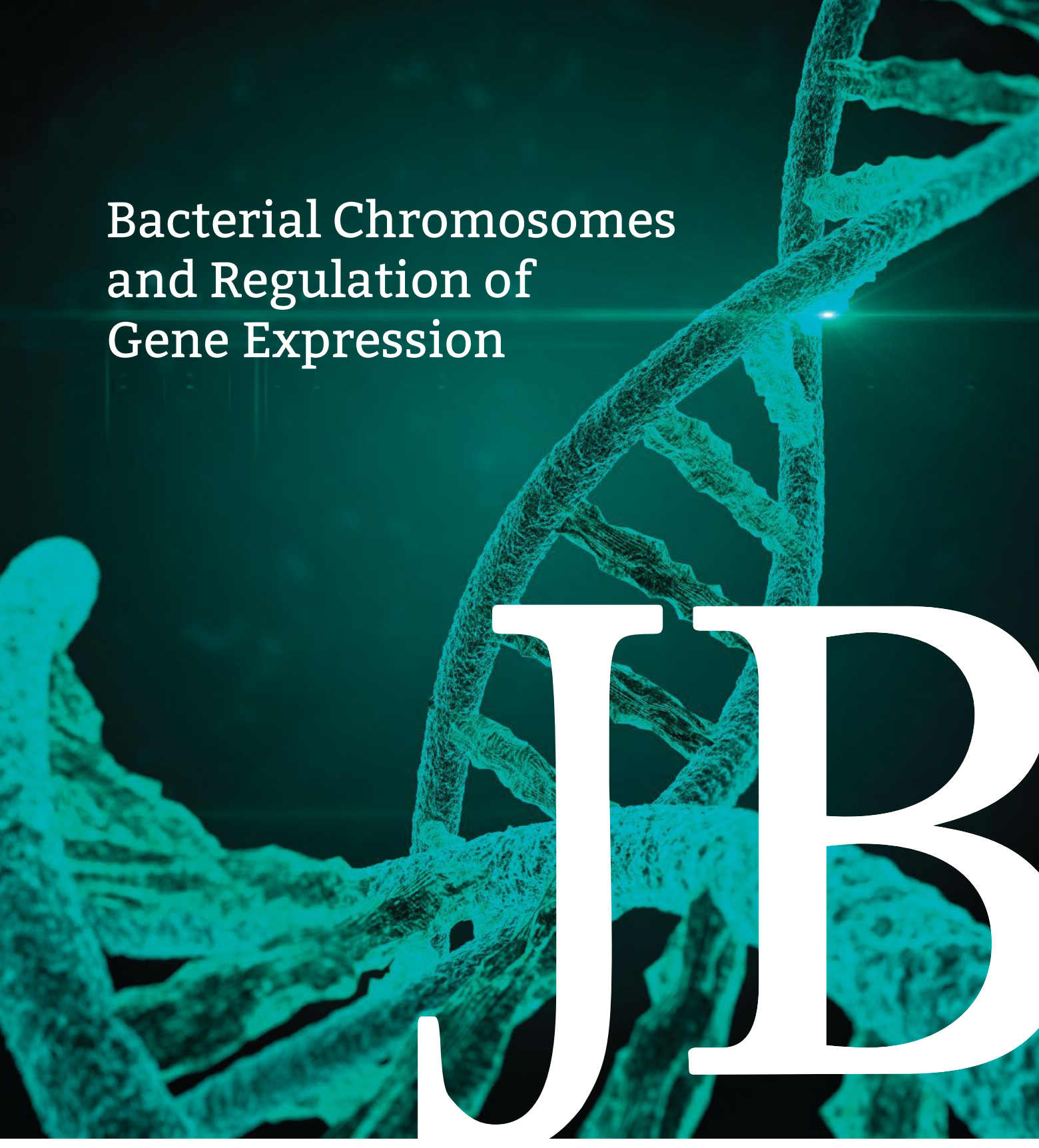
Bacterial Chromosomes and Regulation of Gene Expression



AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

BACTERIAL CHROMOSOMES AND REGULATION OF GENE EXPRESSION
A Special Collection of JB Minireviews

