sites in the dnaAΔC mutant but not the dnaA+ strain is consistent with the notion that wild type DnaA is present at most or all Rok-bound regions but is below the limit of detection at some of these regions.

Figure 11. Relative DnaA protein levels in mutant strains.
Wild-type cells (strain AG174, oriC+), a strain expressing full-length dnaA (CAS221), and a strain expressing dnaAΔC (CAS231) were grown to mid-exponential phase at 30 °C in minimal glucose medium. Relative DnaA protein levels were measured by quantitative western blot, corrected for differences in molecular weight, and normalized to OD (Experimental Procedures). Data are averages of at least three biological replicates, and error bars represent the standard error.
The dispensability of the DnaA C-terminus for association with Rok-bound regions indicated that DnaA did not associate with these regions via canonical DNA binding. Furthermore, the binding peaks of full-length DnaA and DnaAΔC were coincident (representative regions shown in Fig. 12), indicating that DnaA-DNA contacts were not necessary for association of even wild-type DnaA with these regions, and that DnaA likely associates indirectly with these DNA regions via direct protein-protein interactions with Rok.

Figure 12. ChIP-seq reads for selected regions bound by DnaA and DnaAΔC. DnaA was immunoprecipitated from strains wild-type dnaA (CAS221) or mutant dnaAΔC (CAS231). Each panel shows a magnified view of the data from Fig. 10. The selected chromosomal regions include (A) sunA, (B) sboA, and (C) yuzB/yutJ. Numbers of reads are shown on the y-axis. Reads mapping to the forward strand are shown on the upper y-axis in blue, and reads mapping to the reverse strand are shown on the lower y-axis in green. Chromosomal position is shown on the x-axis. The ChIP-seq point of symmetry indicates the inferred center of the binding site (Belitsky and Sonenshein, 2013). The center of each DnaA binding peak, as determined by SISSRs, is indicated by a bold vertical line. Positions of DnaA boxes are indicated by dashed red vertical lines and red points on each x-axis. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions.
DnaA and Rok interact *in vitro*

We found that DnaA and Rok interacted *in vitro* on a DNA fragment. We purified Rok and the DnaA DNA-binding mutant (DnaAΔC) and measured their effects on the mobility of a DNA fragment during gel electrophoresis. The DNA fragment contained the *rok* promoter region, which was associated with DnaA and Rok *in vivo*.

As expected, there was no detectable binding of DnaAΔC to the DNA probe (**Fig. 13, lanes 1, 2**). In contrast, there was binding by Rok to the DNA fragment that resulted in a detectable shift in mobility of the DNA probe (**Fig. 13, lanes 3, 5**), as expected based on previous work (Albano et al., 2005). When DnaAΔC and Rok were both incubated with the DNA fragment, mobility of the probe was further reduced (**Fig. 13, lane 4**). These results indicated that Rok alone binds DNA and that DnaA can act on this Rok-DNA complex to cause a supershift. These results are consistent with those obtained from analyses of DNA association *in vivo*. We interpret these results as evidence of direct interaction between DnaA and Rok.

The nucleotide-bound state of DnaA affects cooperative binding to DNA. DnaA-ATP binding is more cooperative than that of DnaA-ADP (Bonilla and Grossman, 2012; Merrikh and Grossman, 2011). In contrast, the association of DnaA with Rok-bound DNA did not depend on the nucleotide bound to DnaA. For both full-length DnaA and DnaAΔC, binding to Rok was similar whether DnaA was incubated with ATP or ADP (data not shown). These results are consistent with the ChIP-seq results, and together indicate that DnaA association with Rok-bound DNA was different from direct binding of DnaA to DNA.
Figure 13. DnaAΔC requires Rok for association with DNA in vitro.

Purified DnaAΔC-His (2 µM) was incubated with a DNA probe (0.1 nM) in the presence or absence of purified Rok-His (200 nM). The DNA probe was a 32P-labeled PCR product corresponding to a 342-bp region of the *rok* promoter. The gel shown is representative of at least three independent replicates.

**Rok and replication initiation**

We postulated that Rok might have an effect on initiation of chromosomal replication, perhaps directly through its association with the DNA unwinding element in *oriC*, or indirectly by titrating DnaA away from direct DNA binding sites. H-NS, the Rok analog in *E. coli*, has an indirect positive effect on replication initiation (Atlung and Hansen, 2002), as discussed in Chapter 1. Additionally, the *E. coli* nucleoid-associated proteins Fis, HU, and IHF play important roles in binding *oriC* during replication initiation (Gille et al., 1991; Hwang et al., 1992; Polaczek, 1990; Ryan et al., 2004), so we surmised that Rok might contribute to replication in *B. subtilis*. 
We found little or no detectable effect of a *rok* null mutation on replication initiation. We measured the origin-to-terminus ratio (*oriC/terC*) by qPCR in cells growing exponentially in minimal and rich media. During growth in LB, defined minimal glucose medium, or defined minimal succinate medium, the deletion of *rok* did not significantly affect the *oriC/terC* ratio (Appendix A). These results indicate that if *rok* affects replication initiation, then the effects are either too small to measure under the conditions tested, or there are other mechanisms that might compensate for the loss of *rok*. By binding to the DUE, perhaps bending or occluding it, and by affecting the localization of DnaA, Rok might affect the precise timing of replication initiation or the ease of DUE melting. Such effects might be redundant with the functions of other DNA-binding proteins or masked by the several regulators of replication initiation.

**Gene regulation by DnaA and Rok**

DnaA and Rok are both transcription factors. Recent analyses of gene expression in a *dnaA* null mutant indicated that many genes normally repressed by Rok had increased expression in the absence of *dnaA* (T.A. Washington, J.L. Smith, and ADG, in preparation). In *B. subtilis*, DnaA regulates genes involved in DNA replication (Ogura et al., 2001), sporulation (Burkholder et al., 2001; Ishikawa et al., 2007), and the response to replication stress ((Goranov et al., 2005; Ishikawa et al., 2007); T.A. Washington, J.L. Smith, and ADG, in preparation). Rok regulates genes involved in competence (Hoa et al., 2002) and extracellular functions (Albano et al., 2005; Kovács and Kuipers, 2011; Marciniak et al., 2012). DnaA regulates some genes directly and other genes indirectly ((Burkholder et al., 2001; Ishikawa et al., 2007; Ogura et al., 2001); T.A. Washington, J.L. Smith, and ADG, in preparation). Many of the indirect effects of DnaA on gene expression are due to activation of the checkpoint gene *sda* ((Burkholder et al., 2001); T.A.
Washington, J.L. Smith, and ADG, in preparation). Similarly, Rok can regulate genes directly (Albano et al., 2005; Hoa et al., 2002; Marciniak et al., 2012) or indirectly (Hoa et al., 2002; Kovács and Kuipers, 2011). The co-association of DnaA and Rok at several chromosomal regions, including many promoters, suggested that one or both of these transcription factors might contribute to regulating gene expression at these regions.

We directly compared the effects of dnaA and rok null mutations on gene expression. We used a strain in which oriC was replaced with oriN and dnaA was removed. dnaN was expressed constitutively at its native locus from a derivative of the promoter Ppen (Experimental Procedures). We compared the dnaA null mutant to an isogenic dnaA+ strain. For a direct comparison with the dnaA null mutant, we also used the dnaA+ oriN background to study the effects of a rok null mutant and dnaA rok double mutant.

Using microarrays, we analyzed global gene expression (mRNA levels) during exponential growth in defined minimal glucose media. We used an adjusted P-value corresponding to at most one false positive as our significance criterion for identifying differentially expressed genes. Expression of dnaA was not significantly altered in the rok null mutant, and expression of rok was not significantly altered in the dnaA null mutant. We focused on 55 genes representing the 36 regions of Rok-dependent DnaA association. We considered these genes candidates for direct regulation by DnaA and Rok.

From the 55 candidate genes, we found three operons with at least one gene with significantly altered expression in the dnaA null mutant and the rok null mutant: sunA, yxaI, and yybN. Expression of several other genes was also altered in both mutants, but the effects fell slightly below our criteria for significance. Expression of target genes was increased in the dnaA and rok mutants (Table 2). For example, mRNA levels of five genes in the sun operon were
increased in both mutants (Table 2, lines 1-5). The sun genes are in the prophage SPβ and function in the biosynthesis and transport of the antibiotic sublancin. mRNA levels were increased approximately 2- to 5-fold in the dnaA null mutant and 6- to 20-fold in the rok null mutant (Table 2), indicating that DnaA and Rok normally act to repress expression of the sun operon. These inhibitory roles were consistent with previous findings from a dnaA mutant in a slightly different oriN background (T.A. Washington, J.L. Smith, and ADG, in preparation) and findings in a rok mutant replicating from oriC (Albano et al., 2005). Expression of the yybN operon and yxaJ was also increased in both the dnaA null mutant and the rok null mutant (Table 2, lines 7-9). Again, the inhibitory roles of DnaA and Rok were generally consistent with previous findings ((Albano et al., 2005); T.A. Washington, J.L. Smith, and ADG, in preparation), except yxaJ was not previously identified as differentially expressed in a dnaA null mutant.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Δrok (oriN)</th>
<th>ΔdnaA Δrok</th>
<th>ΔdnaA</th>
<th>dnaAΔC</th>
</tr>
</thead>
<tbody>
<tr>
<td>sunT</td>
<td>20.48</td>
<td>13.86</td>
<td>4.69</td>
<td>3.82</td>
</tr>
<tr>
<td>sunA</td>
<td>5.74</td>
<td>5.08</td>
<td>1.95</td>
<td>4.98</td>
</tr>
<tr>
<td>sunS (yolJ)</td>
<td>14.34</td>
<td>9.15</td>
<td>3.54</td>
<td>2.93</td>
</tr>
<tr>
<td>bdbA</td>
<td>17.07</td>
<td>10.53</td>
<td>2.70</td>
<td>2.75</td>
</tr>
<tr>
<td>bdbB</td>
<td>14.34</td>
<td>10.49</td>
<td>2.98</td>
<td>3.97</td>
</tr>
<tr>
<td>yxaI</td>
<td>-6.25</td>
<td>-7.58</td>
<td>-2.45</td>
<td>-2.79</td>
</tr>
<tr>
<td>yxaJ</td>
<td>2.60</td>
<td>2.81</td>
<td>1.80</td>
<td>2.23</td>
</tr>
<tr>
<td>yybM</td>
<td>1.95</td>
<td>3.78</td>
<td>1.91</td>
<td>2.62</td>
</tr>
<tr>
<td>yybN</td>
<td>2.32</td>
<td>5.70</td>
<td>2.54</td>
<td>2.53</td>
</tr>
</tbody>
</table>

Table 2. Gene regulation by DnaA and Rok.
Global gene expression in an oriN background was measured for the following strains: Δrok (CAS196), ΔdnaA Δrok (CAS192), ΔdnaA (CAL2074), dnaAΔC (CAS231). Values represent linear fold changes in gene expression, relative to a dnaA+ oriC- oriN+ strain (CAL2083) unless otherwise indicated. Values in bold are statistically significant.

yxaI, directly downstream of yxaJ and oriented in the opposite direction, had decreased expression in both mutants (Table 2, line 6), indicating that DnaA and Rok normally function to activate expression of yxaI either directly or indirectly. We suspect that the regulation of yxaI by
Rok is indirect, perhaps similar to the indirect activation of yuaB (bslA) by Rok (Kovács and Kuipers, 2011). Rok is not known to activate transcription, and yxaI (unlike yxaJ) is not significantly affected in a rok null mutant lacking the Rok-repressed transcriptional regulator comK (Albano et al., 2005).

To better characterize the potential co-regulation of genes by DnaA and Rok, we analyzed gene expression in a dnaA rok double mutant. For most genes, rok was epistatic to dnaA; that is, the gene expression effects of DnaA depended on Rok. For example, expression of yxaI and yxaJ appeared to be similar in the rok dnaA double mutant to that in the rok single mutant (Table 2), indicating that for these genes, rok was epistatic to dnaA. Expression of sunA, sunT, sunS, bdbA, and bdbB, was similar or slightly lower in the rok dnaA double mutant compared to the rok single mutant. For these genes, rok was likely epistatic to dnaA, although it is unclear why the double mutant had somewhat lower expression than the rok single mutant. In contrast, expression of yybM and yybN in the rok dnaA double mutant seemed to be additive from the two single mutants, indicating that Rok and DnaA were having independent effects.

For genes where rok was epistatic or nearly epistatic, we suspect that DnaA and Rok are the major regulators of gene expression and that DnaA affects the function of Rok. Previous studies of gene expression and Rok binding in vivo and in vitro support direct transcriptional repression of these genes by Rok (Albano et al., 2005; Smits and Grossman, 2010). For the genes where dnaA and rok had additive effects, despite the Rok-dependent association of DnaA with these regions, other regulators downstream of DnaA and/or Rok likely also contribute to gene expression. Both DnaA and Rok can have indirect effects on gene expression by affecting other transcription factors. DnaA activates expression of the sporulation checkpoint regulator sda (Ishikawa et al., 2007), and Rok represses the master regulator of competence development.
comK (Hoa et al., 2002). Additionally, the positioning of DnaA and Rok at a promoter (e.g., sunA) versus a coding sequence (e.g., the yybN operon) might contribute to the different effects of the single and double mutants on different operons.

Gene expression in the dnaAΔC mutant was similar to that of a dnaA null strain (Table 2). This result was expected for genes that are directly regulated by DnaA binding to DnaA boxes (e.g., ywcI, yydA). For chromosomal regions with Rok-dependent DnaA binding (e.g., the sun operon), this result indicates that the C-terminus of DnaA is required, directly or indirectly, for the effects of DnaA on gene expression, even though the C-terminus was not required for association with these genomic regions. The C-terminus of DnaA could be structurally important for influencing gene expression. Alternatively, the effects of DnaA on these genes could be indirect and due to a downstream regulator that is directly regulated by DnaA, such as Sda (T.A. Washington, J.L. Smith, and ADG, in preparation).

Our results indicate that control of gene expression by DnaA and Rok is not easily predicted from sequence features or protein association. The sun operon and several other regions were predicted to be directly regulated by DnaA because the promoters contained putative DnaA binding sites (Goranov et al., 2005). We found, however, that association of DnaA with the sunA promoter and other candidate binding regions did not require direct binding of DnaA to DNA (Fig. 10C).

There were several regulatory regions bound by both DnaA and Rok for which gene expression was not detectably altered in the mutants (e.g., the sboA operon). The positioning of DnaA and Rok might not allow effective transcriptional regulation. Alternatively, certain bound genes might not have been expressed under our experimental conditions, and effects of dnaA and rok might be revealed under different conditions, for example, during stationary phase.
Transcriptional response to replication stress

In *B. subtilis*, and likely other organisms, DnaA regulates a transcriptional response to replication stress ((Breier and Grossman, 2009; Goranov et al., 2005); T.A. Washington, J.L. Smith, and ADG, in preparation). Several of the regions associated with both DnaA and Rok were upstream of genes previously identified as responsive to replication stress (Goranov et al., 2005). These regions include *sunA*, *ahpC*, *katA*, *yydD*, *yxkC*, and *yjcM*. Based on the Rok-dependent association of DnaA with these regions, the effects of replication stress on these genes could be due to changes in association of DnaA with Rok, or they could be due to indirect effects.

Conservation of DnaA and Rok

DnaA is highly conserved in bacteria (Messer, 2002; Ogasawara et al., 1991; Zakrzewska-Czerwinska et al., 2007), but Rok has only been identified in *Bacillus* species and is believed to have been acquired by horizontal gene transfer (Albano et al., 2005; Singh et al., 2012). Although Rok is not widely conserved, we speculate that the general connection between DnaA and chromosome architecture proteins may extend to other organisms. Conservation of DnaA and Rok is discussed in additional detail in Appendix B.

A major function of Rok is as a nucleoid associated protein. Rok binds A+T-rich sequences and is an analog of the nucleoid-associated protein H-NS of gamma-proteobacteria and Lsr2 of *Mycobacteria* (Smits and Grossman, 2010). Binding of H-NS to DNA is complex; it involves bridging separate DNA sequences and is influenced by other proteins (Amit et al., 2003; Arold et al., 2010; Dame et al., 2000, 2006; Lang et al., 2007; Liu et al., 2010; Noom et al., 2007; Tippner and Wagner, 1995). These potential aspects of Rok function have not yet been
described, but it seems possible that Rok may be analogous to H-NS in these ways and that DnaA might affect an as-yet uncharacterized aspect of Rok function.

We suspect that DnaA in other organisms might interact with other nucleoid-associated proteins, including H-NS analogues. Rok and H-NS are both relatively small proteins (21 kDa and 15 kDa, respectively), predicted to be net positively charged (isoelectric points of 9.3 and 9.5, respectively). The robust conservation of DnaA in bacteria suggests that DnaA proteins from different organisms are likely to have similar biochemical properties. DnaA from *E. coli* and *B. subtilis* are 51% identical and bind the same consensus sequence, although they do not substitute for each other *in vivo* (Krause and Messer, 1999; Krause et al., 1997). Given the interaction between DnaA and Rok, it seems possible that DnaA could also interact with small, basic chromosome architecture proteins in other organisms.

Our results with the DnaA∆C mutant indicate that the C-terminus of DnaA is not required for DnaA-Rok interaction. The DnaA∆C mutant still contained the conserved AAA+ domain. If this domain contributes to the DnaA-Rok interaction, the interaction seems likely to be conserved in other organisms. Alternatively, the interaction might depend on the less conserved N-terminus of DnaA. If so, the interaction with Rok might be specific to the N-terminus of *B. subtilis* DnaA. Perhaps the DnaA N-terminus in different organisms conveys organism-specific specificity to interactions with chromosome architecture or other proteins. Interaction between the replication initiator and nucleoid-associated proteins analogous to Rok might then represent a broader regulatory connection between DNA replication and chromosome architecture.
Experimental Procedures

Strains and alleles

*B. subtilis* strains and relevant genotypes are listed in Table 3. Properties and construction of important alleles are described below.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype (reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG174</td>
<td><em>trpC2 pheA1</em> (wild-type, JH642) (Perego et al., 1988)</td>
</tr>
<tr>
<td>WKS1038</td>
<td>Δrok::cat (Smits and Grossman, 2010)</td>
</tr>
<tr>
<td>AIG200</td>
<td><em>trp+</em>, Δ*[oriC dnaA dnaN]<em>::spc, spoIIIJ::</em>[oriN repN kan], amyE::<em>[PxylA-dnaN cat]</em> (Goranov et al., 2005)</td>
</tr>
<tr>
<td>TAW5</td>
<td><em>trp+</em>, Δ*[oriC dnaA dnaN]<em>::spc, spoIIIJ::</em>[oriN repN kan], amyE::<em>[PxylA-dnaN cat]</em>, lacA::<em>[Pspank-dnaA tet]</em> (Merrikh and Grossman, 2011)</td>
</tr>
<tr>
<td>CAL2074</td>
<td>Δ*[oriC dnaA]<em>::</em>[oriN repN Ppen-2028-dnaN cat]</td>
</tr>
<tr>
<td>CAL2083</td>
<td>Δ*[oriC dnaA]<em>::</em>[oriN repN Ppen-2028-dnaN cat], lacA::*[Pspank-dnaA tet]</td>
</tr>
<tr>
<td>CAS192</td>
<td>Δrok::<em>cm::mls, Δ</em>[oriC dnaA]<em>::</em>[oriN repN Ppen-2028-dnaN cat]</td>
</tr>
<tr>
<td>CAS221</td>
<td>Δ*[oriC dnaA]<em>::</em>[oriN repN Ppen-2028-dnaN cat], amyE::*[Pspank-dnaA spc]</td>
</tr>
<tr>
<td>CAS231</td>
<td>Δ*[oriC dnaA]<em>::</em>[oriN repN Ppen-2028-dnaN cat], amyE::*[Pspank-dnaA ΔC spc]</td>
</tr>
</tbody>
</table>

Table 3. *B. subtilis* strains used in this study.

All strains are derived from AG174 and contain the *trpC2 pheA1* alleles unless otherwise indicated.

Δrok::cat. The *rok* open reading frame was replaced with a chloramphenicol resistance cassette (*cat*) by long-flanking homology PCR to generate strain WKS1030 (Smits and Grossman, 2010). Strain WKS1038 was the product of backcrossing genomic DNA from WKS1030 into wild-type AG174 cells. Strain CAS192 contained a derivative of this Δrok allele in which the chloramphenicol resistance gene was disrupted with an MLS (macrolide-lincosamide-streptogramin B) resistance gene (Δrok::cat::mls).