TABLE OF CONTENTS

Continued on following page

Continued from preceding page

Continued on following page

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

Introduction

Despite years of study, there are still many surprises in store for how bacteria and archaea manage their DNA and regulate gene expression. For those interested in these general topics, we have compiled a series of minireviews that provide insight into recent advances. The diverse topics range from transcription regulation in archaea to nonstop translation, from DNA uptake and CRISPR systems to chromosome structure and dynamics. We think that this compilation, as well as forthcoming compilations centered on the bacterial cell and on bacterial pathogenesis and interactions with the host, will make it easier to find useful information about relevant topics, including graphical summaries for teaching.

William Margolin, Editor, *Journal of Bacteriology*
Department of Microbiology & Molecular Genetics
McGovern Medical School
Houston, Texas, USA
William.Margolin@uth.tmc.edu

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

MINIREVIEW

Friendly Fire: Biological Functions and Consequences of Chromosomal Targeting by CRISPR-Cas Systems

Gary E. Heussler, George A. O'Toole

Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

Clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) systems in bacteria and archaea target foreign elements, such as bacteriophages and conjugative plasmids, through the incorporation of short sequences (termed spacers) from the foreign element into the CRISPR array, thereby allowing sequence-specific targeting of the invader. Thus, CRISPR-Cas systems are typically considered a microbial adaptive immune system. While many of these incorporated spacers match targets on bacteriophages and plasmids, a noticeable number are derived from chromosomal DNA. While usually lethal to the self-targeting bacteria, in certain circumstances, these self-targeting spacers can have profound effects in regard to microbial biology, including functions beyond adaptive immunity. In this minireview, we discuss recent studies that focus on the functions and consequences of CRISPR-Cas self-targeting, including reshaping of the host population, group behavior modification, and the potential applications of CRISPR-Cas self-targeting as a tool in microbial biotechnology. Understanding the effects of CRISPR-Cas self-targeting is vital to fully understanding the spectrum of function of these systems.

Bacteria and archaea are under constant threat of viral predation and have evolved numerous mechanisms to defend against infection (1, 2). One such mechanism is the clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) protein system, which provides adaptive immunity against viruses and plasmids, collectively referred to as mobile genetic elements (MGEs) (3–5). While many divergent CRISPR-Cas systems exist, currently divided into 6 distinct types (6, 7), their general function is conserved. A CRISPR array is composed of short repeat sequences flanking unique spacer inserts transcribed by a promoter found in an adjacent AT-rich sequence (termed the leader) into a long precursor RNA molecule known as pre-CRISPR RNA (pre-crRNA). This pre-CRISPR RNA transcript is processed into multiple RNA molecules, known as mature crRNA, through nucleolytic cleavage at specific sequences in the repeats, with one exception being the type II-C system, in which mature crRNA is generated through transcription from promoters in each CRISPR (8). The mature crRNA then associates with Cas proteins to form the targeting CRISPR ribonucleoprotein complex (9, 10). The crRNA molecule is critical in host defense against MGEs, since the transcribed spacer provides the specificity of its target, and once bound through Watson-Crick base pairing, results in either degradation (type I) or cleavage (types II-VI) of the target through Cas protein nuclease activity.

The incorporation of a short sequence (usually 30 to 40 bp) from the invading MGE into the CRISPR array as a new spacer is a process termed CRISPR adaptation, a key step in CRISPR adaptive immunity. Analysis of the sequences of spacers from a given CRISPR array can serve as a history of previous CRISPR-Cas interactions with invading MGEs. The first evidence that CRISPR-Cas systems function as an immune system came from such an analysis, when in 2005, three separate groups examined a variety of CRISPR arrays and found that spacers matched sequences found in phages and plasmids (11–13). However, in addition to the matches with MGEs, spacers were also found that target sites on the bacterial or archaeal genome. For example, in *Yersinia* spp., of 36 spacers analyzed, the majority were of bacteriophage origin, but 8 spacers matched sequences on the *Yersinia* chromosome (13). A similar pattern was observed in a broader analysis of 4,500 spacers across archaea and bacteria, in which 35% of the spacers that matched sequences in the NCBI database were derived from chromosomal DNA and apparently were not related to foreign elements or prophages (12). Since these early observations, self-targeting spacers have consistently been found in CRISPR arrays (14–19), demonstrating that the insertion of a self-targeting spacer is not a rare event.