Δ*[oriC-dnaA-dnaN]*. Strains AIG200 and TAW5 contained a deletion-insertion, in which *dnaA* and most of *dnaN* were replaced with a spectinomycin resistance cassette (*spc*) by long-flanking homology PCR (Goranov et al., 2005; Merrikh and Grossman, 2011). Replication was supported by insertion of the heterologous origin *oriN* and its initiator *repN* near *oriC* at spoIIIJ.
(Berkmen and Grossman, 2007; Hassan et al., 1997; Moriya et al., 1997). *dnaN* was expressed from PxylA at *amyE*. These strains still contained the *dnaA* promoter region, including the cluster of DnaA binding sites. Strains AIG200 and TAW5 contained a deviation in the *ypjG-hepT* region, containing tryptophan biosynthesis genes, as described previously (Berkmen and Grossman, 2007).

Strain CAL2074 and its derivatives contain a deletion-insertion in which *dnaA* and flanking regions are replaced with a product (generated by isothermal assembly (Gibson et al., 2009)) containing a chloramphenicol resistance cassette, *oriN* and *repN*, and a promoter driving expression of *dnaN*. The chloramphenicol resistance cassette (including the transcription terminator with *cat*) was inserted at the left end of *oriC*, upstream of *rpmH*. *oriN* and *repN* were inserted upstream of this cassette such that *oriN-repN, cat*, and *rpmH* were co-directional. A derivative of the constitutive promoter Ppen (Ppen-2028, C. Lee, unpublished data) was cloned upstream of *dnaN*. Ppen is derived from the *B. licheniformis* penicillinase gene and drives *lacI* expression on the integration vector pDR110. Ppen-2028 carries mutations relative to Ppen (Ppen-2028 sequence in lowercase) between the putative -35 and -10 sequences (underlined): 5’- TTGCATTTAattcggtggtgTGAATACTTTTCAA-3’.

*dnaAΔC*. Strain CAS231 contained a truncated version of *dnaA* corresponding to the first 355 amino acids of DnaA and missing the C-terminal 91 amino acids comprising most of the DNA binding domain, as annotated by the Conserved Domains feature of PubMed Protein (http://www.ncbi.nlm.nih.gov/protein/16077069). The *dnaA* sequence and flanking regions from the integration vector pDR110 were amplified using genomic DNA from strain AMB89 and the following primers: oCS105 with oCS111 (upstream) and oCS112 with oCS106 (downstream). *dnaAΔC* was cloned downstream of the IPTG-inducible promoter Pspank in the Nhel-linearized
integration vector pDR110 using isothermal assembly. The resulting product was transformed into AG174 cells for double-crossover integration at *amyE*. This intermediate strain was then transformed with CAL2074 genomic DNA to produce strain CAS231. The same strategy was used for ectopic expression of full-length *dnaA* in CAS221. The full-length *dnaA* sequence was amplified using primers oCS105 and oCS106.

**Media and growth conditions**

Unless otherwise specified, all strains were grown at 30 °C in S7 defined minimal medium buffered with 50 mM MOPS (Jaacks et al., 1989) and containing 1% glucose, 0.1% glutamate, trace metals, 40 µg/ml tryptophan, and 40 µg/ml phenylalanine. For growth of AIG200 and TAW5, glucose was replaced with 1% arabinose, and 0.5% xylose was used to induce expression of DnaN from the xylose-inducible promoter P*xyrA*. To induce expression of DnaA from the LacI-repressible, IPTG-inducible promoter Pspank, 0.1 mM IPTG was added.

**ChIP-seq**

Immunoprecipitations were performed with anti-DnaA and anti-Rok rabbit polyclonal antibodies, or with mouse monoclonal anti-myc antibodies (Invitrogen) essentially as described (Lin and Grossman, 1998; Merrikh and Grossman, 2011). Briefly, exponentially growing cells were treated with 1% formaldehyde, and the cross-linked lysates were sonicated to shear the DNA. Immunoprecipitations were performed by incubating the cross-linked lysates with antibodies for at least two hours at room temperature, followed by incubation with Protein A sepharose beads for at least one hour at room temperature. A control sample of non-immunoprecipitated lysate was incubated under the same conditions. Immunoprecipitated
material was washed and eluted from the beads, followed by reversal of cross-links by incubation at 65° C overnight. Samples were then treated with proteinase K at 37° C for at least two hours, and DNA was recovered using the Qiagen PCR purification kit.

**Antibody specificity**

We assessed the specificity of each antibody by performing ChIP-seq in appropriate null mutants, missing the protein of interest. There was little or no detectable precipitation of specific chromosomal regions in ChIP-seq experiments with anti-DnaA and anti-Rok antibodies from *dnaA* and *rok* null mutants, respectively. There were weak signals (less than three-fold above background) in the anti-DnaA immunoprecipitations from a *dnaA* null mutant for *yoeC*, *yonT*, and *sboA*. However, sequence reads for these regions were not strongly symmetrical in the forward and reverse directions, indicating that they were most likely artifacts.

**High-throughput sequencing analysis**

We obtained approximately 7-50 million 40-nt reads for each sample. We mapped the reads to the *Bacillus subtilis* strain AG174 genome using bwa (Burrows-Wheeler Aligner) for single-end short reads (Li and Durbin, 2009) and allowing a maximum number of alignments (n) of 2. To make comparisons across samples, we normalized the number of reads at each chromosomal position to the total number of reads for that sample. To calculate coverage at each base pair on the chromosome, we computationally extended each read by the estimated average fragment length of 250 bp (J. L. Smith, ADG, submitted).

We used SISSRs (Jothi et al., 2008) to identify enriched regions in each ChIP sample. We compared each ChIP sample to the corresponding non-immunoprecipitated (total) DNA
sample. We used the following parameters: F (fragment length) = 250, e (sensitivity) = 1, m (fraction mappable) = 1, w (window) = 20, E (required number of reads) = 1 per million sample reads, and L (maximum fragment length) = 400. We then optimized p (P-value) such that the enrichment fold cutoff was approximately 3. From these candidate enriched regions, we selected those with an enrichment of at least five-fold.

To search for a Rok binding motif, we used regions with at least five-fold Rok enrichment as determined by SISSRs. For each region, we extracted 101 nucleotides of sequence, centered on the midpoint identified by SISSRs. These sequences were used as input for DREME, along with the following parameters: comparison source = shuffled sequences, both strands used, and maximum E-value = 1.

**Relative DnaA levels**

We used quantitative western blotting to evaluate DnaA protein levels and normalized these values to OD as a proxy for total protein. Exponentially growing cells were plated on LB in serial dilutions to determine cfu/ml. A culture sample of 15 ml was pelleted, and all but 1 ml supernatant was removed. Samples were frozen at -80 °C and resuspended with the addition of 14 ml TE (10 mM Tris pH 8.0, 10 mM EDTA) with protease inhibitors (AEBSF). The OD of the resuspension was measured in triplicate. Cells were lysed by sonication rather than lysozyme treatment because lysozyme cross-reacts with the LiCor goat anti-rabbit 800 secondary antibody used for quantitation. Samples were sonicated on ice for 6 min per sample, in bursts of 0.3 sec on/off. A 450-μl aliquot was taken, and 50 μl of TE plus protease inhibitors (AEBSF) was added. Equal volumes of each sample were analyzed on a 15% SDS-PAGE gel, along with standards of purified DnaA. Most samples were loaded on the gel in duplicate. Samples were
imaged and quantitated on the LiCor scanner. Protein intensities from the lysates were compared to serial dilutions of the purified protein standards. Relative protein levels (between-strains comparisons) were made by directly comparing the western blot signals, after normalization for OD. Western blot signals typically differed by less than a factor of 2.

**DNA microarrays**

Global mRNA levels were analyzed by hybridization to DNA microarrays as described (Goranov et al., 2009). Exponentially growing cells from at least three replicate cultures were fixed with an equal volume of -20 °C methanol. RNA was purified from lysates using the Qiagen RNEasy kit. Experimental and reference RNA samples (Goranov et al., 2009) were reverse transcribed using Superscript II reverse transcriptase (Invitrogen), random hexamers, and aminoallyl-dUTP (Ambion). The cDNA was labeled by conjugation to monofunctional Cy3 or Cy5 dyes (Amersham) for reference or experimental samples, respectively. Each experimental sample was mixed with an aliquot of reference sample. Salmon testes DNA and yeast tRNA were added, and each cDNA sample was hybridized to a DNA microarray at 42 °C overnight. Microarrays contained PCR products from >95% of the annotated *B. subtilis* ORFs spotted onto Corning GAPS slides. Microarrays were scanned with a GenePix 4000B scanner, and images were analyzed using GenePix 3.0 (Axon Instruments).

Data were analyzed using the R statistical software package Linear Models for Microarray Data (LIMMA) (Smyth, 2005). Spot intensities were normalized within and between arrays, and gene expression values were corrected for multiple hypothesis testing using the Benjamini-Hochberg correction option. Genes were identified as differentially expressed if the adjusted P-value corresponded to no more than one false positive within the given set of genes.
Protein expression and purification

dnaAΔC was cloned between the Ncol and NotI sites of pET-28b (generating pCAS254) that produces DnaAΔC with a C-terminal hexa-histidine tag (DnaAΔC-his). The construct was transformed into E. coli BL21 pLysS, and expression was induced with 1 mM IPTG for 5 h during growth in LB medium at 37 °C. A pellet from 1 liter of cells was frozen at -80 °C, thawed, and resuspended in lysis buffer (50 mM NaPO4 pH 7, 5 mM imidazole, 300 mM NaCl) containing AEBSF protease inhibitor (Sigma). MgCl2 (10 mM) and 3 µl Benzonase Nuclease (EMD Millipore) were added, and the lysate was stirred for 10 min. The lysate was cleared by centrifugation. The supernatant was loaded onto a 1-ml HisTALON column (Clontech). The column was washed with Talon buffer A (50 mM NaPO4 pH 7, 300 mM NaCl, 10% glycerol), followed by Talon buffer A containing 4.5 mM imidazole. The column was then eluted with a linear gradient of 4.5 to 150 mM imidazole in Talon buffer A. Fractions containing DnaAΔC were identified by SDS polyacrylamide gel electrophoresis, pooled, and combined with two volumes of buffer containing 45 mM HEPES-KOH pH 7.6, 0.75 mM EDTA, 15 mM magnesium acetate, 1.5 mM DTT, and 5% sucrose. The solution was loaded onto a 5-ml HiTrap Q FF column (GE Healthcare Life Sciences) and washed with Q buffer A (45 mM HEPES-KOH pH 7.6, 0.5 mM EDTA, 10 mM magnesium acetate, 1 mM DTT, 5% sucrose, 100 mM potassium glutamate). The column was eluted with a linear gradient to Q buffer B (Q buffer A containing 1 M potassium glutamate). The fractions containing DnaAΔC were found to have eluted in 100% Q buffer B. These fractions were pooled, and aliquots were stored at -80 °C.

Wild-type (full-length) DnaA was purified essentially as described in (Bonilla and Grossman, 2012; Merrikh and Grossman, 2011). dnaA was cloned between the Ncol and NotI sites of pET-28b (generating pCAL1293) to encode recombinant DnaA with a C-terminal hexa-
histidine tag (DnaA-his). The construct was transformed into E. coli BL21 pLysS, and expression was induced with 0.4 mM IPTG for 3.5 h during growth in LB medium at 30 °C. Full-length DnaA was purified according to the same protocol as DnaAΔC, except DNase I (50 µl) was used instead of Benzonase Nuclease, final pooled fractions were concentrated in an Amicon Ultra-15 50 kDa MWCO concentrator unit (GE Healthcare Life Sciences), and final storage buffer was estimated to contain 700 mM potassium glutamate.

Rok was expressed with a C-terminal hexa-histidine tag from pED428 in E. coli M15 (Hoa et al., 2002). Expression was induced with 1 mM IPTG for 3 h during growth in LB medium at 37 °C. A pellet from 500 ml of cells was resuspended in lysis buffer (50 mM NaPO4 pH 7, 4.5 mM imidazole, 300 mM NaCl) and frozen at -80 °C. The cell suspension was thawed, and MgCl2 (10 mM), AEBSF protease inhibitor (Sigma), and 1.5 ul Benzonase Nuclease (EMD Millipore) were added. The suspension was stirred for 15 min, lysed with lysozyme (1 mg/ml), and stirred for an additional 15 min. The lysate was cleared by centrifugation. The supernatant was loaded onto a 1-ml HisTALON column (Clontech). The column was washed with Talon buffer A (50 mM NaPO4 pH 7, 300 mM NaCl, 10% glycerol), followed by Talon buffer A containing 25 mM imidazole. The column was then eluted with a linear gradient to Talon buffer B (Talon buffer A containing 2 M imidazole). Fractions containing Rok were identified by SDS-PAGE, pooled, and loaded onto a HiTrap Heparin HP column (GE Healthcare Life Sciences). The column was washed with Talon buffer A and eluted with a linear gradient to Heparin buffer B (Talon buffer A containing 2 M NaCl). Fractions containing Rok were identified by SDS-PAGE and pooled. The buffer for these fractions was calculated to contain 750 mM NaCl, and glycerol was added to a final concentration of 10%. Protein was concentrated using a Vivaspin 6
5 kDa MWCO concentrator unit (GE Healthcare Life Sciences), and aliquots were stored at -80 °C.

**Gel electrophoresis mobility shift assay**

A 342-bp region of the rok promoter was amplified from B. subtilis genomic DNA using primers HM57 (5’-CGGGATCCGCTTCTCTCTTCCAATACCAT-3’) and HM58 (5’-CGGAATTCTGATTTTTGCTCCTTATTTAG-3’). The PCR product was purified with a PCR purification column (Qiagen) and analyzed on a 6% polyacrylamide gel (37:1 acrylamide/bisacrylamide) containing 1X TAE and run in 1X TAE. The PCR product corresponding to the expected size was extracted and purified. The purified probe was end-labeled with radioactive gamma-32P-ATP (Perkin Elmer) using T4 Polynucleotide Kinase (New England Biolabs). Excess label was removed using a MicroSpin G-50 column (GE Healthcare Life Sciences). The DNA probe was used at a final concentration of 0.1 nM in each reaction. For experiments measuring the effects of nucleotide-bound DnaA, purified DnaA was incubated with 5 mM ATP or ADP on ice for 2 h. In other experiments, this incubation step was omitted.

Gel shift reactions were performed in gel shift buffer (5 mM NaPO4 pH 7, 15 mM HEPES-KOH pH 7.6, 10 mM magnesium acetate, 300 mM NaCl, 100 mM potassium glutamate, 0.5 mM EDTA, 10% glycerol, 50 ug/ml bovine serum albumin, 0.5% sucrose, 0.1 mM DTT) with protein concentrations as indicated in the figure legends. Reactions with full-length DnaA also contained 2.5 mM ATP or ADP and a final DTT concentration of 1 mM. Reactions were incubated at room temperature for 20 min. Samples of each reaction were run on 5% polyacrylamide gels (37:1 acrylamide/bisacrylamide) containing 0.5X Tris-borate-EDTA (TBE) and 2.5% glycerol. Gels were run in 0.5X TBE at approximately 12 V/cm for 3 h at 4 °C. Gels
were dried, exposed to a storage phosphor screen (GE Healthcare Life Sciences), and imaged on a Typhoon scanner (GE Healthcare Life Sciences).

Acknowledgments

We thank Janet Smith, Catherine Lee, Houra Merrikh, and Tracy Washington for useful discussions. This work was supported, in part, by the Department of Defense, Air Force Office of Scientific Research, National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a, to CAS and Public Health Service grant GM41934 to ADG.
Chapter 3

Recruitment of the bacterial replication regulator YabA to several chromosomal regions by the nucleoid-associated protein Rok in *Bacillus subtilis*

Charlotte A. Seid, Tracy A. Washington, Houra Merrikh, and Alan D. Grossman

This chapter is being prepared for publication.
Abstract

We found a connection between a negative regulator of DNA replication (YabA) and the nucleoid-associated protein Rok of *Bacillus subtilis*. YabA interacts with the replication initiator DnaA and the sliding clamp of the replication machinery, DnaN. YabA associates with the replisome and with chromosomal regions bound by DnaA, including the origin of replication. Using ChIP-seq, we found that YabA also associated with chromosomal regions that are bound by the nucleoid-associated protein Rok, an analogue of H-NS from gamma-proteobacteria. Association of YabA with these Rok-bound regions was dependent on Rok but not on DnaA. Rok and DnaA remained associated with these regions in the absence of YabA. Our results indicate that YabA may be involved in cellular physiology beyond its known role as a negative regulator of replication initiation.

Introduction

YabA is a negative regulator of replication initiation in *B. subtilis* (Goranov et al., 2009; Hayashi et al., 2005; Noirot-Gros et al., 2002) and likely acts similarly in the other Gram-positive organisms in which it is found. YabA interacts with DnaA, the replication initiator. DnaA is a AAA+ protein that binds to sequences (DnaA boxes) found in oriC and other chromosomal regions. *In vitro*, YabA inhibits cooperative binding of DnaA to DNA (Merrikh and Grossman, 2011; Scholefield and Murray, 2013). *In vivo*, YabA decreases the amount of DnaA associated with oriC (Merrikh and Grossman, 2011), which is thought to account for the negative effects of YabA on replication initiation.

YabA also interacts with DnaN (Noirot-Gros et al., 2002, 2006), the processivity clamp of DNA polymerase. The interaction between YabA and DnaN likely serves to antagonize
YabA-mediated inhibition of DnaA by removing YabA from chromosomal regions bound by DnaA (Goranov et al., 2009; Merrikh and Grossman, 2011). Some YabA in the cell is associated with the replisome, likely through its interaction with the processivity clamp (Goranov et al., 2009; Hayashi et al., 2005; Noirot-Gros et al., 2006; Soufo et al., 2008).

YabA is associated with at least some of the chromosomal regions bound by DnaA, including oriC (Merrikh and Grossman, 2011). YabA alone appears incapable of binding DNA, or binds with very low affinity, and its association with chromosomal regions bound by DnaA in vivo depends on DnaA (Merrikh and Grossman, 2011). Most of the work on YabA since its discovery has focused on its role in replication.

We took a genomic approach to define the chromosomal regions that are associated with YabA in vivo. Using chromatin immunoprecipitation and deep sequencing (ChIP-seq), we found that YabA was associated with many chromosomal regions, including all those bound by DnaA and most if not all those bound by the nucleoid-associated protein Rok. Rok binds A+T-rich regions throughout the chromosome and has a role in silencing some regions of horizontally acquired DNA (Smits and Grossman, 2010). In these ways, Rok is analogous, but not homologous, to H-NS from E. coli and other gamma-proteobacteria (Dorman, 2007, 2010; Navarre et al., 2006) and to Lsr2 from Mycobacteria (Gordon et al., 2008, 2010). Rok is a repressor of comK and thus a negative regulator of competence development in B. subtilis (Hoa et al., 2002). Rok also represses expression of several other genes (Albano et al., 2005) and binds directly to their promoters, although a defined binding site has not been identified (Albano et al., 2005; Smits and Grossman, 2010).

Because of the association of YabA with some of the chromosomal regions bound by Rok, and its previously known association with at least some of the chromosomal regions bound
by DnaA (Merrikh and Grossman, 2011), we compared the genome-wide association of these three proteins. Previously, we analyzed the genome-wide association of Rok and DnaA. This work uncovered two types of regions bound by DnaA: 1) regions where DnaA binds directly to DnaA boxes via the DNA binding domain in DnaA; and 2) regions where DnaA associates with Rok, independently of the presence of DnaA boxes and independently of the DNA binding domain in DnaA (Chapter 2). Using ChIP-seq, we found that, in addition to its association with the previously identified DnaA box regions, YabA was associated with some of the chromosomal regions bound by Rok. Furthermore, we found that this association was largely dependent on Rok. That is, deletion of *rok* greatly reduced association of YabA with these regions. DnaA is also present at these regions (Chapter 2). However, in contrast to the DnaA-dependent binding of YabA to oriC and other direct DnaA targets, YabA did not depend on DnaA for association with the Rok-bound regions. Our results indicate that the subcellular location of YabA is largely determined by the binding of DnaA and Rok to genomic regions. Our findings indicate that Rok and YabA may modulate each other's functions.

**Results**

**Overview of ChIP-seq analysis of YabA and Rok**

YabA was previously known to associate with *oriC* and several other regions bound by DnaA (Merrikh and Grossman, 2011). Preliminary ChIP-chip experiments indicated that YabA was associated with some of the chromosomal regions bound by Rok (HM, TAW, ADG). To examine the chromosome-wide association of YabA and Rok at high resolution, we performed ChIP-seq experiments from exponentially growing *B. subtilis* cells in defined minimal glucose medium. We used the peak-calling algorithm SISSRs (Jothi et al., 2008) and a five-fold
enrichment cutoff (Experimental Procedures) to define chromosomal regions that were associated with YabA or Rok.

To evaluate the specificity of the antibodies and the immunoprecipitation reactions, we performed analogous ChIP-seq experiments with anti-YabA and anti-Rok antibodies in the appropriate null mutants (Experimental Procedures). The anti-Rok antibodies were highly specific. That is, there was little or no specific DNA in the anti-Rok immunoprecipitates from the *rok* null mutant. There was a small amount of enrichment of some chromosomal regions in anti-YabA immunoprecipitates from a *yabA* null mutant (Experimental Procedures). Further verification of the ChIP-seq results with YabA is described in more detail below.

**ChIP-seq analysis of genome-wide binding by YabA**

We found that YabA associated with many previously described DnaA-bound regions but was not restricted to these chromosomal regions (Fig. 1A). Eight DnaA binding regions, each containing at a cluster of DnaA-boxes, have been defined by ChIP-chip (Breier and Grossman, 2009), chromatin affinity purification and hybridization to microarrays (ChAP-chip) (Ishikawa et al., 2007), and more recently, ChIP-seq (Chapter 2). These DnaA box clusters are located in the intergenic regions between: *rpmH/dnaA* (the *dnaA* promoter region), *dnaA/dnaN* (directly upstream of the DNA unwinding element in *oriC*), *gcp/ydiF*, *yqeG/sda*, *ywLC/ywlB*, *ywCI/vpr*, *yydA/yycS*, and *trmE/jag*. YabA was previously found to associate with four of these regions (*rpmH/dnaA*, *yydA/yycS*, *ywLC/ywlB*, and *ywCI/vpr*) in a DnaA-dependent manner (Merrikh and Grossman, 2011). We observed YabA at all eight regions to some extent (Fig. 2), but only seven of them were recognized by the peak-calling algorithm SISSRs, and only five of these met our five-fold enrichment cutoff (Table 1). It is likely that YabA is also associated with the
remaining DnaA box cluster regions (\textit{yqeG/sda}) but below detectable levels in our ChIP-seq experiments (see below). Alternatively, YabA might not associate with DnaA at this region.

*Figure 1. Genome-wide binding of YabA and Rok in wild-type cells.*

Wild-type cells (strain AG174) were grown to mid-exponential phase, and YabA (A) and Rok (B) were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of ori\(C\) is set as 0. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.

Asterisks and black labels in (A) indicate previously identified DnaA binding sites (Breier and Grossman, 2009; Ishikawa et al., 2007) at which YabA was detected. Dashed lines and red labels in (B) indicate selected previously identified Rok binding sites (Albano et al., 2005; Smits and Grossman, 2010). Data represent single biological experiments.
Figure 2. Association of YabA and DnaA with DnaA box cluster chromosomal regions in wild-type cells.

YabA and DnaA were immunoprecipitated from wild-type cells (strain AG174). Each panel shows a magnified view of the data from Fig. 1. The eight DnaA box cluster regions include (A) rpmH-dnaA, (B) dnaA-dnaN, (C) gcp-ydiF, (D) yqeG-sda, (E) ywlC-ywlB, (F) ywcI-vpr, (G) yydA-yycS, and (H) trmE-jag (Breier and Grossman, 2009; Ishikawa et al., 2007). ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. DnaA is plotted on a larger y-axis scale in panels (A) and (B). The position of oriC is set as 0. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions. DnaA boxes were defined by a position-specific site matrix (J. L. Smith, ADG, submitted). DnaA boxes on the plus and minus strands are plotted as right-facing and left-facing arrows, respectively.
Table 1. Association of YabA with DnaA box cluster regions.

Enrichment values were calculated using SISSRs. For reference, the replication terminus terC is located 1.99 Mb from oriC. Enrichment values with asterisks were below the five-fold cutoff used in our Experimental Procedures. Brackets indicate that the sda region was not called as a YabA binding peak by SISSRs (enrichment was <2.21-fold), so enrichment was estimated from ChIP-seq coverage of YabA relative to non-immunoprecipitated DNA.

<table>
<thead>
<tr>
<th>Region</th>
<th>DnaA enrichment</th>
<th>YabA enrichment</th>
<th>Chromosomal position, Mb</th>
<th>Distance from oriC, Mb</th>
</tr>
</thead>
<tbody>
<tr>
<td>dnaA</td>
<td>49.85</td>
<td>15.36</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>dnaN</td>
<td>68.34</td>
<td>14.43</td>
<td>0.0018</td>
<td>0.0018</td>
</tr>
<tr>
<td>jag</td>
<td>24.66</td>
<td>9.52</td>
<td>4.18</td>
<td>0.0026</td>
</tr>
<tr>
<td>yydA</td>
<td>13.67</td>
<td>5.60</td>
<td>4.11</td>
<td>0.07</td>
</tr>
<tr>
<td>ywcI</td>
<td>11.72</td>
<td>9.18</td>
<td>3.88</td>
<td>0.30</td>
</tr>
<tr>
<td>gcp</td>
<td>8.67</td>
<td>2.66*</td>
<td>3.77</td>
<td>0.41</td>
</tr>
<tr>
<td>sda</td>
<td>17.64</td>
<td>3.1*</td>
<td>0.63</td>
<td>0.63</td>
</tr>
</tbody>
</table>

In addition to the five DnaA box cluster regions, we found that YabA was associated with 44 other chromosomal regions (Table 2; Fig. 1A). These additional regions were previously found to be associated with the nucleoid-associated protein Rok (Chapter 2 and (Smits and Grossman, 2010)). These novel YabA-bound regions included several promoter regions that are regulated by Rok (Albano et al., 2005) and bound by Rok in vivo (Chapter 2 and (Smits and Grossman, 2010)) and in vitro (Albano et al., 2005). Some of the genes in these binding regions included: sunA, sboA, yxaJ/yxaI, lytA/tagU, yucB/yutJ, yxkC/D, yybN, and yydD (Fig. 3).

These YabA-bound, Rok-bound chromosomal regions were essentially the same regions that were associated with DnaA in a Rok-dependent manner (Fig. 4). Thirty-four of the 44 regions showed Rok-dependent association of DnaA (Chapter 2). The remaining ten regions still showed DnaA enrichment (3.63-fold to 4.97-fold) but fell slightly below the five-fold enrichment cutoff used in our Experimental Procedures. Similarly, two of the 36 regions with Rok-dependent DnaA binding (argG and sunT cds) still showed YabA enrichment (4.59-fold and 4.52-fold, respectively) but fell slightly below the five-fold cutoff.
<table>
<thead>
<tr>
<th>Nearest gene(s)</th>
<th>Annotated function(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yceB/C</td>
<td>monooxygenase; stress adaptation protein</td>
</tr>
<tr>
<td>dtpT/yclG</td>
<td>di-tripeptide-proton ABC symporter; uronase</td>
</tr>
<tr>
<td>yddM</td>
<td>helicase in ICEBs1</td>
</tr>
<tr>
<td>pspA</td>
<td>phage shock protein A</td>
</tr>
<tr>
<td>ssuB/katA</td>
<td>aliphatic sulfonate ABC transporter ATP-binding protein, catalase</td>
</tr>
<tr>
<td>citR/citA *</td>
<td>transcriptional regulator; citrate synthase</td>
</tr>
<tr>
<td>comK</td>
<td>competence transcription factor</td>
</tr>
<tr>
<td>appD</td>
<td>oligopeptide ABC transporter ATP-binding protein</td>
</tr>
<tr>
<td>trpS/oppA *</td>
<td>tRNA synthetase; oligopeptide ABC transporter binding lipopeptide</td>
</tr>
<tr>
<td>yjcM/N</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>rok</td>
<td>repressor of comK</td>
</tr>
<tr>
<td>yknZ/fruR *</td>
<td>permease; transcriptional regulator</td>
</tr>
<tr>
<td>BSUA_01899 *</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>ynfC/alsT *</td>
<td>hypothetical protein; amino acid carrier protein</td>
</tr>
<tr>
<td>ppsA</td>
<td>plipastatin synthetase (antifungal peptide)</td>
</tr>
<tr>
<td>yobI cds</td>
<td>NTPase with transmembrane helices</td>
</tr>
<tr>
<td>yobJ cds</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>odhA *</td>
<td>2-oxoglutarate dehydrogenase E1</td>
</tr>
<tr>
<td>yosX</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yonX</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>sunA</td>
<td>lantibiotic precursor peptide</td>
</tr>
<tr>
<td>yqgA/B</td>
<td>hypothetical protein; factor involved in motility</td>
</tr>
<tr>
<td>yqxI</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yraN/M</td>
<td>LysR transcriptional regulator, hypothetical protein</td>
</tr>
<tr>
<td>araA</td>
<td>L-arabinose isomerase</td>
</tr>
<tr>
<td>iscS/braB</td>
<td>cysteine desulfurase; branched chain amino acid/Na+ symporter</td>
</tr>
<tr>
<td>nupN</td>
<td>lipoprotein</td>
</tr>
<tr>
<td>comP cds *</td>
<td>competence two-component sensor histidine kinase</td>
</tr>
<tr>
<td>yuzB/yutJ</td>
<td>hypothetical protein, NADH dehydrogenase</td>
</tr>
<tr>
<td>lytA/tagU</td>
<td>membrane-bound lipoprotein; membrane-bound transcriptional regulator</td>
</tr>
<tr>
<td>ggaB cds, 1 of 2 *</td>
<td>poly(glucosyl NAG 1-phosphate) glucosyltransferase</td>
</tr>
<tr>
<td>ggaB cds, 2 of 2</td>
<td>poly(glucosyl NAG 1-phosphate) glucosyltransferase</td>
</tr>
<tr>
<td>ggaA</td>
<td>poly(glucosyl NAG 1-phosphate) glucosyltransferase</td>
</tr>
<tr>
<td>glyA</td>
<td>serine hydroxymethyltransferase</td>
</tr>
<tr>
<td>sboA</td>
<td>subtilosin (and associated operon including transporter)</td>
</tr>
<tr>
<td>yxkC/D</td>
<td>hypothetical protein, efflux transporter</td>
</tr>
<tr>
<td>yxaI/J</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>ahpC</td>
<td>alkyl hydroperoxide reductase</td>
</tr>
<tr>
<td>yydH cds</td>
<td>hypothetical protein</td>
</tr>
</tbody>
</table>
**Table 2. Rok-bound chromosomal regions associated with YabA.**

Regions are listed in order of chromosomal position. Association within a coding sequence is indicated by “cds.” Asterisks indicate regions where DnaA was enriched in a ChIP-seq analysis (Chapter 2), but below five-fold. All other regions were associated with DnaA in a Rok-dependent manner (Chapter 2).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>yydD cds</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yybN</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yybM cds</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>yybK cds</td>
<td>hypothetical protein</td>
</tr>
</tbody>
</table>

**Figure 3. Association of YabA and Rok with Rok-bound chromosomal regions in wild-type cells.**

YabA and Rok were immunoprecipitated from wild-type cells (strain AG174). Each panel shows a magnified view of the data from Fig. 1. Eight selected chromosomal regions include (A) sunA, (B) sboA, (C) yxaI/yxaJ, (D) lytA/tagU, (E) yuzB/yutJ, (F) yxkC/D, (G) yybN, and (H) yydD. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. YabA and Rok are plotted on a larger y-axis scale in panels (A) and (B). The position of oriC is set as 0. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions.
Figure 4. Comparison of chromosomal regions bound by YabA, DnaA, and Rok.
The Venn diagram includes all chromosomal regions at which YabA, DnaA, or Rok was
enriched at least five-fold in a ChIP-seq analysis of genome-wide binding in wild-type cells
(Chapters 2, 3). Underlined numbers represent DnaA box cluster regions (Breier and
Grossman, 2009; Ishikawa et al., 2007). At regions other than DnaA box clusters, association of
DnaA and YabA depended on Rok (see below and Chapter 2). Single asterisks indicate regions
at which YabA enrichment was detectable but below the five-fold cutoff. Double asterisks
indicate regions at which DnaA enrichment was detectable but below the five-fold cutoff.

Because YabA appeared to be associated with several of the chromosomal regions bound
by Rok, we were concerned about the specificity of the anti-YabA antibodies used in the
immunoprecipitation. To verify that the apparent association of YabA with these regions was
due to immunoprecipitation of YabA and not some cross-reacting protein (e.g., Rok), we
performed ChIP-seq analysis with anti-YabA antibodies in a yabA null mutant. The anti-YabA
antibodies associated with three chromosomal regions, which were also bound by Rok (Chapter
2), in a yabA null mutant (Experimental Procedures). The enrichment of these chromosomal
regions was greatly reduced in the yabA null mutant relative to the yabA+ strain, but was not
eliminated (Experimental Procedures), indicating that most (but not all) of the DNA that was co-
precipitated with the anti-YabA antibodies was actually due to association with YabA. We
further verified the association of YabA with Rok-bound chromosomal regions, using ChIP-qPCR of YabA containing a 3x-myc-tag (N-terminal fusion) and anti-myc monoclonal antibodies. These results confirmed that YabA was associated with these chromosomal regions, either directly or indirectly.

Together, our results indicate that YabA interacts with several chromosomal regions known to be directly bound by DnaA, as well as several regions bound by the nucleoid associated protein Rok. We considered three possibilities to explain the association of YabA and Rok with several of the same chromosomal regions. 1) Since the ChIP-seq analyses were done on a population of cells, different subpopulations could have one or another of the proteins associated with a given region. In this way, each protein could be associated with the same region, but just not in the same cells. 2) The association of these proteins could simply be a coincidence of independent binding to the same chromosomal regions, although YabA is not known to bind DNA directly. 3) Association of YabA and Rok at the same chromosomal regions could depend on YabA and Rok. In particular, we considered whether YabA might depend on Rok at these Rok-bound regions since Rok is a DNA binding protein and YabA is not known to bind DNA on its own.

**ChIP-seq analysis of genome-wide YabA binding in a rok null mutant**

We found Rok was required for association of YabA with some chromosomal regions, but not with others. We analyzed the genome-wide binding of YabA, using ChIP-seq, in a rok null mutant. There was little or no effect of a rok deletion on YabA association with the DnaA box cluster regions (**Fig. 5A, B**). That is, we detected YabA at the same five DnaA box cluster regions regions in the rok null mutant as in the wild-type strain (**Fig. 4A**). This result was
expected since Rok was not detectably associated with most of these regions (Chapter 2). YabA did not associate with additional DnaA box cluster regions in the absence of rok.

In contrast, Rok was required for association of YabA at the other chromosomal regions. In the rok null mutant, we did not detect association of YabA at the 44 regions that showed co-association of YabA and Rok in wild-type cells (Fig. 5A, C). No additional YabA-bound regions were detected in the rok null mutant, so we do not believe that Rok occludes any potential YabA binding sites on the chromosome. We confirmed that YabA protein levels were similar between the wild-type and rok-null strains (see below), indicating that the changes in binding did not reflect a decrease in the amount of YabA in cells. We also confirmed that this result was not due to cross-reactivity of the YabA antibodies with another protein. We verified the Rok-dependent association of YabA with selected chromosomal regions using ChIP-qPCR with myc-YabA and anti-myc antibodies (Experimental Procedures). The anti-myc antibody showed negligible nonspecific association with these chromosomal regions in a rok+ strain and a rok null strain (Experimental Procedures), indicating that the anti-myc antibody did not cross-react with Rok.

The Rok-dependent association of YabA with certain chromosomal regions indicated that both proteins were present at the target regions in the same cells and not in separate subpopulations of cells. DnaA also associates with all 44 of these regions to some extent (see above and Chapter 2). Since DnaA also depends on Rok for association with these regions (Chapter 2), we considered whether there might be dependence between DnaA and YabA at the Rok-bound regions.
Figure 5. Genome-wide binding of YabA in wild-type cells and a rok null mutant. Wild-type cells (AG174) and a rok null mutant (WKS1038) were grown to mid-exponential phase, and YabA was immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC is set as 0. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments. (A) Genome-wide binding of YabA in wild-type cells (upper y-axis, black) and a rok null mutant (lower y-axis, red). Asterisks indicate the DnaA box cluster regions at which YabA was detected. Association of YabA with the eight DnaA box cluster regions (B) and selected Rok-bound regions (C) in wild-type cells (solid lines) and a rok null mutant (dashed lines). Each panel shows a magnified and superimposed view of the data from (A). The selected Rok-bound chromosomal regions include sunA, sboA, yuzB/yutJ, and yybN. DnaA boxes (upper row) and gene annotations (lower row) are shown below the corresponding chromosomal positions.
ChIP-seq analysis of genome-wide binding in a dnaA-null mutant

YabA is known to depend on DnaA for association with several DnaA box cluster regions (Merrikh and Grossman, 2011). We analyzed the genome-wide binding of YabA and Rok, using ChIP-seq, in a dnaA null mutant. dnaA is usually essential because of its role in replication initiation. We used a strain background in which dnaA is no longer essential because replication initiates from a heterologous origin, oriN, using its cognate initiator, RepN (Hassan et al., 1997; Moriya et al., 1997). We used a strain in which oriN was inserted near oriC, and the oriC region, including dnaA and dnaN, was removed. dnaN was expressed from a xylose-inducible promoter (PxyldnaN), integrated at another region of the chromosome, essentially as described (Merrikh and Grossman, 2011). These cells express ~3.5-fold more DnaN than wild-type cells under our experimental conditions (Merrikh and Grossman, 2011). We compared this dnaA null mutant to an isogenic strain expressing dnaA from an IPTG-inducible promoter (PspankdnaA) integrated into the chromosome at an ectopic site (Experimental Procedures). In a ChIP-seq analysis of DnaA in this dnaA+ strain, DnaA was associated with the same regions (except for the oriC region, which is missing) as in the wild-type strain (Chapter 2).

Loss of DnaA affected association of YabA with several, but not all, chromosomal regions. As expected and previously reported for four DnaA box cluster regions (Merrikh and Grossman, 2011), DnaA was required for association of YabA with the five DnaA box cluster regions where YabA was detected in wild-type cells. We did not detect YabA at any of these regions by ChIP-seq in the dnaA null mutant (data not shown), although YabA was still detectably enriched at other regions (Fig. 6). We confirmed these negative results for selected DnaA box cluster regions using qPCR (data not shown).
Figure 6. Genome-wide binding of YabA and Rok in a dnaA null mutant.

A dnaA null mutant (strain AIG200) was grown to mid-exponential phase, and YabA (upper y-axis, black) and Rok (lower y-axis, red) were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC in the wild-type strain is set as 0, but oriC was deleted from strain AIG200. This strain contained a deletion of dnaA and dnaN, replicated from the heterologous origin oriN, and expressed dnaN from Pxyl-dnaN. Data represent single biological experiments.

Loss of DnaA did not abolish association of YabA with the chromosomal regions that were also associated with Rok. In the dnaA null mutant, YabA was still detected by ChIP-seq at Rok-bound chromosomal regions, but not at any other places (Fig. 6). The ChIP-seq signal for YabA was very low in the dnaA+ and dnaA- strains replicating from oriN, likely due to DnaN overexpression (Merrikh and Grossman, 2011). We used ChIP-qPCR to confirm YabA association with Rok-bound regions in these strains. We chose six chromosomal regions (sunA, sboA, yybN, yjcM, iscS/braB, yuzB/yutJ) representing a range of YabA enrichment levels, binding patterns (single or multiple ChIP-seq peaks), and chromosomal positions. We found that YabA association with these six chromosomal regions bound by Rok was similar between the
DNA+ and DNA- strains (Fig. 7A). We confirmed that this result was not due to cross-reactivity of the YabA antibodies with another protein. We immunoprecipitated myc-YabA with anti-myc monoclonal antibodies in DNA+ and DNA- strains, and using qPCR, we observed DNA-independent association of YabA with the six selected regions (Fig. 7B).

Together, our results indicate that YabA depends on DnaA for association with DnaA box cluster regions, but that YabA associates with other chromosomal regions even in the absence of DnaA. As discussed above, YabA was dependent on Rok for association with the regions other than DnaA box clusters. At these Rok-dependent regions, Rok might associate with both YabA and DnaA in all cells, or Rok might associate with YabA in some cells and with DnaA in others. The population-based ChIP-seq results cannot distinguish between these possibilities. Rok and DnaA directly interact in vitro, and this interaction likely accounts for the co-association of Rok and DnaA in vivo (Chapter 2). We considered whether YabA might affect this interaction or otherwise influence the genome-wide binding of Rok and DnaA.
Figure 7. YabA does not require DnaA for association with some chromosomal regions. ChIP-qPCR was used to determine the relative association of the indicated chromosomal regions with YabA or myc-YabA in strains with or without dnaA. PCR primers were designed to amplify portions of seven chromosomal regions at which YabA and DnaA depend on Rok for association. Enrichment of these chromosomal regions was normalized to non-immunoprecipitated DNA and to a control region (yhaX) that is not bound by YabA, DnaA, or Rok. A value of one indicates no enrichment relative to the control region. Data represent averages of three biological replicates, and error bars represent the standard error. 