Given that at least one of the six types of CRISPR-Cas systems are present in 84% and 45% of sequenced archaeal and bacterial genomes, respectively (20), it is clear that these systems are as widespread as they are diverse. As research on these systems expands, and since a surprisingly large percentage of CRISPR spacers have been shown to have sequence identity to bacterial chromosomal targets, it is becoming clear that CRISPR-Cas systems can play a role in biological functions beyond adaptive immunity (21) and that self-targeting spacers can, at least in part, drive these alternative functions. This review aims to summarize current research on CRISPR-Cas self-targeting in prokaryotes and the role these events can play in important biological functions.

SELF-TARGETING WITH 100% COMPLEMENTARITY CAN DRIVE EVOLUTION

The most likely outcome of a 100% complementary, self-targeting spacer is cell death (Fig. 1A); such events have been experimentally demonstrated in multiple CRISPR-Cas types (22–24). For this reason, there are mechanisms that prevent a spacer from targeting the CRISPR array from which it was transcribed, such as the requirement of a protospacer adjacent motif (PAM) in the type I, II,

Accepted manuscript posted online 29 February 2016

Citation Heussler GE, O'Toole GA. 2016. Friendly fire: biological functions and consequences of chromosomal targeting by CRISPR-Cas systems. J Bacteriol 198:1481–1486. doi:10.1128/JB.00086-16.

Editor: W. Margolin

Address correspondence to George A. O'Toole, georgeo@dartmouth.edu.