(A) dnaA+ cells (strain TAW5, gray bars) and a dnaA null mutant (strain AIG200, white bars) were grown to mid-exponential phase, and YabA was immunoprecipitated after cross-linking with formaldehyde. 

(B) Strains expressing 3x-myc-yabA in a dnaA+ background (CAS190, gray bars) and a dnaA null background (CAS147, white bars) were grown to mid-exponential phase, and myc-YabA was immunoprecipitated with anti-myc antibodies after cross-linking with formaldehyde.
ChIP-seq analysis of genome-wide binding in yabA null mutant

We found that neither Rok nor DnaA was dependent on YabA for association with any chromosomal regions. We analyzed the genome-wide binding of Rok and DnaA, using ChIP-seq, in a yabA null mutant compared to an isogenic yabA+ strain. yabA mutants over-initiate replication, and altered replication can affect the activity of DnaA as a transcription factor (Breier and Grossman, 2009; Goranov et al., 2005). To avoid possible complications of comparing two strains that have different frequencies of replication initiation, we used strains with oriC inactivated and replicating from the heterologous origin oriN+. oriN is integrated into the chromosome near oriC, and oriC is inactivated by a small deletion of the DnaA binding sites adjacent to the DNA unwinding element (Berkmen and Grossman, 2007). YabA does not affect replication in this strain background (Goranov et al., 2009), so we could directly compare a yabA null mutant to an isogenic yabA+ strain without being concerned about effects on replication initiation. ChIP-seq of DnaA and Rok in yabA+ and yabA mutant strains was performed as described above. We found that the genomic regions bound by Rok (Fig. 8A) and DnaA (Fig. 8B) were essentially identical between the yabA+ and yabA mutant strains.

Together, our results indicate that YabA associates with chromosomal regions that are bound by Rok and DnaA. At the Rok-bound regions, YabA association depends on Rok but not DnaA. At regions having DnaA boxes that are bound directly by DnaA, association of YabA is dependent on DnaA. We did not detect effects of YabA on the genome-wide association profiles of Rok or DnaA.
Figure 8. Genome-wide binding of Rok and DnaA in yabA+ and yabA null mutant cells. In an oriC- oriN+ background, yabA+ cells (strain MMB170, upper y-axis, black) and a yabA null mutant (strain AIG185, lower y-axis, red) were grown to mid-exponential phase, and Rok (A) and DnaA (B) were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC in the wild-type strain is set as 0, but a portion of oriC was deleted from these strains. These strains replicated from the heterologous origin oriN. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.
**Comparison of ChIP-seq binding profiles of YabA, DnaA, and Rok**

To better characterize the association of YabA with genomic regions bound by DnaA and Rok, we compared the binding profiles at high resolution. Specifically, we wondered if YabA was associated with DNA sequences adjacent to those bound by Rok and DnaA, or if YabA associated with the same sequences. We examined two categories of regions: those bound by only YabA and DnaA, and those bound by YabA, DnaA, and Rok.

**Chromosomal regions bound by YabA and DnaA.** DnaA associates with eight DnaA box cluster regions \((rpmH/dnaA, dnaA/dnaN, gcp/ydiF, yqeG/sda, ywlC/ywlB, ywcI/vpr, yydA/yycS,\) and \(trmE/jag\)), which each contain at least four DnaA-box motifs (Chapter 2 and (Breier and Grossman, 2009; Ishikawa et al., 2007)). YabA was also associated with most of these regions, and the degree of YabA enrichment met our peak-calling criteria for five of them \((rpmH/dnaA, dnaA/dnaN, ywcI/vpr, yydA/yycS,\) and \(trmE/jag;\) **Table 1**). Where DnaA and YabA co-associated, the centers of the estimated binding peaks coincided (within the 20-nt window used by SISSRs for peak detection). For proteins that bind to distinct nucleotide sequences within a given chromosomal region, their expected ChIP-seq binding peaks would be offset. If one protein does not directly bind DNA, but is associated with a chromosomal region via protein-protein interactions with a DNA-binding protein, the binding peaks would be expected to be identical. The essentially complete overlap between DnaA and YabA peaks indicates that the two proteins do not bind separate DNA sequences within the region, but rather that one protein (DnaA) binds DNA directly and the other (YabA) binds indirectly (via DnaA). The coincidence of the DnaA and YabA peaks is consistent with the presence of DnaA boxes and the DnaA-dependent binding of YabA to these regions (this work and (Merrikh and Grossman, 2011)).
We believe that the variable extent of YabA association with these DnaA-bound regions can be partially explained by limited detection but is likely also influenced by other factors. For these eight regions, YabA enrichment roughly correlated with DnaA enrichment, a proxy for the amount of DnaA bound (Fig. 9A). The two regions with the lowest detectable YabA association (ywlC/ywlB and yqeG/sda) also had correspondingly low DnaA association. Chromosomal position or copy number might also influence association of YabA with DnaA-bound regions, since YabA enrichment correlated well with chromosomal distance from oriC (Table 1). The sequence, spacing, and orientation of DnaA boxes within each region may also affect the ability of YabA to associate with DnaA at these regions.

![Figure 9](image)

**Figure 9. Relative enrichment of YabA, DnaA, and Rok at co-associated regions in wild-type cells.**
Each data point represents one chromosomal region. Enrichment values were determined using the peak-calling algorithm SISSRs. Selected regions are labeled.

(A) YabA enrichment vs. DnaA enrichment at seven of the eight DnaA box cluster regions, at which association of YabA depended on DnaA and did not depend on Rok. The sda region was omitted because it did not show detectable YabA enrichment (<2.2-fold).

(B) YabA enrichment vs. Rok enrichment at co-associated chromosomal regions other than DnaA box clusters. The minimum enrichment for inclusion in this dataset was three-fold.
Chromosomal regions bound by YabA, DnaA, and Rok. At the 44 chromosomal regions where YabA, DnaA, and Rok co-associated in a Rok-dependent manner, the binding peaks were closely coincident (Fig. 10). These observations, together with the dependence results, previous results for Rok and DnaA (Chapter 2), and in vitro binding of Rok to DNA (Albano et al., 2005), indicate that Rok directly binds DNA at these regions, whereas YabA and DnaA associate with these regions indirectly, via Rok.

We observed many chromosomal regions bound by Rok but not YabA. YabA might associate with only a subset of Rok-bound regions, or YabA might be present at all Rok-bound regions, but below the limit of detection at some of them. We favor the latter explanation because the relative enrichment of Rok closely correlated with the relative enrichment of YabA (Fig. 9B). That is, the regions with the strongest Rok enrichment were also associated with YabA, whereas the Rok-bound regions without detectable YabA enrichment had the lowest Rok enrichment. A similar correlation is true for Rok and DnaA (Chapter 2). Therefore, we believe that YabA likely associates with all Rok-bound regions to some extent, but many are below our limit of detection. We did not observe a direct interaction between YabA and Rok in vitro using a gel electrophoretic mobility shift (data not shown). YabA might not interact with Rok, or the interaction might not be apparent under the conditions tested.
Figure 10. ChIP-seq reads for selected regions bound by YabA, DnaA, and Rok in wild-type cells.

YabA, DnaA, and Rok were immunoprecipitated from wild-type cells (strain AG174). Each panel shows a magnified view of the data from Fig. 1 and Chapter 2. The selected chromosomal regions include (A) sunA, (B) sboA, and (C) yuzB/yutJ. Numbers of reads are shown on the y-axis. Reads mapping to the forward strand are shown on the upper y-axis in blue, and reads mapping to the reverse strand are shown on the lower y-axis in green. Chromosomal position is shown on the x-axis. The ChIP-seq point of symmetry indicates the inferred center of the binding site (Belitsky and Sonenshein, 2013). The center of each DnaA binding peak, as determined by SISSRs, is indicated by a bold vertical line. Positions of DnaA boxes are indicated by dashed red vertical lines and red points on each x-axis. DnaA boxes (upper row, red) and gene annotations (lower row, blue) are shown below the corresponding chromosomal positions.
Effects of *rok* and *yabA* on cell viability and protein levels

Rok is known to repress its own transcription (Hoa et al., 2002), and the *rok* promoter was associated with Rok, YabA, and DnaA in a Rok-dependent manner (*Table 2; Fig. 5; Chapter 2*). DnaA also regulates its own transcription (Ogura et al., 2001), and the *dnaA* promoter is associated with DnaA and YabA in a DnaA-dependent manner (Merrikh and Grossman, 2011). Regulation of *yabA* transcription has not been extensively studied. *dnaA* is essential (Kobayashi et al., 2003) and *rok* overexpression is reported to be lethal in *B. subtilis* (Hoa et al., 2002), so altered regulation of these genes by YabA and/or Rok might affect cell viability. We investigated the effects of *rok* or *yabA* null mutations on cell viability and protein levels of DnaA, YabA, and Rok. We found that loss of *rok* did not detectably affect cell viability, the relative concentrations of DnaA or YabA, or YabA protein stability. Loss of *yabA* did not detectably affect Rok protein stability. We did not detect an effect of *yabA* on Rok stability or relative protein levels, beyond the known effects of *yabA* on cell size.

**Viability.** Loss of *rok* did not detectably affect viability. For both wild-type and *rok* null strains, the number of colony-forming units (cfu) per ml during exponential growth was approximately $1.0 \times 10^8$ (*Fig. 11A*). Loss of *yabA* decreased the cfu/ml by a factor of two relative to wild-type (*Fig. 11A*). *yabA* null cells are longer than wild-type cells (Noirot-Gros et al., 2002) and have 40% more mass (Hayashi et al., 2005), so the lower cfu/ml in the *yabA* null mutant may simply reflect larger cell mass.
Figure 11. Viable cell density in rok null and yabA null mutants.
Wild-type cells (strain AG174, white bars), a rok null mutant (strain WKS1038, blue bars), and a yabA null mutant (strain AIG109, yellow bar) were grown to (A) mid-exponential phase or (B) stationary phase at 30 °C in minimal glucose medium. Serial dilutions were plated onto LB to determine colony-forming units per ml. One cfu is assumed to represent one cell. Cfu/ml measurements were normalized to OD. Data are averages of at least three biological replicates, and error bars represent the standard error.

Molecules per cell of DnaA, YabA, and Rok. We estimated the number of DnaA, YabA, and Rok molecules per cell in the wild-type strain using quantitative western blotting and the previously measured cfu/ml measurements. Our estimate of DnaA molecules/cell (~6,000; Fig. 12A) was comparable to previous estimates of ~3,000 molecules/cell in B. subtilis (Merrikh and Grossman, 2011) and ~2,000 molecules/cell in E. coli (Ali Azam et al., 1999). Our estimate of YabA molecules/cell (~8,000; Fig. 12C) was lower than a previous estimate of >20,000 molecules/cell (Merrikh and Grossman, 2011). We suspect that the previous measurement may have been an overestimate, due to the cross-reactivity of lysozyme in cell lysates with at least one type of secondary antibody (Experimental Procedures), plus the presence of a protein that migrates near YabA on a gel and cross-reacts with rabbit anti-YabA antibodies. This protein band does not correspond to YabA, as it is still present in a yabA null lysate. When this cross-reacting signal was included with the true YabA signal in measurements of YabA molecules/cell, the estimate was ~20,000 molecules/cell, similar to the previous value.
Figure 12. Relative YabA, DnaA, and Rok protein levels in wild-type and mutant strains. Wild-type cells (strain AG174, white bars), a rok null mutant (strain WKS1038, blue bars), and a yabA null mutant (strain AIG109, yellow bars) were grown to mid-exponential phase at 30 °C in minimal glucose medium, and protein levels of YabA (A, B), DnaA (C, D), and Rok (E, F) were measured by quantitative western blot. Molecules/cell measurements (A, C, E) represent protein levels normalized to OD and to the cfu/ml for each strain. Relative protein levels (B, D, F) represent protein levels normalized to OD (Experimental Procedures). Data are averages of at least three biological replicates, and error bars represent the standard error.
We calculated >30,000 Rok molecules/cell (Fig. 12E). This estimate is at least 10-fold higher than a previous figure (1,000 - 3,000 Rok molecules/genome), which was based on a myc-tagged version of Rok and a comparison to HBsu (Smits and Grossman, 2010). Other nucleoid-associated proteins are quite abundant, however, so a Rok concentration on the order of $10^4$ molecules/cell would not be unusual. H-NS has ~20,000 molecules/cell (Ali Azam et al., 1999; Hulton et al., 1990), HBsu has ~50,000 molecules/genome (Ragkousi et al., 2000), and HU in *E. coli* has ~40,000 molecules/cell (Ali Azam et al., 1999).

**Relative protein levels in different strains.** In comparing protein levels between different strains, we used two ways of normalizing protein levels: relative to number of cells and relative to OD (a proxy for total protein). The molecules/cell measurement can be influenced by differences in cell size, such as the larger mass of *yabA* null cells relative to wild-type cells (Hayashi et al., 2005). The OD-normalized protein measurement is expected to be more indicative of the intracellular concentration of a given protein, and the simpler calculation involves fewer sources of variation.

Loss of *rok* did not affect DnaA or YabA protein levels. Between the wild-type and *rok* null strains, there was no detectable difference in YabA molecules/cell (Fig. 12A), YabA relative protein concentrations (Fig. 12B), DnaA molecules/cell (Fig. 12C), or DnaA relative protein concentrations (Fig. 12D). As expected for larger cell mass, *yabA* null cells had two- to three-fold higher molecules/cell of DnaA (Fig. 12C) and Rok (Fig. 12E). Loss of *yabA* did not affect the relative protein concentration of Rok (Fig. 12F) but increased the relative protein concentration of DnaA by 40% (Fig. 12F). The effect on DnaA is consistent with a reported 40% increase in *dnaA* transcription in a YabA depletion mutant (Cho et al., 2008) and may
reflect the effects of YabA on replication initiation and consequently *dnaA* copy number and autoregulation.

**Protein stability.** Rok and YabA did not detectably alter the stability of each other. In wild-type cells, both were stable for at least four hours following translational inhibition with chloramphenicol (Fig. 13). Stability of Rok was not substantially different in a *yabA* null strain (Fig. 13A), nor was stability of YabA substantially different in a *rok* null strain (Fig. 13B).

**YabA, DnaA, and Rok in stationary phase**

Since Rok is part of regulatory networks that govern competence and the transition to stationary phase (Hoa et al., 2002), we examined YabA, DnaA, and Rok protein levels and genome-wide binding in stationary phase. The substantially lower protein levels for YabA and DnaA in stationary phase made it difficult to assess or interpret genome-wide binding patterns.

We measured viability and genome-wide binding (ChIP-seq) in wild-type and *rok* null cells at two timepoints: mid-exponential growth and early stationary phase. A representative growth curve is shown in Fig. 14A. In early stationary phase but not during exponential growth, the *rok* null mutant strain had slower growth (Fig. 14A; Appendix A) and approximately 6-fold lower cfu/ml (normalized to OD) than the wild-type strain (Fig. 14B). This reduction in viability is consistent with the *comK*-dependent growth defect of a *rok* null mutant strain (Appendix A). In wild-type cells, relative YabA protein levels decreased by a factor of two from exponential to stationary phase (Fig. 14B), and relative DnaA levels decreased by a factor of three (Fig. 14C). Rok levels remained similar between exponential and stationary phase (Fig. 14D).
Figure 13. Stability of Rok and YabA proteins.
Wild-type cells (strain AG174, closed circles), a yabA null mutant (strain AIG109, open circles), and a rok null mutant (strain WKS1038, open triangles) were grown to mid-exponential phase at 30 °C in minimal glucose medium and treated with chloramphenicol to block new protein synthesis. Protein levels of (A) Rok and (B) YabA were analyzed by quantitative western blot at the indicated timepoints. Volumes loaded on the gel were normalized to culture OD. Data represent individual biological experiments (one timecourse per strain). In (A) data are averages of triplicate technical measurements (each sample loaded on three wells of the same gel) and error bars represent the standard deviation of these technical replicates. In (B) YabA signal included a correction for background signal due to lysozyme, which runs at the same molecular weight and cross-reacts with the fluorescent secondary antibody (Experimental Procedures).
Figure 14. Relative YabA, DnaA, and Rok protein levels in exponential and stationary phase.
Wild-type cells (strain AG174, white bars or closed circles) and a rok null mutant (strain WKS1038, blue bars or open circles) were grown through exponential phase into stationary phase at 30 °C in minimal glucose medium. (A) ChIP-seq samples were taken at the timepoints indicated by red arrows on the growth curve. Protein levels of YabA (B), DnaA (C), and Rok (D) were measured by quantitative western blot. Relative protein levels represent protein levels normalized to OD (Experimental Procedures). Data are averages of at least three biological replicates, and error bars represent the standard error.

The reduced protein levels likely contributed to the low ChIP-seq signal for YabA and DnaA in stationary phase. YabA was not detectably enriched at any chromosomal regions in wild-type cells in stationary phase (Fig. 15A). In a rok null mutant in stationary phase, YabA was detected only at the five DnaA box cluster regions (Fig. 15A) where it was detected in exponentially growing rok null cells (Fig. 5A, B). In the rok null mutant, at least, the genome-wide association of YabA did not change between exponential and stationary phase. In
stationary-phase wild-type cells, DnaA was detectably enriched only at the *rpmH/dnaA* and *dnaA/dnaN* regions (Fig. 15B), which were the regions with the highest DnaA enrichment in exponentially growing wild-type cells (Chapter 2). We suspect that DnaA is present at additional regions in stationary phase but below the limit of detection. In stationary-phase *rok* null cells, DnaA was detected at all eight of the DnaA box cluster regions (Fig. 15B) where it was detected in exponentially growing *rok* null cells (Chapter 2). That is, in the *rok* null mutant, the genome-wide association of DnaA did not change between exponential and stationary phase.

In wild-type cells, Rok was detected at the same chromosomal regions in exponential and stationary phase (Fig. 16). Therefore, the lack of detectable YabA and DnaA enrichment at Rok-dependent regions in stationary-phase wild-type cells cannot be explained by a loss of Rok or a reduction in Rok protein levels (Fig. 14D). Although the ChIP-seq data for wild-type cells in stationary phase are difficult to interpret, we suspect that the binding profiles of YabA and DnaA are not substantially different between exponential and stationary phase, and that the reduction in protein levels prevents detection of YabA and DnaA at all but the highest enriched regions (i.e., the oriC region for DnaA). We also suspect that the stationary-phase *rok* null cells showed higher (and therefore detectable) YabA and DnaA enrichment than wild-type cells at the DnaA box cluster regions because YabA and DnaA could no longer associate with Rok-dependent regions and were redistributed to the DnaA box cluster regions. This type of chromosome-wide redistribution might be more evident or of more regulatory importance during stationary phase when YabA and DnaA levels are low.
Figure 15. Genome-wide binding of YabA and DnaA in stationary phase in wild-type and rok null mutant cells.

Wild-type cells (strain AG174, upper y-axis, black) and a rok null mutant (strain WKS1038, lower y-axis, red) were grown to early stationary phase, and YabA (A) and DnaA (B) were immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC is set as 0. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.
Figure 16. Genome-wide binding of Rok in exponential and stationary phase in wild-type cells.
Wild-type cells (strain AG174) were grown from exponential phase (upper y-axis, black) to early stationary phase (lower y-axis, red), and Rok was immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC is set as 0. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.

Discussion

YabA was initially identified in yeast two-hybrid screens for products that interact with the replication initiator DnaA and the processivity clamp DnaN, and found to be a negative regulator of replication initiation (Noirot-Gros et al., 2002). Subsequent studies of YabA have mainly focused on its role in replication control (Cho et al., 2008; Goranov et al., 2009; Hayashi et al., 2005; Merrikh and Grossman, 2011; Noirot-Gros et al., 2006; Scholefield and Murray, 2013; Soufo et al., 2008). YabA is associated with the replisome (Goranov et al., 2009; Hayashi et al., 2005; Noirot-Gros et al., 2006; Soufo et al., 2008) and with DnaA at regions around the chromosome where DnaA binds (Merrikh and Grossman, 2011).
Our results show that regulation and function of YabA may be more complex than previously thought. We found that YabA associated with many chromosomal regions bound by the nucleoid-associated protein Rok, and association of YabA with these regions depended on Rok, similar to recent findings for DnaA (Chapter 2). We also found that YabA did not require DnaA for association with these Rok-bound regions, nor did DnaA require YabA for association with these regions. We believe it is still unlikely that YabA binds to DNA directly, since all detectable YabA association was dependent on DnaA or Rok, and we did not detect YabA at any new regions in the absence of DnaA or Rok.

Together, the genome-wide binding profiles of YabA, DnaA, and Rok indicate that bound chromosomal regions fall into two categories: (1) regions bound by DnaA and often YabA, at which YabA association depends on the binding of DnaA to DnaA binding sites, and (2) regions bound by YabA, DnaA, and Rok, at which association of YabA and DnaA depends on Rok, which is thought to bind directly to DNA. We believe that the several regions that show association with Rok, but not YabA or DnaA, likely belong to the latter category and represent limited detection of YabA and DnaA. The chromosomal regions with Rok-dependent binding demonstrate that YabA association and DnaA association (this work; J. L. Smith and ADG, submitted) are not limited to chromosomal regions containing DnaA binding sites. Although the physiological role of YabA association with Rok-bound chromosomal regions is unclear, this association could represent a function for YabA beyond its known role in replication initiation.
Experimental Procedures

Strains and alleles

*B. subtilis* strains and relevant genotypes are listed in Table 3. Properties and construction of important alleles are described below.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype (reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG174</td>
<td><em>trpC2 pheA1</em> (wild-type, JH642) (Perego et al., 1988)</td>
</tr>
<tr>
<td>WKS1038</td>
<td>∆rok::cat (Smits and Grossman, 2010)</td>
</tr>
<tr>
<td>WKS624</td>
<td>yabA::[Pspank-3x-myc-yabA spc]</td>
</tr>
<tr>
<td>CAS37</td>
<td>yabA::[Pspank-3x-myc-yabA spc], ∆rok::cat</td>
</tr>
<tr>
<td>AIG200</td>
<td><em>trp+</em>, ∆[oriC dnaA dnaN]::spc, spoIII::[oriN repN kan], amyE::[PxyLA-dnaN cat] (Goranov et al., 2005)</td>
</tr>
<tr>
<td>AIG109</td>
<td><em>trp+</em>, ∆yabA::cat, oriC+ (Goranov et al., 2009)</td>
</tr>
<tr>
<td>MMB170</td>
<td>*trp+, yabA+, oriC-S, spoIII::[oriN repN kan] (Berkmen and Grossman, 2007)</td>
</tr>
<tr>
<td>AIG185</td>
<td>*trp+, ∆yabA::cat, oriC-S, spoIII::[oriN repN kan] (Goranov et al., 2009)</td>
</tr>
<tr>
<td>CAL2074</td>
<td>∆[oriC dnaA]::[oriN repN Ppen-2028-dnaN cat]</td>
</tr>
<tr>
<td>CAL2083</td>
<td>∆[oriC dnaA]::[oriN repN Ppen-2028-dnaN cat], lacA::[Pspank-dnaA tet]</td>
</tr>
<tr>
<td>CAS147</td>
<td>yabA::[Pspank-3x-myc-yabA spc::mls], ∆[oriC dnaA]::[oriN repN Ppen-2028-dnaN cat]</td>
</tr>
<tr>
<td>CAS190</td>
<td>yabA::[Pspank-3x-myc-yabA spc::mls], ∆[oriC dnaA]::[oriN repN Ppen-2028-dnaN cat], lacA::[Pspank-dnaA tet]</td>
</tr>
</tbody>
</table>

Table 3. *B. subtilis* strains used in this study.

All strains are derived from AG174 and contain the *trpC2 pheA1* alleles unless otherwise indicated.

∆rok::cat. The *rok* open reading frame was replaced with a chloramphenicol resistance cassette (*cat*) by long-flanking homology PCR to generate strain WKS1030 (Smits and Grossman, 2010). Strain WKS1038 was the product of backcrossing genomic DNA from WKS1030 into wild-type AG174 cells.

yabA::3x-myc-yabA. The *yabA* open reading frame was replaced with a copy of *yabA* encoding an N-terminal fusion with three copies of the myc epitope, under control of the LacI-repressible, IPTG-inducible promoter Pspank. Strains CAS147 and CAS190 contained a
derivative of this allele in which the spectinomycin resistance gene was disrupted with an MLS (macrolide-lincosamide-streptogramin B) resistance gene (spc::mls).

\( \Delta \text{oriC-dnaA-dnaN} \). Strains AIG200 and TAW5 contain a deletion-insertion, in which dnaA and most of dnaN were replaced with a spectinomycin resistance cassette (spc) by long-flanking homology PCR (Goranov et al., 2005; Merrikh and Grossman, 2011). Replication was supported by insertion of the heterologous origin oriN and its initiator repN near oriC at spoIIIJ (Berkmen and Grossman, 2007; Hassan et al., 1997; Moriya et al., 1997). dnaN was expressed from PxylA at amyE. These strains still contained the dnaA promoter region, including the cluster of DnaA binding sites. Strains MMB170, AIG109, AIG185, AIG200, and TAW5 contained a deviation in the ypjG-hepT region, containing tryptophan biosynthesis genes, as described previously (Berkmen and Grossman, 2007).

Strain CAL2074 and its derivatives contain a deletion-insertion in which dnaA and flanking regions are replaced with a product (generated by isothermal assembly (Gibson et al., 2009)) containing a chloramphenicol resistance cassette, oriN and repN, and a promoter driving expression of dnaN. The chloramphenicol resistance cassette (including the transcription terminator with cat) was inserted at the left end of oriC, upstream of rpmH. oriN and repN were inserted upstream of this cassette such that oriN-repN, cat, and rpmH were co-directional. A derivative of the constitutive promoter Ppen (Ppen-2028, C. Lee, unpublished data) was cloned upstream of dnaN. Ppen is derived from the B. licheniformis penicillinase gene and drives lacI expression on the integration vector pDR110. Ppen-2028 carries mutations relative to Ppen (Ppen-2028 sequence in lowercase) between the putative -35 and -10 sequences (underlined): 5’- TTGCATTTAttcggtggtcGTAATACTTCCA-3’.
oriC-S. The DnaA binding sites and part of the DUE are deleted, and these strains replicate from oriN inserted at spoIIIJ, as described for strains AIG200 and TAW5.

**Media and growth conditions**

Unless otherwise specified, all strains were grown at 30 °C in S7 defined minimal medium buffered with 50 mM MOPS (Jaacks et al., 1989) and containing 1% glucose, 0.1% glutamate, trace metals, 40 µg/ml tryptophan, and 40 µg/ml phenylalanine. For growth of AIG200 and TAW5, glucose was replaced with 1% arabinose, and 0.5% xylose was used to induce expression of DnaN from the xylose-inducible promoter Pxyl. To induce expression of DnaA or 3x-myc-YabA from the LacI-repressible, IPTG-inducible promoter Pspank, 0.1 mM IPTG was added.

**ChIP-seq and ChIP-qPCR**

Immunoprecipitations were performed with anti-DnaA, anti-YabA, and anti-Rok rabbit polyclonal antibodies, or with mouse monoclonal anti-myc antibodies (Invitrogen) essentially as described (Lin and Grossman, 1998).Briefly, exponentially growing cells were treated with 1% formaldehyde, and the cross-linked lysates were sonicated to shear the DNA. Immunoprecipitations were performed by incubating the cross-linked lysates with antibodies for at least 2 hrs at room temperature, followed by incubation with Protein A sepharose beads for at least 1 hr at room temperature. A control sample of non-immunoprecipitated lysate was incubated under the same conditions. Cross-linked DNA-protein complexes were washed and eluted from the beads, followed by cross-link reversal at 65 °C overnight. Samples were then
treated with proteinase K at 37 °C for at least 2 hrs, and DNA was recovered using the Qiagen PCR purification kit.

High-throughput sequencing or real-time quantitative PCR was used to determine the relative enrichment of chromosomal regions. qPCR was performed essentially as described in (Merrikh and Grossman, 2011). Primer sets amplified chromosomal regions that were representative of DnaA-dependent YabA association (e.g., oriC) or Rok-dependent YabA association (e.g., sboA). A control region, yhaX, was used for normalization, as it is not known to be bound by YabA, Rok, or DnaA. Each well in a 96-well white qPCR plate (Qarta Bio) contained 1X LightCycler 480 SYBR Green I Master Mix (Roche), 2.5 μM forward primer, 2.5 μM reverse primer, and template DNA in a final volume of 20 μl. Template DNA was used at a final reaction concentration diluted 20X from the Qiagen PCR purification elution. Three technical replicates (qPCR reactions) were performed for each biological replicate.

Crossing-point cycle (Cp) values were calculated from melting curves using the LightCycler 480 software. Cp values (on a logarithmic scale) were converted to relative DNA concentrations (on a linear scale), using a standard curve generated by dilutions of non-immunoprecipitated (total) DNA. Technical triplicates were averaged to obtain a relative DNA concentration for each PCR product. Enrichment was normalized to yhaX and to total DNA: (IP target concentration/IP yhaX concentration)/(total target concentration/total yhaX concentration).

**Antibody specificity**

Specificity of the anti-DnaA and anti-Rok antibodies was confirmed previously (Chapter 2). We assessed the specificity of the anti-YabA antibody by performing ChIP-seq of YabA in a yabA null strain. We used a wild-type background as well as a strain background in
which replication initiates from a heterologous origin, oriN, using its cognate initiator, RepN (Hassan et al., 1997; Moriya et al., 1997). This strain background was used to control for the effects of YabA on replication (Goranov et al., 2009) and the possible effects of altered replication on YabA binding. In these strains, oriN was inserted near oriC, and oriC was inactivated by a small deletion in the DNA unwinding element (Berkmen and Grossman, 2007; Kadoya et al., 2002).

The anti-YabA negative control did not show enrichment in the oriC+ background (Fig. 17A), but in the oriC- oriN+ background, there was some enrichment at several regions with Rok-dependent YabA association (Fig. 17B). Three chromosomal regions were enriched (using a five-fold cutoff) in the oriC- oriN+ yabA null mutant: ppsA (5.26-fold), sunA (8.29-fold), and sboA (5.06-fold) (Fig. 17B). All of these regions were also bound by Rok (Chapter 2). Three additional regions showed enrichment over three-fold (rok (a Rok-bound region), vpr cds, and yxkO), but vpr cds and yxkO appeared to be artifacts of the peak-calling algorithm, as they did not correspond to increases in ChIP-seq coverage.

The simplest explanation is that the anti-YabA antibody also cross-reacted with Rok. The degree of YabA enrichment in the negative control, however, did not quantitatively account for all the enrichment observed in the wild-type ChIP-seq experiment at ppsA (24.79-fold), sunA (40.84-fold), and sboA (29.85-fold). It is also formally possible that RepN (the cognate initiator protein for oriN), or some native protein preferentially expressed in the oriN background, associated with the Rok-dependent DnaA/YabA-bound regions and cross-reacted with the anti-YabA antibody.
Figure 17. Specificity of YabA immunoprecipitations.
The indicated strains were grown to mid-exponential phase, and YabA was immunoprecipitated after cross-linking with formaldehyde. ChIP-seq coverage, normalized to the total number of reads in each sample, is plotted on the y-axis relative to chromosomal position on the x-axis. The position of oriC in the wild-type strain is set as 0, but a portion of oriC is deleted from strains MMB170 and AIG185. These strains replicated from the heterologous origin oriN. Parallel slash marks indicate truncation of y-values for clarity. Data represent single biological experiments.

(A) In an oriC+ background, YabA immunoprecipitation was compared between the wild-type strain (AG174, upper y-axis, black) and a yabA null mutant (AIG109, lower y-axis, red).

(B) In an oriC- oriN+ background, YabA immunoprecipitation was compared between a yabA+ strain (MMB170, upper y-axis, black) and a yabA null mutant (strain AIG185, lower y-axis, red). The presence of chromosomal regions from YabA immunoprecipitations in the yabA null strain (red asterisks) indicates nonspecific binding (cross-reactivity) of the anti-YabA antibody.
We confirmed the YabA association results by immunoprecipitating YabA using an N-terminal epitope tag instead of the cross-reactive antibody to the native protein. We introduced a second, IPTG-inducible, 3x-myc-tagged copy of the yabA gene at the wild-type yabA locus and immunoprecipitated YabA using a monoclonal anti-myc antibody. In these experiments, we observed association of myc-YabA with the same representative chromosomal regions as observed with immunoprecipitation of untagged YabA with the anti-YabA antibodies (Fig. 18A, B). This association was dependent on the 3x-myc tag (Fig. 18A, B) and on yabA (Fig. 18C).

We used anti-myc ChIP-qPCR to confirm YabA association with two of the previously identified regions (sboA and yuzB/yutJ, representing high and low YabA enrichment by ChIP-seq, respectively), as well as with the known DnaA-dependent region rpmH/dnaA (PdnaA, also labeled as oriC). We then measured dependence on Rok and on DnaA by introducing the 3x-myc-tagged yabA gene into the appropriate strain backgrounds. We confirmed that 3x-myc-YabA association with the sboA and yuzB/yutJ regions was largely Rok-dependent, as enrichment decreased 6.5-fold and 8-fold, respectively, in the rok null strain (Fig. 19A). Similarly, the 3x-myc-YabA results recapitulated the DnaA-independent association of YabA with Rok-bound, Rok-dependent regions (sboA, sunA, yybN, yjcM/yjcN, iscS/braB, and yuzB/yutJ) and the DnaA-dependent association of YabA with DnaA box cluster regions (PdnaA and ywcI) (Fig. 19B). We also confirmed by ChIP-qPCR that DnaA and Rok were enriched at the expected regions in these strains (data not shown).
Figure 18. Verification of YabA association with chromosomal regions.

(A, B) Specificity of the anti-myc antibody in an oriC+ strain background. Wild-type cells (strain AG174) and a strain expressing 3x-myc-yabA (WKS624) were grown to mid-exponential phase, and the indicated immunoprecipitations were performed after cross-linking with formaldehyde. Protein association with the oriC (A) and sboA (B) regions was measured by qPCR.

(C) Specificity of the anti-myc antibody in an oriC- oriN+ strain background. A yabA+ strain (MMB170), and a yabA null mutant (strain AIG185) were grown to mid-exponential phase, and myc was immunoprecipitated after cross-linking with formaldehyde. Association of myc with the indicated regions was measured by qPCR.
Figure 19. Verification of Rok-dependent YabA association with chromosomal regions.

(A) Rok-dependent association of myc-YabA with chromosomal regions. Strains expressing 3x-myc-yabA in a rok+ background (strain WKS624) and a rok null background (strain CAS37) were grown to mid-exponential phase, and myc was immunoprecipitated after cross-linking with formaldehyde. Association of myc-YabA with the representative Rok-dependent regions sboA and yuzB/yutJ was measured by qPCR.

(B) Specificity of the anti-myc antibody in a rok null mutant (strain WKS1038) and a yabA null mutant (strain AIG109). Strains were grown to mid-exponential phase, and myc was immunoprecipitated after cross-linking with formaldehyde. Association of myc with the indicated regions was measured by qPCR.
**High-throughput sequencing analysis**

We obtained approximately 7-50 million 40-nt reads for each sample. We mapped the reads to the *Bacillus subtilis* strain AG174 genome using bwa (Burrows-Wheeler Aligner) for single-end short reads (Li and Durbin, 2009) and allowing a maximum number of alignments (n) of 2. To make comparisons across samples, we normalized the number of reads at each chromosomal position to the total number of reads for that sample. To calculate coverage at each base pair on the chromosome, we computationally extended each read by the estimated average fragment length of 250 bp (J. L. Smith, ADG, submitted).

We used SISSRs (Jothi et al., 2008) to identify enriched regions in each ChIP sample. We compared each ChIP sample to the corresponding non-immunoprecipitated (total) DNA sample. We used the following parameters: F (fragment length) = 250, e (sensitivity) = 1, m (fraction mappable) = 1, w (window) = 20, E (required number of reads) = 1 per million sample reads, and L (maximum fragment length) = 400. We then optimized p (P-value) such that the enrichment fold cutoff was approximately 3. From these candidate enriched regions, we selected those with an enrichment of at least five-fold.

**Protein stability**

Cells were grown in minimal glucose medium at 30 °C to mid-exponential phase and collected at OD ~0.4. Chloramphenicol was added to a concentration of 200 μg/ml, and culture timepoints were taken. Although some strains contained a chloramphenicol resistance cassette, growth at the high chloramphenicol concentration was substantially slowed, indicating that there was still a bacteriostatic effect. Strains with resistance cassettes showed similar growth inhibition and protein stability as the non-resistant wild-type strain, indicating that
chloramphenicol was similarly effective against these strains. Each culture sample was boiled 5 min to kill the cells, and the pellet was concentrated 10X. Lysozyme was added to 0.2 mg/ml and incubated at 37 °C for 10 min. The sample was boiled with SDS-PAGE sample buffer for 5 min. The samples were run on a 15% SDS-PAGE gel, alongside samples of purified Rok or YabA, and transferred to a PVDF membrane. The volume loaded in each lane of the gel was normalized to the culture OD to represent an equal number of cells per lane.

Western blotting was performed as in Appendix A. All steps were performed at room temperature. Briefly, blots were incubated with Odyssey Blocking Buffer for 1 hr, followed by primary antibody (rabbit anti-Rok Test Bleed #2, 1:5,000 in Odyssey Blocking Buffer + 0.2% Tween) for 1 hr. Blots were washed with PBST (phosphate-buffered saline + 0.2% Tween) for at least 4 x 5 min and then incubated for 30 min with secondary antibody (LiCor dye 800 goat anti-rabbit 1:10,000 in Odyssey Blocking Buffer + 0.2% Tween). Blots were imaged and quantitated on a LiCor scanner.

The anti-rabbit secondary antibody was found to cross-react with lysozyme, which runs on a gel at the same molecular weight (14 kDa) as YabA. Therefore, lysozyme contributed to the fluorescence signal in the YabA western blot. YabA fluorescence intensities were corrected by estimating and subtracting the lysozyme signal for each gel lane. Estimated lysozyme signal (per lysate volume loaded on the gel) was calculated from the western blot of Rok from wild-type lysates. This blot was not incubated with anti-YabA antibody, so the intensity of the 14-kDa band was assumed to be due entirely to the secondary antibody recognizing lysozyme. For subsequent western blots involving YabA, the issue of lysozyme background signal was avoided by lysing the cells via sonication.
Relative protein levels

Cells were grown in minimal glucose medium at 30 °C to mid-exponential phase and collected at OD ~0.5 or grown to stationary phase and collected at OD 6.0 (wild type) or 2.0 (rok null mutant). Serial dilutions were plated on LB to determine cfu/ml. A culture sample of 15 ml was pelleted, and all but 1 ml supernatant was removed. Samples were frozen at -80 °C and resuspended with the addition of 14 ml TE (10 mM Tris pH 8.0, 10 mM EDTA) with protease inhibitors (AEBSF). The OD of the resuspension was measured in triplicate. Cells were lysed by sonication rather than lysozyme treatment because lysozyme cross-reacts with the LiCor goat anti-rabbit 800 secondary antibody used for quantitation. Samples were sonicated on ice for 6 min per sample, in bursts of 0.3 sec on/off. A 450-μl aliquot was taken, and 50 μl of TE plus protease inhibitors (AEBSF) was added. Equal volumes of each sample were analyzed on a 15% SDS-PAGE gel, along with standards of purified protein (DnaA, His6-YabA from Houra Merrikh, Rok from Janet Smith). Most samples were loaded on the gel in duplicate. Samples were imaged and quantitated on the LiCor scanner. Protein intensities from the lysates were compared to serial dilutions of the purified protein standards. Molecules-per-cell estimates were calculated by normalization to culture OD and comparison to the cfu/ml for each strain. Relative protein levels (between-strains comparisons) were made by directly comparing the western blot signal after normalization for OD. Western blot signals typically differed by less than a factor of two.
Acknowledgments

We thank Janet Smith and Catherine Lee for useful discussions. This work was supported, in part, by the Department of Defense, Air Force Office of Scientific Research, National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a, to CAS and Public Health Service grant GM41934 to ADG.
Chapter 4

Conclusions and Perspectives
The bacterial replication initiator DnaA is highly regulated to coordinate DNA replication with the cell cycle and other processes. In *Bacillus subtilis*, the several known regulators of replication initiation inhibit the binding of DnaA to DnaA boxes at the origin of replication. In this work, I addressed the localization, function, and regulation of DnaA and one of its regulators, YabA, at chromosomal regions beyond *oriC* and without canonical DnaA boxes. I found that DnaA and its regulator YabA are recruited to chromosomal regions bound by the nucleoid-associated protein Rok, and they depend on Rok for this association. Rok directly interacts with DnaA, and I identified regions of each protein that are sufficient for this interaction. Although YabA depends on DnaA for association with certain chromosomal regions, YabA association with Rok-bound regions did not depend on DnaA. YabA might associate with Rok directly, under conditions that were not captured by the *in vitro* assay in this work, or YabA might be recruited by a different, unidentified adaptor.