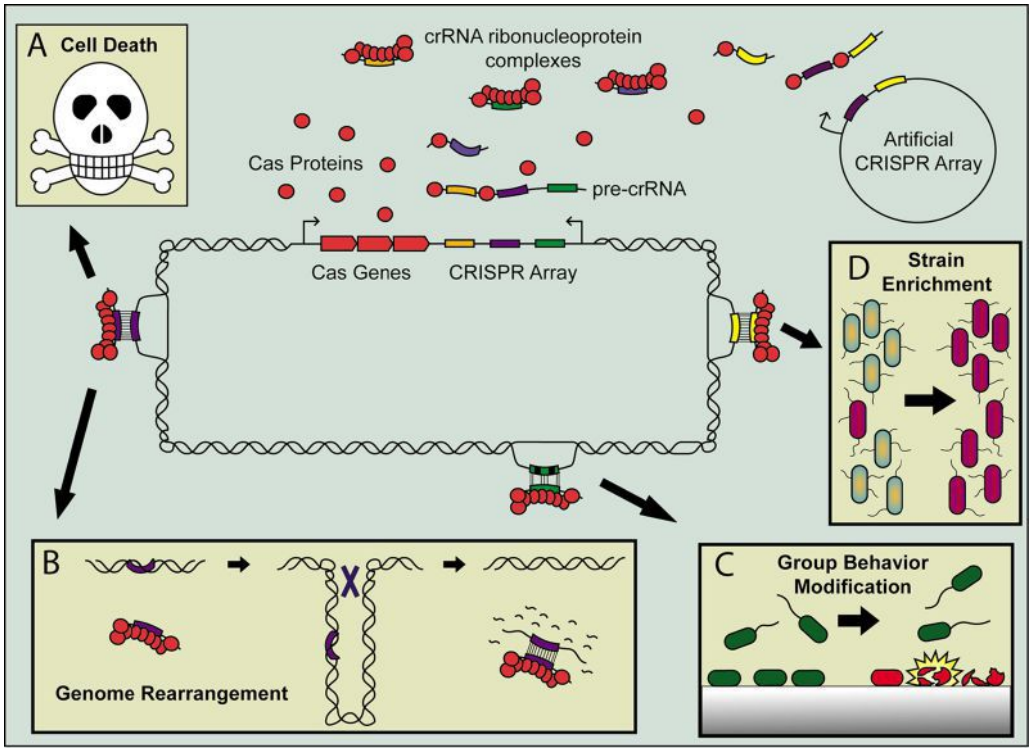


FIG 1 Consequences of CRISPR self-targeting in the type I system. The self-targeting crRNA can be transcribed from either the native CRISPR array or an artificial array on a plasmid, and the crRNA associates with the Cas proteins expressed from the native Cas-encoding operon to form the self-targeting crRNA ribonucleoprotein complex. These self-targeting spacers include both 100% complementary spacers (purple and yellow), as well as partially matching spacers (green and black). The most likely outcome of a 100% complementary match is cell death (A) due to the nucleolytic activity of the CRISPR-Cas system; however, this results in strong selection for mutants that remove the targeted region, for example, by genome rearrangement, such as excision of pathogenicity island or curing of a prophage (B). A partially matching self-targeting spacer can modulate group behavior, such as biofilm formation, by triggering viable planktonic cells (green cells) to die (red cells) upon surface binding (C). Bacterial populations transformed with a plasmid containing an artificial self-targeting spacer can enrich for bacterial strains that do not contain the target on their genome (purple cells, D).

V, and VI systems (25) or inactivation of the targeting CRISPR ribonucleoprotein complex through base pairing between the 5′ handle of the crRNA and the CRISPR repeats in the type III system (26). However, these mechanisms would not prevent cell death if a spacer is acquired that targets the host chromosome, and accordingly, evidence from the type I-E system of *Escherichia coli* suggests that the bacteria have evolved to preferentially sample foreign DNA over chromosomal DNA during naive CRISPR adaptation. It was observed previously in CRISPR adaptation studies using *E. coli* as a model system that the cell is roughly 100 to 1,000 times more likely to incorporate plasmid DNA over chromosomal DNA into its CRISPRs after normalizing for their respective size (27, 28). Recent evidence suggests one such mechanism for this preferential incorporation of foreign DNA into CRISPR regions during naive CRISPR adaptation is the involvement of the RecBCD complex. During DNA replication, when a double-stranded break occurs, RecBCD participates in the generation of single-stranded DNA, which can serve as a substrate for spacer acquisition by the Cas1-Cas2 complex. Chi sites, which are 8-nucleotide motifs well represented in the *E. coli* genome, limit the extent of single-stranded DNA generated by the RecBCD complex, and it was found the CRISPR-Cas system avoids the acquisition of DNA near Chi sites (27). Therefore, the lower prevalence of Chi sites on plasmids versus the *E. coli* genome, along with the preference of RecBCD to degrade linear DNA, such as recently injected phage DNA, might explain the observed preferential incorporation of foreign DNA. Cas1 has been shown to physically interact with RecB and RecC (29), and recently, a second group confirmed the requirement of the RecBCD complex during naive CRISPR adaptation in *E. coli* (30), further supporting this model.

Nevertheless, the frequency at which self-targeting spacers are found in CRISPR arrays suggests that, although less common than the incorporation of foreign DNA, incorporation of a self-targeting spacer can occur, perhaps impacting the biology of the host bacterial cell. This idea was directly tested in *Pectobacterium atrosepticum*. The native CRISPR2 array of *P. atrosepticum* contains a self-targeting spacer 100% complementary to a chromosomal gene within an ~100-kb horizontally acquired island named HAI2, but CRISPR lethality is abrogated due to a nonconsensus PAM (22). To assay the consequences of CRISPR self-targeting, *P. atrosepticum* was transformed with an inducible plasmid containing a truncated type I-F CRISPR1 leader, appropriate repeats, and an engineered spacer targeting the same gene in HAI2 as the native spacer but with the correct PAM. Induction of the self-targeting spacer-containing plasmid resulted in a cessation of bacterial growth and elongation of the bacterial cell indicative of DNA damage and the subsequent SOS response, confirming that a self-targeting spacer is cytotoxic (22). Interestingly, when the bacteria were left for 36 h, suppressor mutants arose in which CRISPR targeting was ablated, including deletions of the chromosomal

target. The size of the deletion differed between the mutants and included excision of the entire ~100-kb HAI2 island. A similar effect was observed when the spacer was designed to target a core region of the genome instead of HAI2, demonstrating that while a self-targeting spacer will normally result in cell death, it can also reshape the bacterial genome. Furthermore, these findings show that self-targeting is not dependent on the target being horizontally acquired, such as the integrative conjugative element HAI2.