I described the association of DnaA with two types of chromosomal regions (Fig. 1). In addition to the canonical mode of binding to DnaA boxes, DnaA can associate with many, if not all, of the A+T-rich, often horizontally acquired genomic regions bound by Rok. At least some of these genes are directly regulated by Rok. I have shown that DnaA also affects the regulation of certain Rok-regulated genes, although indirect regulation via the sporulation checkpoint regulator Sda is also important (T.A. Washington, J.L. Smith, and ADG, in preparation).

Most, if not all, of the previous work on DnaA has focused on its sequence-specific binding to DnaA boxes. This work builds on findings from a genome-wide comparison of DnaA binding *in vivo* and *in vitro* (J.L. Smith and ADG, submitted) and illustrates the types of insights that can be gained from comparing protein binding in these two contexts. At Rok-dependent chromosomal regions, for example, association of DnaA does not even require the ability to bind
DNA. *In vivo* binding can be positively or negatively affected by other DNA-binding proteins, protein-protein interactions, and the three-dimensional structure and topology of the chromosome, whereas *in vitro* binding reflects direct interactions between purified proteins and DNA that is typically linearized, fragmented, or of a defined sequence. Comparing these binding profiles can distinguish between direct and indirect binding and can indicate potential interaction partners at indirect binding regions, even for proteins as long-studied and highly conserved as DnaA.

The relationships among DnaA, YabA, and Rok indicate potential mechanisms for replication to influence or be influenced by chromosome architecture, gene expression, or other aspects of Rok function. The results in this work suggest two general avenues of future investigation: better defining the physiological functions of the DnaA-YabA-Rok interactions in *B. subtilis* and exploring parallels and generalizations in other organisms.

**Figure 1. Model for DnaA association with different types of chromosomal regions.**
(A) At Rok-dependent regions, Rok binds directly to A+T-rich DNA. DnaA associates with Rok via protein-protein interaction that does not require the DNA-binding activity of DnaA. YabA depends on Rok but associates with Rok independently of DnaA. DnaA boxes may or may not be present. Regions in this category include genes that are co-regulated by DnaA and Rok. (B) At regions of direct binding, DnaA binds to DNA via DnaA boxes. Association of YabA, if any, depends on DnaA. Regions in this category include the eight DnaA box cluster regions, such as *oriC*. 
Functions of DnaA-YabA-Rok interactions in *B. subtilis*

The interaction between DnaA and Rok suggests a possible role for Rok in replication initiation, but I did not detect a replication phenotype of a *rok* null mutant. Given the multiple regulatory mechanisms affecting replication initiation and multiple genetic networks involving *rok*, it is possible that a replication phenotype of Rok could be masked. For example, in *B. subtilis*, the chromosome partitioning protein and replication regulator Soj does not have a detectable partitioning phenotype when deleted because SMC (structural maintenance of chromosomes) masks the effect of Soj on partitioning. In the absence of *smc*, however, a *soj* null mutant has a drastic effect, a ten-fold increase in anucleate cells (Lee and Grossman, 2006). In this case, a chromosome organization protein masks the chromosome-partitioning function of a replication regulator. Perhaps, conversely, in the case of Rok, a known or unknown replication protein might mask the replication-regulating function of a chromosome architecture protein.

As a nucleoid-associated protein with functional similarity to H-NS of *E. coli*, Rok likely affects the three-dimensional structure of the chromosome. This spatial structuring, perhaps maintained or regulated by Rok, might be important for the function of DnaA in replication initiation, the transcriptional response to replication stress, gene expression, or other functions. Rok might mediate long-range interactions, similar to H-NS (Dame et al., 2002; Noom et al., 2007; Wang et al., 2011), by looping together A+T-regions of the chromosome. With ongoing advances in chromosome conformation capture (3C) and its derivatives (Gavrilo et al., 2009; Stadhouders et al., 2013), it is becoming easier to study the chromosome as a dynamic three-dimensional entity, and these types of approaches may be instrumental in fully understanding the nucleoid-associated function of Rok.
The evolutionary history of Rok suggests that its original function, before being integrated into the *B. subtilis* genome, may have been contributed to plasmid conjugation (Singh et al., 2012). Other plasmid-encoded H-NS-like proteins, such as the *Shigella* protein Sfh, mimic and often complement the function of H-NS. This “stealth” function avoids the detrimental consequences of H-NS titration on global gene expression (Doyle et al., 2007). If the ancestral Rok protein also interacted with DnaA, the ability to directly affect a widely conserved and essential protein may have allowed a wide host range for the plasmid. pLS20, the conjugative plasmid recently discovered to encode a small *rok* homolog (Singh et al., 2012), is presumed to replicate via DnaA, as the plasmid contains DnaA boxes in its origin of replication (Meijer et al., 1995; Rösch et al., 2014; Tanaka et al., 2005). Perhaps Rok-LS20 and/or the host Rok regulate the recruitment of DnaA (or regulators such as YabA) to the plasmid, or they might affect the relative distribution of DnaA on the chromosome versus the plasmid. The recruitment of DnaA could be not only beneficial but essential if *rok* is (or was) found on plasmids with weak initiators. In *E. coli*, the nucleoid-associated proteins IHF and HU affect DnaA binding to DnaA boxes at oriC and the titration locus *datA* by bending DNA (Dixon and Kornberg, 1984; Donczew et al., 2014; Nozaki et al., 2009; Polaczek, 1990). Rok might have a similar function, perhaps with different topological effects on the plasmid and host chromosome. Transcriptional autoregulation and perhaps cross-repression by the chromosomal and plasmid Rok proteins might contribute to homeostasis or the timing of replication. These models have not been investigated, but it could be interesting to investigate the potential roles of plasmid-encoded and chromosome-encoded Rok proteins in plasmid replication.

The structural basis for the protein-protein interaction of DnaA and Rok has yet to be precisely defined. I found that the N-terminal region of DnaA (Domains I-II-III) was sufficient
for association with Rok in vivo and in vitro. A clear next step would be to analyze the interaction of Rok with additional DnaA truncation mutants, as well as with YabA mutants. If the Rok-interacting region of DnaA falls within a highly conserved region or structural feature, DnaA proteins from additional species might be expected to interact with B. subtilis Rok, its homologues, or proteins with similar properties.

**DnaA, YabA, and Rok in other organisms**

Rok is functionally analogous to H-NS and a diverse set of H-NS-like proteins with strikingly convergent functions. The H-NS analogues show remarkable interchangeability and cross-species complementation (Dorman et al., 1999), including one instance of a mammalian protein complementing a Mycobacterium regulator (Timchenko et al., 1996). This versatility suggests that the common properties of H-NS-like proteins are easily evolved and perhaps arise from fairly generic biophysical or biochemical properties. A key set of experiments would be to test for protein-protein interactions between DnaA and H-NS (or its analogues) in organisms other than B. subtilis, starting with E. coli. It would also be interesting to test whether Rok can complement the function of H-NS and its analogues with respect to DNA binding, gene expression, and interaction with other proteins such as heterologous DnaAs or H-NS antagonists.

The Hha-YmoA family of H-NS-interacting proteins shares some loose parallels with YabA. The 8.5-kDa Hha protein of E. coli does not bind DNA on its own, similar to the 14-kDa YabA. Hha depends on H-NS for association with the hemolysin promoter and forms a heteromeric complex with H-NS (Nieto et al., 2000, 2002). The Hha family member Cnu in E. coli binds part of oriC and forms a complex with H-NS, but its role, if any, in replication initiation appears cryptic, since cnu or hha single mutants do not have a detectable replication
phenotype and under-replication in the double mutant is modest (Kim et al., 2005). Although Hha proteins are not known to be found in Gram-positive organisms (Madrid et al., 2007), YabA might share functional or structural characteristics with this group of regulators. Given the widespread functional convergence of H-NS-like proteins from evolutionarily distant species, I suspect that the function and properties of Rok, such as its interaction with the conserved initiator DnaA, are likely to extend to other members of the diverse and versatile group of H-NS-like nucleoid-associated proteins.
Appendix A

Growth and replication phenotypes of a rok null mutant
Summary

I present supporting information describing the negligible direct effects of rok on growth and replication initiation in B. subtilis. There was no detectable difference in growth between wild-type and rok null cells under exponential growth. The growth defect of the rok null mutant during the transition to stationary phase was dependent on comK. There was no detectable replication phenotype of rok null cells under a variety of conditions: during different growth phases, in rich or minimal media, under changes in DnaN levels (which led to altered replication initiation and were expected to perturb free YabA levels), or in a yabA null background (to control for the dual effects of Rok on DnaA and YabA). If Rok does affect replication initiation, the role of Rok appears to be cryptic and may be compensated by other regulators of replication initiation.

Introduction

The Rok-dependent recruitment of DnaA and YabA to chromosomal regions bound by Rok suggested a possible mechanism for Rok to affect replication initiation. Rok might modulate the free concentrations of DnaA and YabA in the cell and thereby couple their activity to chromosome architecture, competence, or other aspects of cellular physiology affected by Rok. I found no detectable effect of a rok null mutation on replication initiation of cells under various growth conditions and in genetic backgrounds that were expected to sensitize cells to perturbations in DnaA and YabA levels.
Results and Discussion

The growth defect of *rok* null cells upon entering stationary phase depends on *comK*

During exponential growth, there was no detectable difference in growth between wild-type and *rok* null cells. Doubling times for both strains were approximately 75 min at 30 °C in minimal glucose medium (Fig. 1). Similarly, no growth differences between wild-type and *rok* null cells were observed during exponential growth in minimal arabinose medium (Fig. 2A; doubling time approximately 115 min) or LB medium (Fig. 2B; doubling time approximately 45 min).

Upon entry to stationary phase (around OD 1.0 in both strains), *rok* null cells slowed growth more gradually than wild type cells in minimal glucose medium (Fig. 3A) and in minimal succinate medium (Fig. 3B). The *rok* null cells reached a maximum OD much later than wild-type cells. When stationary cultures were diluted into fresh media, the *rok* null mutant showed a longer lag, although this comparison might reflect the lower ODs of the stationary phase *rok* null inocula rather than a recovery phenotype per se (Fig. 4A). No recovery defect (Fig. 4B) or stationary phase defect (Fig. 4C) was observed in LB medium, perhaps because nutrient-rich conditions do not promote competence development, a developmental switch relevant to Rok. Since Rok represses *comK* and competence development occurs around the break from exponential growth (Hamoen et al., 2003), the growth defect upon entry to stationary phase might relate to *comK* derepression in the *rok* null mutant.
Figure 1. No growth phenotype of a *rok* null mutant during exponential growth in minimal glucose medium.

Wild-type cells (strain AG174, circles) and a *rok* null mutant (strain WKS1038, triangles) were grown at 30 °C in minimal glucose medium from (A) early to (B) late exponential phase. Each curve represents an independent biological replicate. All replicates in a given experiment (i.e., within each panel) were diluted to an equal starting density from exponentially growing overnight cultures. Time is arbitrary and relative to a dilution point specific to each experiment.
Figure 2. No growth phenotype of a *rok* null mutant during exponential growth in minimal arabinose or LB medium.

Wild-type cells (strain AG174, circles) and a *rok* null mutant (strain HM57, triangles) were grown at 30 °C in (A) minimal arabinose medium or (B) LB medium. Each curve represents an independent biological experiment. All cultures in a given experiment (i.e., within each panel) were diluted to an equal starting density from exponentially growing overnight cultures. Time is arbitrary and relative to a dilution point specific to each experiment. See Table 1 for an explanation of the HM57 allele.
Figure 3. Growth defect of a *rok* null mutant upon entry to stationary phase.
Wild-type cells (strain AG174, circles) and a *rok* null mutant (strain WKS1038, triangles) were grown at 30 °C in (A) minimal glucose medium or (B) minimal succinate medium. Each curve represents an independent biological replicate. All replicates in a given experiment (i.e., within each panel) were diluted to an equal starting density from exponentially growing overnight cultures. Time is arbitrary and relative to a dilution point specific to each experiment.
Figure 4. No growth defect of a *rok* null mutant during recovery from stationary phase or in LB medium.

Wild-type cells (strain AG174, circles) and a *rok* null mutant (strain WKS1038, triangles) were grown at 30 °C in (A) minimal glucose medium or (B) LB medium after dilution to OD 0.025 from a stationary phase culture having the OD indicated at 0 hrs. (C) Wild-type cells (strain AG174, circles) and a *rok* null mutant (strain HM57, triangles) were grown at 30 °C in LB medium after dilution to an equal starting density from exponentially growing overnight cultures. Each curve represents an independent biological replicate. Time is arbitrary and relative to a dilution point specific to each experiment.
There was no growth defect of \textit{rok} null cells relative to \textit{rok+} cells in a \textit{comK} null background, indicating that the growth phenotype of the \textit{rok} null mutant was a consequence of downstream \textit{comK} regulation. During growth and entry into stationary phase, the doubling time and maximum OD were similar between \textit{rok} null and \textit{rok+} cells in a \textit{comK} null background in minimal glucose medium (\textbf{Fig. 5A}). The growth curves in the \textit{comK} null mutant were similar to those of the \textit{comK+} wild-type strain (\textbf{Fig. 3A}), indicating that \textit{comK} upregulation confers a stationary phase growth defect. Likewise, there was negligible growth difference between the \textit{rok} null and \textit{rok+} strains in a \textit{comK} null background upon recovery from stationary phase (\textbf{Fig. 5B}). These results indicate that \textit{comK} regulation accounts for the stationary phase growth phenotype of a \textit{rok} null mutant.

\textbf{Rok does not detectably affect replication initiation}

I found no detectable effect of a \textit{rok} null mutation on replication initiation under various conditions, including media and mutations that were expected to enhance detection of altered replication.

\textbf{Growth phases.} Using the chromosomal origin/terminus ratio (\textit{oriC/terC}) as a measure of replication initiation, I found no observable difference between wild-type and \textit{rok} null cells during different growth phases. \textit{oriC/terC} ratios were similar between the strains during exponential growth in minimal glucose or minimal succinate medium (\textbf{Fig. 6}). To control for the \textit{comK}-dependent growth defect of \textit{rok} null cells in stationary phase, I used a \textit{comK} null background for experiments involving later growth phases. In minimal glucose, there was still no difference in \textit{oriC/terC} between \textit{rok} null and \textit{rok+} cells during the transition to stationary phase (\textbf{Fig. 7A}) or during recovery from stationary phase (\textbf{Fig. 7B}).
Figure 5. No growth defect of a rok null mutant in a comK null background.

In a comK null background, a rok+ strain (HM61, circles) and a rok null strain (HM71, triangles) were grown at 30 °C in minimal glucose medium (A) after dilution to an equal starting density from exponentially growing overnight cultures or (B) after dilution to OD 0.025 from a stationary phase culture having the OD indicated at 0 hrs. Each curve represents an independent biological replicate. Time is arbitrary and relative to a dilution point specific to each experiment.
Figure 6. No replication phenotype of a rok null mutant during exponential growth.
Wild-type cells (strain AG174, white bars) and a rok null mutant (strain WKS1038, blue bars) were grown to mid-exponential phase at 30 °C in (A) minimal glucose medium or (B) minimal succinate medium. Origin/terminus ratio (oriC/terC) was measured by a qPCR assay of genomic DNA purified from the indicated strains. Primers amplified genomic regions of oriC and terC. Data are averages of biological triplicates and error bars represent the standard error. Each value (oriC or terC) was measured in triplicate, and oriC/terC for each biological replicate represents the ratio of these technical averages. qPCR values were fitted to a standard curve prepared from genomic DNA of a dnaB temperature-sensitive mutant, which has an oriC/terC of one at the restrictive temperature (Experimental Procedures). oriC/terC values less than one likely indicate the use of an imperfect standard having oriC/terC greater than one.
Figure 7. No replication phenotype of a rok null mutant in a comK null background.

In a comK null background, a rok+ strain (HM61, circles) and a rok null strain (HM71, triangles) were grown at 30 °C in minimal glucose medium (A) after dilution to an equal starting density from exponentially growing overnight cultures or (B) after dilution to OD 0.025 from a stationary phase culture having the OD indicated at 0 hrs. Origin/terminus ratio (oriC/terC, solid lines) is plotted on the left y-axis and OD (dashed lines) is plotted on the right y-axis. Time is arbitrary and relative to a dilution point specific to each experiment. oriC/terC was measured by a qPCR assay of genomic DNA purified from the indicated strains. Primers amplified genomic regions of oriC and terC. Data are averages of biological triplicates and error bars represent the standard error. Each value (oriC or terC) was measured in triplicate, and oriC/terC for each biological replicate represents the ratio of these technical averages. qPCR values were fitted to a standard curve prepared from genomic DNA of a dnaB temperature-sensitive mutant, which has an oriC/terC of one at the restrictive temperature (Experimental Procedures). oriC/terC values less than one likely indicate the use of an imperfect standard having oriC/terC greater than one.
Replication perturbation. Although Rok modulates the chromosomal localization of YabA (Chapter 3), perturbations of YabA did not detectably sensitize cells to the loss of rok. If Rok serves as a sink for free YabA in the cell, a rok null mutant would be expected to have a higher concentration of free YabA and perhaps decreased replication initiation. Although I did not observe such a phenotype in a wild-type background, regulatory mechanisms such as the predicted titration of YabA by DnaN might mask the effect of Rok. I altered the expression of DnaN as an approach to sensitize cells to the possible effects of Rok. Overexpression of dnaN leads to increased replication initiation, likely by titrating YabA and thereby allowing cooperative binding of DnaA to oriC. Similarly, underexpression of dnaN leads to decreased replication initiation, likely due to higher levels of YabA and consequent inhibition of DnaA (Goranov et al., 2009; Merrikh and Grossman, 2011). I hypothesized that, according to the “sink model,” deletion of rok would exacerbate the under-replication phenotype of dnaN depletion by further increasing levels of free YabA. By the same model, deletion of rok and a subsequent increase in free YabA levels would be expected to suppress over-replication in a dnaN overexpression background.

I compared rok null and rok+ strains under conditions of DnaN overexpression or underexpression. Expression of dnaN was placed under control of a xylose-inducible promoter at the native dnaN locus. For these experiments, I measured oriC/terC ratio during exponential growth in minimal and rich media. Over-replication phenotypes can be easier to observe during slow growth and under-replication phenotypes can be easier to observe during fast growth. In minimal arabinose medium, there was no significant difference in oriC/terC between rok null and rok+ cells under any dnaN background (Fig. 8A). Similarly, in LB, there was little or no difference in oriC/terC between rok null and rok+ cells, except for a slight increase in the rok
null mutant under conditions of dnaN underexpression (Fig. 8B). The lack of a substantial observed replication phenotype of a rok null mutant is consistent with unpublished oriC/terC results by Houra Merrikh.

I confirmed by western blotting that DnaN protein levels responded to changes in xylose. In LB and minimal arabinose, DnaN levels were increased or decreased as expected; xylose depletion led to DnaN levels at least two-fold lower than wild-type, and increased xylose led to DnaN levels at least five-fold higher than wild-type (Fig. 9). DnaN levels were generally similar between rok null and rok+ strains. Where differences occurred, the differences should have exacerbated any phenotype of the rok null mutant. That is, DnaN levels were slightly lower in the rok null for underexpression and higher in the rok null for overexpression. Perturbations in DnaN levels generally had the expected effects on replication initiation. dnaN overexpression increased initiation in both strain backgrounds in minimal arabinose medium (Fig. 8A) and LB (Fig. 8B). dnaN depletion did not detectably decrease oriC/terC, perhaps because the fold decrease in DnaN levels was much lower than the fold increase in DnaN levels, and a greater fold change may be necessary for regulation of YabA.
Wild-type (rok+) cells (strain AG174), a rok null mutant (strain HM57), a rok+ Pxyl-dnaN strain (CAS14), and a rok null Pxyl-dnaN strain (CAS15) were grown to mid-exponential phase at 30 °C in (A) minimal arabinose medium or (B) LB medium. AG174 and HM57 were grown in the absence of xylose and represent wild-type DnaN levels. CAS14 and CAS15 were grown in the absence of xylose for DnaN underexpression or grown in the presence of 0.5% xylose for DnaN overexpression. Origin/terminus ratio (oriC/terC) was measured by a qPCR assay of genomic DNA purified from the indicated strains. Primers amplified genomic regions of oriC and terC. Data are averages of at least three biological replicates and error bars represent the standard error. Each value (oriC or terC) was measured in triplicate, and oriC/terC for each biological replicate represents the ratio of these technical averages. qPCR values were fitted to a standard curve prepared from genomic DNA of a dnaB temperature-sensitive mutant, which has an oriC/terC of one at the restrictive temperature (Experimental Procedures). See Table 1 for an explanation of the HM57 allele.
Figure 9. Response of DnaN protein levels to variable xylose.
Wild-type (rok+) cells (strain AG174), a rok null mutant (strain HM57), a rok+ Pxyl-dnaN strain (CAS14), and a rok null Pxyl-dnaN strain (CAS15) were grown to mid-exponential phase at 30 °C in (A) minimal arabinose medium or (B) LB medium. AG174 and HM57 were grown in the absence of xylose and represent wild-type DnaN levels. CAS14 and CAS15 were grown in the absence of xylose for DnaN underexpression or grown in the presence of 0.5% xylose for DnaN overexpression. Protein levels were measured by quantitative western blotting (Experimental Procedures) with a fluorescent secondary antibody. Data are averages of biological triplicates (except for CAS14 in LB + xylose; smearing of bands for the other two replicates prevented quantitation). All sample intensities except the extreme highest ones (CAS15 in LB + xylose) were confirmed to be in linear range. See Table 1 for an explanation of the HM57 allele.
**YabA perturbation.** I found no detectable effect of *rok* on replication initiation in a *yabA* null background, which was used to control for the dual effects of Rok on DnaA and YabA. Since Rok recruits DnaA and YabA independently of each other to certain chromosomal regions, perhaps Rok could regulate the ratio of free DnaA/YabA. If Rok regulates relative DnaA/YabA levels, the simultaneous increase in free DnaA and free YabA levels in a *rok* null strain could account for lack of an observable phenotype in the previous oriC/terC experiments. I measured oriC/terC in *rok* null and *rok*+ cells in a *yabA* null background. The *yabA* null mutation increased replication as expected, but there was no further detectable effect of *rok* in minimal glucose media (Fig. 10A) or minimal succinate media (Fig. 10B).

If Rok does affect replication initiation, its effects are likely to be subtle, condition-specific, or masked by any of several mechanisms of initiation regulation. In addition to YabA, known regulators of replication initiation in *B. subtilis* include SirA, Soj, and DnaD (Skarstad and Katayama, 2013). SirA is only expressed during sporulation (Rahn-Lee et al., 2009; Wagner et al., 2009), so it is unlikely to affect the function of Rok during exponential growth. DnaD is an essential protein (Kobayashi et al., 2003), so it is difficult to control for its possible involvement in regulating replication in combination with Rok.
Figure 10. No replication phenotype of a rok null mutant in a yabA null background.
Wild-type (rok+) cells (strain AG174), a rok null mutant (strain HM57), a rok+ yabA null mutant (strain AIG109), and a rok null yabA null double mutant (strain HM67) were grown to mid-exponential phase at 30 °C in (A) minimal glucose medium or (B) minimal succinate medium. Origin/terminus ratio (oriC/terC) was measured by a qPCR assay of genomic DNA purified from the indicated strains. Primers amplified genomic regions of oriC and terC. Data are averages of at least three biological replicates and error bars represent the standard error. Each value (oriC or terC) was measured in triplicate, and oriC/terC for each biological replicate represents the ratio of these technical averages. qPCR values were fitted to a standard curve prepared from genomic DNA of a dnaB temperature-sensitive mutant, which has an oriC/terC of one at the restrictive temperature (Experimental Procedures). See Table 1 for an explanation of the HM57 allele.
Experimental Procedures

Strains and alleles

*B. subtilis* strains and relevant genotypes are listed in Table 1. Properties and construction of important alleles are described below.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype (comment and/or reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG174</td>
<td><em>trpC2 pheA1</em> (wild-type, JH642) (Perego et al., 1988)</td>
</tr>
<tr>
<td>WKS1038</td>
<td>Δ<em>rok::cat</em> (double-crossover integration; see Chapter 2) (Smits and Grossman, 2010)</td>
</tr>
<tr>
<td>HM57</td>
<td><em>rok::mls</em> (single-crossover integration)</td>
</tr>
<tr>
<td>HM61</td>
<td>Δ<em>comK::cat</em></td>
</tr>
<tr>
<td>HM71</td>
<td>Δ<em>comK::cat, rok::mls</em> (single-crossover integration)</td>
</tr>
<tr>
<td>CAS14</td>
<td><em>dnaN::[PxyLA-dnaN cat::tet]</em></td>
</tr>
<tr>
<td>CAS15</td>
<td><em>dnaN::[PxyLA-dnaN cat::tet], rok::mls</em> (single-crossover integration)</td>
</tr>
<tr>
<td>AIG109</td>
<td><em>trp+, ΔyabA::cat</em> (HM2; see Chapter 3) (Goranov et al., 2009)</td>
</tr>
<tr>
<td>HM67</td>
<td>Δ<em>yabA::cat, rok::mls</em> (single-crossover integration)</td>
</tr>
</tbody>
</table>

Table 1. *B. subtilis* strains used in this study.
All strains are derived from AG174 and contain the *trpC2 pheA1* alleles unless otherwise indicated.

Δ*rok*. Strain HM57 contains a major disruption of the *rok* coding sequence but not a true null allele. The 3’ two-thirds of the *rok* coding sequence was replaced with a macrolide-lincosamide-streptogramin (MLS) resistance cassette using the vector pDG641-roksco (*rok* single crossover). pDG641 targets the region of the *rok* coding sequence amplified by primers oWKS202 and oWKS203. pDG641-roksco was integrated into *B. subtilis* 168 to generate strain WKS907 (HM47). Strain HM57 was generated by back-crossing genomic DNA from WKS907 into the AG174 (wild-type) background. The single-crossover insertion is expected to leave intact a portion of the *rok* sequence encoding a 64-aa N-terminal fragment with a predicted molecular weight of 7.62 kDa. This fragment corresponds to nearly one-third of the Rok protein and two-thirds of the N-terminal Domain I. This portion of Rok does not bind DNA (Smits and Grossman, 2010) and it is not required for Rok-dependent recruitment of DnaA or YabA to Rok-
bound chromosomal regions (Chapters 2, 3). Nevertheless, strain WKS1038, which contains a true \textit{rok} null allele, is preferable to strain HM57. Strains HM67, HM71, and CAS15 contain the HM57 allele. WKS1038 was used for all of my experiments requiring a \textit{rok} null mutant after July 24, 2012.

\textit{ΔyabA::cat}. Strain AIG109 contained a deviation in the \textit{ypjG-hepT} region, containing tryptophan biosynthesis genes, as described previously (Berkmen and Grossman, 2007).

\textit{ΔcomK::cat}. Strains HM61 and HM71 contain a \textit{ΔcomK::cat} allele from strain JMA174 (HM46). Genomic DNA from HM46 was transformed into the wild-type background of AG174 to generate strain HM61.

\textit{dnaN::[PxylA-dnaN cat::tet]}. Strains CAS14 and CAS15 contain a xylose-inducible \textit{dnaN} allele derived from strain AIG260. AIG260 is a sister isolate of strain AIG261 from (Goranov et al., 2009). The only full copy of \textit{dnaN} is placed under control of the xylose-inducible promoter PxylA. The chloramphenicol resistance gene of AIG261 was converted to tetracycline resistance in anticipation of additional genetic manipulations, although the change was not necessary for experiments described in this work. AIG260 was transformed with the chloramphenicol-to-tetracycline conversion plasmid from strain AG1549 to generate strain CAS14. CAS14 was transformed with genomic DNA from strain HM57 to generate strain CAS15.

**Growth conditions**

Unless otherwise specified, all strains were grown at 30 °C in S7 defined minimal medium buffered with 50 mM MOPS (Jaacks et al., 1989) and containing 1% glucose or other carbon source, 0.1% glutamate, trace metals, 40 \(\mu\)g/ml tryptophan, and 40 \(\mu\)g/ml phenylalanine.
All carbon sources were used at a concentration of 1% unless otherwise specified. Xylose was added at 0.5% to induce dnaN overexpression or omitted to represent dnaN underexpression (basal levels). Strains without the Pxyl-dnaN allele were grown in the absence of xylose to represent wild-type levels of DnaN. Exponentially growing cells were collected at OD 0.3-0.4.

**Origin/terminus assay**

Relative copy numbers of the origin of replication (oriC) and terminus (terC) were measured by quantitative PCR (qPCR). Cells, typically 5 ml of exponentially growing culture, were fixed with an equal volume of -20 °C methanol and incubated at -20 °C for at least 15 min before pelleting. Genomic DNA was purified by phenol-chloroform extraction and resuspended in 100 μl TE (10 mM Tris pH 8.0, 0.1 mM EDTA) per 5 ml of the original culture. The following primer sets were used for amplification: oHM126 with oHM127 to amplify the dnaA promoter region of oriC and oWKS141 with oWKS142 to amplify the rtp region of terC. Each well in a 96-well white qPCR plate (Qarta Bio) contained 1X LightCycler 480 SYBR Green I Master Mix (Roche), 2.5 μM forward primer, 2.5 μM reverse primer, and template DNA (experimental samples or standards) in a final volume of 20 μl. Template DNA from experimental samples was used at a final reaction concentration diluted 5,000X from the initial TE stock. Three technical replicates (qPCR reactions) were performed for each biological replicate.

A standard curve was generated using 8 x 3-fold serial dilutions of genomic DNA from a dnaB temperature-sensitive mutant strain, KPL69 (HM5). When grown under restrictive conditions, this mutant cannot initiate new rounds of replication but can complete previously initiated rounds, resulting in a nearly uniform copy number for all regions of the genome. These
standards are therefore assumed to have an oriC/terC ratio of one. DNA for standards was
prepared by growing KPL69 to early exponential phase in LB at 30 °C, followed by growth for
at least 1 hr at the restrictive temperature, 48 °C. Cells were killed by heating to >60 °C for at
least 5 minutes until the internal temperature of the flask reached 60 °C. Even transient re-
exposure to permissive temperatures can allow re-initiation in some cells, which can lead to
DNA standards with an actual oriC/terC greater than one. Genomic DNA was prepared in the
same way as experimental samples. qPCR was performed on a Roche LightCycler 480.

Crossing-point cycle (Cp) values were calculated from melting curves using the
LightCycler 480 software. Cp values (on a logarithmic scale) were converted to relative DNA
concentrations (on a linear scale), using a standard curve generated by the KPL69 dilutions.
Technical triplicates were averaged to obtain a relative DNA concentration for oriC and for terC.
The oriC/terC ratio for each biological replicate is the ratio of these two technical averages.

**DNA/protein assay**

The DNA/protein ratio was measured essentially as described in (Goranov et al., 2009).
Exponentially growing cells were harvested at approximately OD 0.4. DNA and protein were
extracted from the same cell pellets. DNA concentrations were measured using the
diphenylamine reaction with a standard curve prepared from salmon sperm DNA. Protein
concentrations were measured using the Lowry Bio-Rad assay with a standard curve prepared
from bovine serum albumin.
Western blotting

Exponentially growing cells were lysed with lysozyme and protease inhibitors in ChIP Buffer A (Chapter 2). Lysates were run on 12% polyacrylamide gel and transferred to a PVDF membrane. To represent the same number of cells in each lane, the amounts loaded in each lane were normalized to culture OD. All steps were performed at room temperature. According to the manufacturer’s instructions, blots were blocked with Odyssey Blocking Buffer for 1 hr and incubated with primary antibody (rabbit anti-DnaN antibody 1:20,000 in Odyssey Blocking Buffer + 0.2% Tween) for 1 hr. Blots were washed with PBST (phosphate-buffered saline + 0.2% Tween) for at least 4 x 5 min and then incubated for 30 min with secondary antibody (LiCor dye 800 goat anti-rabbit 1:10,000 in Odyssey Blocking Buffer + 0.2% Tween). Blots were imaged and quantitated on a LiCor scanner. Dilutions of AG174 and CAS14 lysates were used to generate a standard curve and determine the linear range of the fluorescence signal. All sample intensities except the extreme highest ones (CAS15 in LB + xylose) were confirmed to be in linear range.
Appendix B

Conservation of DnaA, YabA, and Rok
YabA and Rok-LS20-like Proteins

The small plasmid-encoded Rok-LS20 and its analogues retain much of Domain III, which I found is sufficient for recruitment of DnaA and YabA to Rok-bound chromosomal regions. It is possible, then, that small Rok proteins maintain the ability to interact with DnaA from *Bacillus* species. I examined the YabA protein sequences of species with large, small, or both *rok* genes, and the YabA sequences were highly similar, with 50-100% identity to *B. subtilis* YabA (*Table 1*). Organisms with only small *rok* genes showed more divergent YabA proteins, but these differences are likely to reflect overall evolutionary distance, and I did not notice any drastically different sequence features in the YabA proteins of these species.

DnaA in a *Bacillus* species that lacks Rok

If the DnaA-Rok interaction is conserved and reflected in the evolutionary history of these proteins, I would expect the DnaA protein sequences to be different in organisms with or without Rok. Such differences might suggest a Rok-interacting region of the DnaA protein. I identified *B. anthracis* as the closest organism to *B. subtilis* that does not have *rok*. Of the *Bacillus* clade, *B. cereus* and *B. anthracis* were identified as lacking *rok* (Albano et al., 2005; Singh et al., 2012). I used BLAST (nucleotide or protein) of *B. subtilis* 168 *rok* or Rok against *B. cereus* and *B. anthracis*. (A BLAST search against all organisms gave similar results as (Singh et al., 2012).) I found that *B. cereus* has a hypothetical protein somewhat similar to Rok (*Fig. 1A*). The proteins share 17.7% identity and are most similar in Domains II and III (*Fig. 1B*).
<table>
<thead>
<tr>
<th>Rok Protein, Species</th>
<th>Rok Accession</th>
<th>Rok Size</th>
<th>Two Roks</th>
<th>YabA Accession</th>
<th>% Identity to B. subtilis 168 YabA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rok_L_subW23, B. subtilis subsp. spizizenii W23</td>
<td>YP_003865820.1</td>
<td>L</td>
<td>Y</td>
<td>YP_003864402.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_L_spizAT, B. subtilis subsp. spizizenii ATCC 6633</td>
<td>ZP_06872236.1</td>
<td>L</td>
<td>Y</td>
<td>ZP_06872020.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_L_spizTU, B. subtilis subsp. spizizenii TU-B-10</td>
<td>YP_004877032.1</td>
<td>L</td>
<td>Y</td>
<td>YP_004875440.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_LlicATC, B. licheniformis ATCC 14580</td>
<td>YP_078814.1</td>
<td>L</td>
<td>Y</td>
<td>YP_077319.1</td>
<td>82</td>
</tr>
<tr>
<td>Rok_L_subBSn5, B. subtilis BSn5</td>
<td>YP_004207469.1</td>
<td>L</td>
<td>Y</td>
<td>YP_004205990.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_L_168, B. subtilis subsp. subtilis 168</td>
<td>NP_389307.1</td>
<td>L</td>
<td>Y</td>
<td>NP_387914.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_L_amyloF, B. amyloliquefaciens FZB42</td>
<td>YP_001420994.1</td>
<td>L</td>
<td>Y</td>
<td>YP_001419713.1</td>
<td>95</td>
</tr>
<tr>
<td>Rok_L_amyloD, B. amyloliquefaciens DSM 7</td>
<td>YP_003920096.1</td>
<td>L</td>
<td>Y</td>
<td>YP_003918629.1</td>
<td>94</td>
</tr>
<tr>
<td>Rok_L_B. atroph, B. atropheus 1942</td>
<td>YP_003972862</td>
<td>L</td>
<td>Y</td>
<td>YP_003975603.1</td>
<td>94</td>
</tr>
<tr>
<td>Rok_L_pumSAF, B. pumilus SAFR-032</td>
<td>YP_001486564.1</td>
<td>L</td>
<td>Y</td>
<td>YP_001485277.1</td>
<td>76</td>
</tr>
<tr>
<td>Rok_L_BpumAT, B. pumilus ATCC 7061</td>
<td>ZP_03052836.1</td>
<td>L</td>
<td>Y</td>
<td>ZP_03056604.1</td>
<td>76</td>
</tr>
<tr>
<td>Hypothetical protein BSNT_02376, B. subtilis subp. natto BEST195</td>
<td>BAI85036.1</td>
<td>L</td>
<td>Y</td>
<td>BAI83479.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_S_subW23, B. subtilis subsp. spizizenii W23</td>
<td>YP_003864911</td>
<td>S</td>
<td>Y</td>
<td>YP_003864402.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_Ssubspi, B. subtilis subsp. spizizenii TU-B-10</td>
<td>4876112.1</td>
<td>S</td>
<td>Y</td>
<td>YP_004875440.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_S_spizAT, B. subtilis subsp. spizizenii ATCC 6633</td>
<td>ZP_06875354.1</td>
<td>S</td>
<td>Y</td>
<td>ZP_06872020.1</td>
<td>100</td>
</tr>
<tr>
<td>Rok_S_Lich14, B. licheniformis ATCC 14580</td>
<td>YP_079175.1</td>
<td>S</td>
<td>Y</td>
<td>YP_077319.1</td>
<td>82</td>
</tr>
<tr>
<td>Rok_S_bacill, Bacillus sp. BT1B_CT2</td>
<td>ZP_08000071.1</td>
<td>S</td>
<td>Y</td>
<td>EFV70137.1</td>
<td>82</td>
</tr>
<tr>
<td>Rok_S_pseudo, B. pseudomycooides DSM 12442</td>
<td>ZP_04153718.1</td>
<td>S</td>
<td>Y</td>
<td>ZP_04149235.1</td>
<td>63</td>
</tr>
<tr>
<td>Rok_S_coag 36, B. coagulans 36D1</td>
<td>YP_004861201.1</td>
<td>S</td>
<td>Y</td>
<td>YP_004859255.1</td>
<td>54</td>
</tr>
<tr>
<td>Rok_S_Coag 2, B. coagulans 2-6</td>
<td>YP_004568723</td>
<td>S</td>
<td>Y</td>
<td>YP_004567463.1</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 1. YabA proteins of Bacillus species with large or small Rok proteins.
Rok proteins and species of interest are from (Singh et al., 2012). YabA protein sequences from these species were obtained from PubMed.
A >gi|507025268|ref|WP_016097327.1| MULTISPECIES: hypothetical protein [Bacillus cereus group] MRKNKICLVSVCATALTLYSTLSIFSEKVAADTIENTTTIQQTQESNDLQNIKIPINDINDLDEALGFTP SQINDFKNSTSPNSLAPLAALKTQYVYWNKAKLQGVLCRSLSLSTPTGKLGDWLISQGIALLGPAKSV ALSLAFINTALKRVLSFLQNALTQVRFGQATGIRIAIEPNPQGYPAYVTMSIY

B

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MRKNKICLVSVCATALTLYSTLSIFSEKVAADTIENTTTIQQTQESNDLQNIKIPINDINDLDEALGFTP SQINDFKNSTSPNSLAPLAALKTQYVYWNKAKLQGVLCRSLSLSTPTGKLGDWLISQGIALLGPAKSV ALSLAFINTALKRVLSFLQNALTQVRFGQATGIRIAIEPNPQGYPAYVTMSIY</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>K RELDRNISPSEIKKDFREKPSLVPVLAELAAEQEIRSYQPSQQSVQPSQLQSISSL</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NDLDLEALGFTPSQINDFKNSTPSNSLAPLAALKTQYVYWNKAKLQGVLCRSLSLSTPTGKLGDWLISQGIALLGPAKSV ALSLAFINTALKRVLSFLQNALTQVRFGQATGIRIAIEPNPQGYPAYVTMSIY</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>PAGIPDGTTRRRRTGSAKLREA---AIKT-LKRHNAIKSSELQKEIEKESGLE</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>PTGKPLGDWLSQGIALLGPAKSVGLSLAFINTALKRVLSFLQNALTQVRFGQATGIR</td>
<td></td>
</tr>
</tbody>
</table>

152 IPNMTTFMQLKMYPEVKPPGCRQYILEGEIESANE 191 sub
175 IAIE----------------P--NPQGYPAYVTMSIY-------- 194 cer

Figure 1. Sequence alignment of *B. subtilis* Rok protein and a hypothetical protein from *B. cereus*.

(A) Protein sequence of the candidate Rok homologue in *B. cereus*. (B) Sequences of *B. subtilis* Rok (sub) and the *B. cereus* hypothetical protein (cer). Alignment was performed using the Uniprot aligner (http://www.uniprot.org/align). Identical residues are indicated with (*), residues with similar properties are indicated with (:), and residues with weakly similar properties are indicated with (.).

*B. anthracis* had no detectable *rok*, so I considered it the closest *rok*-lacking relative of *B. subtilis*. I then compared the DnaA protein sequences of *B. anthracis* and *B. subtilis* and found that they are 84% identical (Fig. 2). Testing *B. anthracis* DnaA for interaction with *B. subtilis* Rok could help define the contribution, if any, of these divergent residues to Rok interaction.
Figure 2. Sequence alignment of DnaA proteins from *B. subtilis* and *B. anthracis*.
Alignment was performed using the Uniprot aligner (http://www.uniprot.org/align). Identical residues are indicated with (*), residues with similar properties are indicated with (:), and residues with weakly similar properties are indicated with (.).