A similar result was demonstrated using the native type II-A CRISPR-Cas system of *Streptococcus thermophilus*. When *S. thermophilus* was transformed with a plasmid carrying genes encoding artificial spacers targeting genomic islands within the transformed strain, >99% of the transformed bacteria were killed, but the surviving transformants contained large deletions of the targeted regions (31). The authors showed that the deletions were the result of recombination between insertion sequence elements (IS) on the *S. thermophilus* chromosome flanking the targeted region, and that the IS-dependent recombination events were occurring at a low frequency in the wild-type population. Taken together, these data indicate that the incorporation of a self-targeting spacer can drive evolution, but likely through selection of a small population of bacteria in which excision of the targeted region has occurred naturally (Fig. 1B).

SELF-TARGETING WITH 100% COMPLEMENTARITY CAN SELECT FOR PROPHAGE-CURED BACTERIA

When a bacterium is lysogenized by a particular phage, the phage genome is for all intents and purposes part of the bacterial genome. Thus, another potential benefit of a self-targeting spacer is its ability to incorporate a spacer targeting the prophage, which will reshape a bacterial population by selecting for cells that have excised the prophage from their chromosome (Fig. 1B). This notion was tested in an *hns* mutant strain of *E. coli* in which the native CRISPR system, normally heat-stable nucleoid-structuring (H-NS) silenced, was active. The *E. coli* strain harboring a lambda phage with a temperature-sensitive *c*I gene allowing controlled lytic induction was transformed with an inducible plasmid containing either an artificial spacer targeting the prophage or a control spacer with no target. Upon simultaneous induction of both the prophage and the artificial spacer-containing plasmid, the majority of cells (>99%) were killed; however, the survival frequency of cells transformed with the prophage-targeting spacer was 500-fold higher than the survival frequency of cells transformed with the control spacer (32). These data suggest that the CRISPR-Cas system, under the right circumstances, can select for cells that have excised the phage.

The ability to select for cells that have cured a prophage is likely not limited to *E. coli*. In *Streptococcus pyogenes*, 13 sequenced strains were assayed for CRISPR arrays, and of the 8 CRISPR-positive strains, 41 distinct spacers were identified, with 26 spacers found to target *S. pyogenes* prophages. However, no strain contained a spacer targeting a prophage on its own genome (33). Additionally, it was found the CRISPR-negative strains contained significantly more prophages than the CRISPR-positive strains, and furthermore, within the 8 CRISPR-positive *S. pyogenes* strains, an inverse correlation was found between the number of spacers per genome and the number of prophages within each genome. Similarly, recently, the CRISPR-Cas systems of the genus *Bifidobacterium* were analyzed, and 25 out of 32 characterized type I and type II CRISPR-Cas systems contained at least one spacer sequence targeting a *Bifidobacterium* lysogenic prophage, including two instances in which the *Bifidobacterium* species included a spacer targeting a prophage on its own genome (16). Similar to *S. pyogenes*, a positive correlation was found between strains lacking a CRISPR-Cas system and the number of times prophages were found on the chromosome targeted by spacers in other *Bifidobacterium* CRISPR-Cas systems. For example, *Bifidobacterium angulatum* contains 26 unique spacers that target prophages on 11 other *Bifidobacterium* species genomes while harboring no spacer-targeted prophages on its own genome.