Potential shared protein sequence motifs in DnaA and Rok

DnaA and Rok share two short protein sequence motifs whose similarity suggests a possible recognition for a shared interactor (such as YabA) or perhaps interaction with each other by cross-recognition (Fig. 3). These sequences show low statistical significance (high E-value), but the separate matter of biological significance has yet to be investigated. Motif 1 occurs in Domain IIIb of DnaA (ATP binding) and in the DNA binding domain (III) of Rok. Motif 2 occurs in the DNA binding domain (IV) of DnaA and Domain II of Rok. These motifs are still present in DnaA from *B. anthracis*, the closest known relative of *B. subtilis* that lacks *rok*, so the
presence of these motifs in DnaA is not necessarily indicative of the presence of a *rok* gene in that species.

**Motif 1**
Score = 17.7 bits (34), Expect = 0.32, Method: Compositional matrix adjust.
Identities = 6/9 (67%), Positives = 7/9 (78%), Gaps = 0/9 (0%)

<table>
<thead>
<tr>
<th></th>
<th>DnaA</th>
<th>Rok</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KAEGLDIPN</td>
<td>KESGLEIPN</td>
</tr>
<tr>
<td>290</td>
<td>146</td>
<td></td>
</tr>
</tbody>
</table>

**Motif 2**
Score = 13.1 bits (22), Expect = 8.7, Method: Compositional matrix adjust.
Identities = 4/9 (44%), Positives = 8/9 (89%), Gaps = 0/9 (0%)

<table>
<thead>
<tr>
<th></th>
<th>DnaA</th>
<th>Rok</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADLAAEALK</td>
<td>AELAAQEIR</td>
</tr>
<tr>
<td>335</td>
<td>66</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3. Potential shared protein sequence motifs in DnaA and Rok.**
Motif 1 (A) in black text and Motif 2 (B) in red text were identified using BLAST. Occurrences of these motifs in the sequences of DnaA (C) and Rok (D) are underlined.

**Acknowledgments**

The potential motifs shared by DnaA and Rok were originally identified by Houra Merrikh.
Appendix C

The pyruvate dehydrogenase complex subunit PdhC is not a direct regulator of replication initiation
Summary

The metabolic enzyme PdhC, which functions in the pyruvate dehydrogenase complex upstream of the citric acid cycle, has been suggested to regulate replication initiation in *Bacillus subtilis*. I confirmed and characterized the severe growth defect of a *pdhC* null mutant in rich media, even with the addition of supplemental carbon sources that were expected to bypass the metabolic defect. I observed under-initiation of replication and a low DNA/protein ratio in the *pdhC* null mutant, suggesting a positive role in initiation, in contrast to previous reports suggesting a negative regulatory function. Finally, I showed that the replication phenotype of the *pdhC* null mutant is indirect and not specific to the PdhC subunit. A null mutation in *pdhB*, which abolishes pyruvate dehydrogenase complex activity but does not perturb PdhC, showed a similar effect on replication initiation as the *pdhC* null mutation. This indirect role for the pyruvate dehydrogenase complex in regulating replication is supported by a recent finding that PdhA (but not PdhC) couples pyruvate levels to cell division in *B. subtilis*. Together, these results indicate that pyruvate metabolism influences the cell cycle, although not through the initially proposed mechanism of regulation by PdhC.

Introduction

The pyruvate dehydrogenase complex subunit PdhC has been proposed to regulate DNA replication initiation in *B. subtilis*. The pyruvate dehydrogenase (PDH) complex catalyzes the conversion of pyruvate to acetyl-CoA, which enters the citric acid cycle (White, 2007). The PDH-catalyzed reactions connect glycolysis with the citric acid cycle, although the citric acid cycle is not essential for *B. subtilis* exponential growth (Fortnagel and Freese, 1968). The PDH complex contains three subunits, encoded by the *pdhABCD* operon in *B. subtilis*: E1 (pyruvate
decarboxylase, with alpha and beta subunits encoded by \textit{pdhA} and \textit{pdhB}, respectively), E2 (dihydrolipoamide acetyltransferase, encoded by \textit{pdhC}), and E3 (lipoamide dehydrogenase, encoded by \textit{pdhD}) (\textbf{Fig. 1}). The PDH complex in Gram-positive bacteria is composed of 30 E1 subunits, 60 E2 subunits, and 6 E3 subunits (Neveling \textit{et al.}, 1998).

\begin{center}
\begin{tikzpicture}
\node[draw, shape=rectangle, minimum width=1cm, minimum height=0.5cm] (pdhA) at (0,0) {\textit{pdhA}};
\node[draw, shape=rectangle, minimum width=1cm, minimum height=0.5cm] (pdhB) at (2,0) {\textit{pdhB}};
\node[draw, shape=rectangle, minimum width=1cm, minimum height=0.5cm] (pdhC) at (4,0) {\textit{pdhC}};
\node[draw, shape=rectangle, minimum width=1cm, minimum height=0.5cm] (pdhD) at (6,0) {\textit{pdhD}};
\end{tikzpicture}
\end{center}

\textbf{Figure 1. Map of genes in the \textit{pdh} operon in \textit{B. subtilis}.}
The \textit{pdhABCD} operon encodes three subunits of the pyruvate dehydrogenase complex. \textit{pdhA} encodes the alpha subunit of E1, \textit{pdhB} encodes the beta subunit of E1, \textit{pdhC} encodes E2, and \textit{pdhD} encodes E3. \textit{pdhA} is essential, and PdhD is involved in other metabolic processes in addition to the pyruvate dehydrogenase complex.

PdhC was proposed to inhibit replication initiation. A 64-kDa protein, later identified as PdhC (Stein and Firshein, 2000), was associated with the \textit{B. subtilis} membrane and found to associate strongly with a plasmid probe containing \textit{oriC} DNA (Laffan and Firshein, 1987). This protein was suspected to negatively regulate replication because an antibody to the 64-kDa protein enhanced replication initiation in an \textit{in vitro} replication system prepared from \textit{B. subtilis} membrane extracts (Laffan and Firshein, 1988). Subsequent identification and purification of PdhC showed an inhibitory effect on the same assay, consistent with the model of negative regulation (Stein and Firshein, 2000). Additionally, PdhC was found to interact with the DNA primase, DnaG, in a yeast two-hybrid screen; it was speculated that the acyltransferase activity of PdhC might be involved in modification of DnaG (Noirot-Gros \textit{et al.}, 2002).
Results and Discussion

The pdhC null mutant has a severe growth defect due to nutrient limitation

The pdhC null mutant had a severe growth defect, even in LB, the optimal growth medium of several that were tested. In LB at 37 °C, the pdhC null mutant doubled in 30 min compared to 23 min for wild-type (Fig. 2A), and the pdhC null mutant reached a maximum OD of only 0.5 compared to 3.6 for wild-type (Fig. 2B). Previous work showed that the addition of citric acid cycle intermediates such as acetate, citrate, or succinate could partially restore or enhance growth of PDH mutants (Freese and Fortnagel, 1969). I tested the effects of adding individual or combined intermediates (0.2% acetate, citrate, or succinate) to LB. None substantially improved doubling time or maximum density in the pdhC null mutant or wild-type strain. Citrate, alone or in combination with the other intermediates, had a detrimental effect on growth in both strains. Doubling time increased ~50% in wild-type and ~20% in the pdhC null mutant, and the maximum OD of the wild-type strain fell to a level comparable to the pdhC null mutant (Fig. 2). Difco Sporulation Medium (DSM) with or without acetate did not improve growth in either strain (Fig. 2).
Figure 2. Growth of a *pdhC* null mutant in media with citric acid cycle intermediates.
Wild-type cells (strain AG174, white bars) and a *pdhC* null mutant (strain JLS1, blue bars) were grown at 37 °C in LB medium supplemented with indicated citric acid cycle intermediates at 0.2% each. Data represent single biological experiments.

(A) Doubling times were calculated by fitting an exponential regression equation to the exponential portion of each growth curve.

(B) Maximum OD reflects the highest optical density measurement observed in stationary phase.
This defect is consistent with previous observations. In Nutrient Sporulation Medium, the *pdhC* null mutant grows slowly and reaches stationary phase at one-third the density of the wild-type strain. Addition of 0.2% acetate, glutamate, succinate, or citrate in NSM only slightly increases the density to half that of wild-type (Gao *et al.*, 2002). Previous work by Janet Smith showed that the *pdhC* null mutant did not grow in minimal media supplemented with three citric acid cycle intermediates (0.2% each of succinate, citrate, and acetate, as described in (Freese and Fortnagel, 1969)). The same combination of citric acid cycle intermediates did not substantially affect the growth rate of the *pdhC* null strain in LB, nor did addition of 4.5 mM Tris-HCl, pH 7.0, to buffer metabolically induced pH changes.

The growth defect was due to nutrient limitation rather than accumulation of a toxic byproduct because the maximum culture density increased in more concentrated media. Increasing the LB concentration from 0.5X to 5X increased the maximum OD of both the *pdhC* null mutant (by five-fold) and the wild-type strain (by two-fold) (**Fig. 3**). Supplying additional nutrients increased the maximum density, as expected for nutrient-limited growth but not if growth had been limited by toxic byproduct accumulation.
Figure 3. Effect of medium concentration on growth of a pdhC null mutant.
Wild-type cells (strain AG174, white bars) and a pdhC null mutant (strain JLS1, blue bars) were grown at 37 °C in LB medium prepared at the indicated concentrations. Data represent single biological experiments. Maximum OD reflects the highest optical density measurement observed in stationary phase.

Reduced replication initiation in a pdhC null mutant

I observed a 10-20% reduction in replication initiation in a pdhC null mutant compared to a wild-type strain. I measured replication by two assays: DNA/protein ratio and a qPCR assay of chromosomal origin/terminus ratio \((oriC/terC)\). The \(oriC/terC\) assay is more representative of initiation, whereas the DNA/protein assay can also reflect altered replication elongation. In the \(oriC/terC\) assay, the pdhC null mutant showed a 50% reduction in replication initiation relative to the wild-type strain (Fig. 4A). In the DNA/protein assay, the pdhC null mutant showed little effect relative to the wild-type strain (Fig. 4B). These results are consistent with a role for pdhC in positively regulating replication initiation.
This role was supported by the phenotype of a strain overexpressing \textit{pdhC} as a C-terminal 3x-myc fusion. Overexpression of this \textit{pdhC} fusion partially complemented the replication phenotype of the \textit{pdhC} null mutant in the ori\textit{C}/ter\textit{C} assay (\textbf{Fig. 4A}) and reversed the phenotype in the DNA/protein assay (\textbf{Fig. 4B}). It is difficult to interpret results from the overexpression construct, however, because it was unable to complement the growth phenotype of a \textit{pdhC} null mutant (\textbf{Fig 5}). The myc-tag or altered expression levels may compromise the function of PdhC.

**The replication phenotype of a \textit{pdhC} null mutant is indirect and not specific to the PdhC subunit**

To distinguish direct (PdhC-specific) vs. indirect (metabolic) effects on replication, I examined the effect of abolishing PDH function without specifically disrupting \textit{pdhC}. I found that disrupting PDH activity via a \textit{pdhB} null mutant had a similar replication phenotype as the \textit{pdhC} null mutant, indicating the replication defect of the \textit{pdhC} deletion represents an indirect metabolic effect rather than a specific regulatory function of PdhC.

I disrupted PDH complex activity independently of \textit{pdhC} by instead deleting \textit{pdhB}. The other \textit{pdh} genes were not good candidates for deletion because \textit{pdhA} is essential (Gao \textit{et al.}, 2002; Kobayashi \textit{et al.}, 2003) and the E3 subunit encoded by \textit{pdhD} is shared with other metabolic complexes (Neveling \textit{et al.}, 1998), so a \textit{pdhD} null would not specifically represent a PDH loss of function. The E1 subunit encoded by \textit{pdhB} is not known to participate in other metabolic processes, so this mutation was not expected to have off-target effects. In previous work, deletions of \textit{pdhB}, \textit{pdhC}, and \textit{pdhD} each abolished PDH activity, and immunoblotting showed that each mutation is not polar on the other \textit{pdh} genes (Gao \textit{et al.}, 2002).
Figure 4. The replication phenotype of PDH complex disruption is not specific to *pdhC*. Wild-type cells (strain AG174), a *pdhC* null mutant (strain JLS1), a *pdhB* null mutant (strain CAS2), and a strain expressing Pspank-*pdhC-myc* (HM105) were grown at 37 °C in LB medium to mid-exponential phase. Strain HM105 overexpressed *pdhC-myc* under control of the IPTG-inducible promoter Pspank integrated at *amyE*. Data are averages of biological triplicates and error bars represent the standard error.

(A) Origin/terminus ratio (*oriC/terC*) was measured by a qPCR assay of genomic DNA purified from the indicated strains. Primers amplified genomic regions of *oriC* and *terC*. Each value (*oriC* or *terC*) was measured in triplicate, and *oriC/terC* for each biological replicate represents the ratio of these technical averages. qPCR values were fitted to a standard curve prepared from genomic DNA of a *dnaB* temperature-sensitive mutant, which has an *oriC/terC* of one at the restrictive temperature (Experimental Procedures).

(B) DNA/protein was measured by standard colorimetric assays (Experimental Procedures). For each biological replicate, DNA concentration was measured in triplicate and protein concentration was measured in quadruplicate (duplicate measurements of two dilutions).
Figure 5. Overexpression of pdhC-myc does not complement the growth defect of a pdhC null mutant.

Wild-type cells (strain AG174), a strain expressing Pspank-pdhC-myc (HM105), a pdhC null mutant (strain JLS1), and a pdhC null mutant expressing Pspank-pdhC-myc (strain HM131) were grown at 37 °C in LB medium. Strains HM105 and HM131 overexpressed pdhC-myc under control of the IPTG-inducible promoter Pspank integrated at amyE. Data represent single biological experiments.

(A) Doubling times were calculated by fitting an exponential regression equation to the exponential portion of each growth curve.

(B) Maximum OD reflects the highest optical density measurement observed in stationary phase.
As expected, the \textit{pdhB} null mutant had a growth defect relative to the wild-type strain. Similar to the \textit{pdhC} null mutant, the doubling time in the \textit{pdhB} null mutant was higher than wild-type (Fig. 6A) and the maximum OD was lower than wild-type (Fig. 6B). There were no detectable polar effects of this deletion because expression of \textit{pdhAB} under the native \textit{pdh} promoter could complement the growth defects of the \textit{pdhB} mutant (Fig. 6).

The \textit{pdhB} null mutant had a similar decrease in \textit{oriC/terC} as a \textit{pdhC} null mutant (Fig. 4A). If \textit{pdhC} affected replication independently of its metabolic function, this alternative PDH-disrupted strain would not be expected to show an initiation phenotype. Therefore, the role for \textit{pdhC} in replication can be explained by metabolic effects rather than direct regulation of initiation. There was a slight increase in DNA/protein in the \textit{pdhB} null mutant relative to the \textit{pdhC} null mutant and wild-type strain (Fig. 4B), but this assay is less specific to replication initiation than \textit{oriC/terC}, and the difference might reflect dysregulation of PdhA protein levels in the absence of PdhB. The indirect role of \textit{pdhC} is consistent with recent work showing that PdhA (but not PdhC) couples pyruvate levels to cell division in \textit{B. subtilis}. PdhA localizes to the nucleoid in a pyruvate-dependent manner and may enhance Z-ring formation in nutrient-rich conditions (Monahan \textit{et al.}, 2014). The pyruvate dehydrogenase complex does connect replication to growth and metabolism, but not, as originally suspected, through the PdhC subunit.
Figure 6. Overexpression of *pdhAB* complements the growth defect of a *pdhB* null mutant. Wild-type cells (strain AG174), a *pdhB* null mutant (strain CAS2), a strain expressing *pdhAB* from *amyE* (CAS3), and a *pdhB* null mutant expressing *pdhAB* from *amyE* (strain CAS5) were grown at 37 °C in LB medium. Strains CAS3 and CAS5 expressed *pdhAB* under control of the native *pdh* promoter integrated at *amyE*. Data represent single biological experiments.

(A) Doubling times were calculated by fitting an exponential regression equation to the exponential portion of each growth curve.

(B) Maximum OD reflects the highest optical density measurement observed in stationary phase.
Experimental Procedures

Strains and alleles

*B. subtilis* strains and relevant genotypes are listed in Table 1. Properties and construction of important alleles are described below.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype (comment and/or reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG174</td>
<td>trpC2 pheA1 (wild-type, JH642) (Perego et al., 1988)</td>
</tr>
<tr>
<td>JLS1</td>
<td>ΔpdhC::cat</td>
</tr>
<tr>
<td>CAS1</td>
<td>ΔpdhB::kan</td>
</tr>
<tr>
<td>CAS3</td>
<td>amyE::[Ppdh-pdhAB spc] (single-crossover integration)</td>
</tr>
<tr>
<td>CAS5</td>
<td>ΔpdhB::kan, amyE::[Ppdh-pdhAB spc] (single-crossover integration)</td>
</tr>
<tr>
<td>HM105</td>
<td>amyE::[Pspank-pdhC-3x-myc spc]</td>
</tr>
<tr>
<td>HM131</td>
<td>ΔpdhC, amyE::[Pspank-pdhC-3x-myc spc]</td>
</tr>
</tbody>
</table>

Table 1. *B. subtilis* strains used in this study.
All strains are derived from AG174 and contain the *trpC2 pheA1* alleles unless otherwise indicated.

**ΔpdhC.** Strain JLS1 is a *pdhC* null strain constructed by double-crossover replacement of the *pdhC* coding sequence with a chloramphenicol resistance cassette using the long-flanking homology PCR method. It is derived from back-crossing the genomic DNA from the original transformed strain into wild-type (AG174) cells.

Strain HM105 contains a C-terminal 3x-myc fusion of *pdhC* under the control of the IPTG-inducible promoter Pspank. This insert was cloned into the integration vector pDR111 to generate pCAL1312, which was transformed into AG174. Strain HM131 was constructed by transforming the *pdhC* knockout cassette (used to generate strain JLS1) into strain HM105.

**ΔpdhB.** Strain CAS1 is a *pdhB* null strain constructed by double-crossover replacement of the *pdhB* coding sequence with a kanamycin resistance cassette, including a transcriptional terminator, using the long-flanking homology PCR method. Flanking genomic regions were amplified using primers oCS1 and oCS3 (left flank) with oCS2 and oCS4 (right flank). The kanamycin resistance cassette was amplified from pGK67 from strain AG1637.
Strain CAS3 contains a complementation construct in which the native pdh promoter drives expression of pdhAB at an ectopic locus, amyE. The pdhAB genes and their promoter were amplified using primers oCS5 and oCS6. This PCR product was inserted into the integration vector pDR111 (which had been digested with EcoRI and BamHI) via isothermal assembly. The resulting plasmid was transformed into AG174. The integration is a single crossover, as judged by the ability of this strain to digest starch, indicating a functional copy of amyE. Strain CAS5 was created by transforming the pdhB null PCR product into CAS3.

**Growth conditions**

All strains were grown at 37 °C. Citric acid cycle intermediates were prepared from stock solutions of sodium citrate, sodium succinate, or sodium acetate and added at the indicated concentrations, which were based on methods from (Gao et al., 2002). Minimal medium refers to S7 defined minimal medium buffered with 50 mM MOPS (Jaacks et al., 1989) and containing 1% glucose or other carbon source, 0.1% glutamate, trace metals, 40 µg/ml tryptophan, and 40 µg/ml phenylalanine. Doubling times were calculated by fitting an exponential regression equation to the exponential portion of each growth curve. Maximum OD reflects the highest measurement observed in stationary phase.

**DNA/protein assay**

The DNA/protein ratio was measured essentially as described in (Goranov et al., 2009). Exponentially growing cells were harvested at approximately OD 0.7 (for wild-type and pdhC overexpression strains) or OD 0.1 (for the pdhC and pdhB null mutants). DNA and protein were extracted from the same cell pellets. DNA concentrations were measured using the
diphenylamine reaction with a standard curve prepared from salmon sperm DNA. Protein concentrations were measured using the Lowry Bio-Rad assay with a standard curve prepared from bovine serum albumin.

**Origin/terminus assay**

The origin/terminus qPCR assay was performed as described in **Appendix A**.

**Acknowledgments**

Catherine Lee performed the molecular cloning for strain JLS1 and the myc-tagged *pdhC* fusion. Houra Merrikh constructed strains HM105 and HM131. Janet Smith constructed strain JLS1 and performed preliminary experiments to optimize growth of this strain.
References


Brendler, T., and Austin, S. (1999). Binding of SeqA protein to DNA requires interaction between two or more complexes bound to separate hemimethylated GATC sequences. EMBO J. 18, 2304–2310.


214