Additionally, a bioinformatic analysis of 365 *Pseudomonas aeruginosa* CRISPR-Cas systems from 672 sequenced *P. aeruginosa* strains found that 55% of the 2,823 unique spacers matched sites in the *P. aeruginosa* genome, the vast majority of which are predicted to target the *P. aeruginosa* accessory genome, including potential prophages (17). The *P. aeruginosa* strains lacking CRISPR-Cas systems were on average 300 kbp larger than those with a CRISPR-Cas system. Thus, it is possible the CRISPR-Cas systems of *S. pyogenes*, *Bifidobacterium* spp., and *P. aeruginosa* are either actively preventing the uptake of MGEs (including prophages) or incorporating self-targeting spacers that select for cells that have purged these MGEs in a manner analogous to the experimentally demonstrated removal of genomic regions in *P. atrosepticum* (22) and *S. thermophilus* (31).

SELF-TARGETING WITH 100% COMPLEMENTARITY IN THE TYPE III-A SYSTEM CAN HELP MICROBES TOLERATE TEMPERATE PHAGES BY BLOCKING THE ENTRY OF PROPHAGES INTO THE LYTIC CYCLE

While the presence of a lysogenic phage on a bacterial chromosome can be detrimental to bacterial fitness (34), it has been well established that prophages can provide beneficial functions to, and increase the fitness of, the bacterial host (35, 36). For this reason, curing a phage from the bacterial chromosome can have a selective disadvantage. Unlike the type I and II systems discussed above, a type III-A/B CRISPR-Cas array can potentially include a self-targeting spacer without associated cytotoxicity due to general transcription-dependent targeting of the type III-A/B systems (37–39).

A recent study of *Staphylococcus aureus* showed that when spacers targeted lytic genes silenced during lysogeny, no CRISPR interference through the type III-A system was detected. CRISPR interference could only be achieved when the target gene was expressed; this finding was demonstrated by integrating a CRISPR-targeted gene under the control of a tightly regulated inducible promoter on the *S. aureus* chromosome. This strain was then transformed with a plasmid expressing the spacer targeting the integrated gene, and CRISPR interference was detected only upon induction of the chromosomal promoter driving expression of the targeted gene. Additionally, only spacers targeting the nontemplate strand of the targeted gene can generate CRISPR interference, illustrating transcription-dependent targeting, and furthermore, distinguishing the targeting requirements of the type III system from the type I and II systems (37, 40). The implication of these findings is that bacteria with spacers capable of self-targeting a prophage can tolerate a 100% complementary self-match and reap the putative benefits of a prophage while also preventing phage-mediated lysis.

SELF-TARGETING WITH PARTIAL COMPLEMENTARITY CAN DRIVE GROUP BEHAVIOR

The research discussed so far has focused on the interaction of a spacer that is 100% complementary with a DNA target on (or associated with) the host chromosome, but a partially matching spacer can still elicit an effect without the lethality typically associated with a 100% complementary spacer. Generally, partially matching spacers are thought to be important in CRISPR priming, in which the presence of a spacer partially matching an MGE will greatly increase the likelihood of incorporating new spacers against the MGE, and particularly sequences that match regions adjacent to the partially matched target (41–45). Additionally, it was recently found that spacers with a mismatch in regions considered indispensable for CRISPR interference, such as the seed sequence or PAM, may still be capable of functioning directly in CRISPR-mediated interference (46). Therefore, self-targeting spacers with partial complementarity can likely still drive biological functions.

The best example of a partially complementary spacer impacting microbial behavior is the modification of group behavior by the type I-F CRISPR-Cas system in *P. aeruginosa*. When *P. aeruginosa* is lysogenized by the bacteriophage DMS3, the interaction of a partially complementary spacer encoded by the *P. aeruginosa* CRISPR2 array with a target on the DMS3 prophage inhibits biofilm formation and swarming motility, two group behaviors important for virulence (47, 48). These altered biofilm and swarming phenotypes are dependent on the nickase activity of the effector protein Cas3, since a mutation in the catalytic region of the nuclease domain of Cas3 fully restores biofilm formation and swarming motility (48). The 2 mismatches at the +9 and +11 positions on the spacer prevent a lethal self-targeting event but promote sufficient CRISPR targeting activity to induce RecA, which in turn directly induces the expression of SOS-regulated, phage-related autolysis genes that induce cell death upon the bacteria attaching to a surface. Induction of the phage-related autolysis genes effectively inhibits biofilm and swarming motility through death of the surface-associated bacteria while having no noticeable effect on the planktonic population (Fig. 1C) (49). Given that numerous *P. aeruginosa* Mu-like phages contain the same protospacer target as DMS3 (50, 51), it is unclear if this interaction was directly selected or simply a side effect of CRISPR-Cas immunity. Nevertheless, the demonstration that a partially matching spacer can elicit such a strong biological response, in combination with the frequency of self-targeting spacers present in CRISPR arrays across both bacteria and archaea, suggests the potential of other partial matching interactions impacting important aspects of bacterial biology. It is unlikely that the biological impact of partial complementarity of a CRISPR spacer is limited to *P. aeruginosa*.

SELF-TARGETING CAN BE EXPLOITED FOR USE IN BIOTECHNOLOGY

CRISPR-Cas research has received a great deal of attention, largely for use of the type II system (52, 53) and the recently described type V system (54), in the genetic engineering of eukaryotic systems. Recent work has demonstrated that CRISPR-Cas technology is not limited to the engineering of eukaryotic systems and can be used in bacterial and archaeal systems as well. For example, for the type I-E system, it has been shown that *E. coli* lacking the effector enzyme Cas3 can be transformed with plasmids harboring self-targeting spacers without lethal effects, and designing these spacers to target the promoter of a particular gene will result in silencing of that particular gene due to the binding of the Cascade complex (55, 56). Exploiting CRISPR-Cas systems for such engineering approaches is likely not limited to the type I-E system and would potentially allow the endogenous type I systems of any number of bacteria to be coopted for CRISPR-mediated genetic control. Such approaches may be particularly useful in nonmodel organisms.

As another example, it has been shown that the cytotoxic effects of CRISPR self-targeting, such as those described earlier (22), can be used to remove particular strains of bacteria from a population. A mixed but equal population of *E. coli* K-12 and *E. coli* B, which share 99% sequence identity, was transformed with plasmids harboring self-targeting spacers unique to either sequences of strain K-12 or B (Fig. 1D), and the resulting transformants were composed almost entirely of the nontargeted strain (>99.9%) due to CRISPR interference from the self-targeting spacers (57). This cytotoxicity is the basis of "CRISPR antimicrobials," in which antibiotic strains of a particular bacterial species, including *E. coli* (58) and *S. aureus* (59), are selected out of a population through CRISPR self-targeting of antibiotic resistance genes.

CONCLUSIONS AND FUTURE PERSPECTIVES

While the incorporation of a self-targeting spacer into a CRISPR array is typically a lethal event for a microbe, the surprisingly high frequency of self-targeting spacers identified from sequenced CRISPR arrays indicates that the occurrence of CRISPR-Cas self-targeting is not rare in nature. Self-targeting spacers with 100% complementarity in a type I or II system present the microbe with the challenge of avoiding CRISPR-induced lethality, which can reshape bacterial populations by selecting for bacteria in which genome rearrangement has occurred, such as excising pathogenicity islands or curing of a prophage. There are several examples in bacteria in which CRISPR-positive strains have a smaller genome than that of CRISPR-negative strains, and it is unclear if the reduction in genome size (specifically MGEs) is the result of the CRISPR-Cas system actively preventing the acquisition of MGEs or if incorporation of self-targeting spacers is selecting for bacteria in the population that have purged their genome of any of these invading sequences. It is important to note that microbes containing a type III CRISPR system are more tolerant of self-targeting spacers, since target transcription is required for CRISPR interference, allowing the microbe to tolerate a prophage when the targeted gene is silenced, which is the case for most phage genes during lysogeny.

Self-targeting is not limited to 100% complementary spacers; a self-targeting spacer can still have meaningful biological consequences even with multiple mismatches between the spacer and its genomic target, such as group behavior modification in *P. aeruginosa*. This is a largely unexplored area of research, and with the sheer number of potential partially matching spacers already identified in bacterial genomes, it is tempting to speculate that other microbial behaviors can be attributed to CRISPR self-targeting, especially in light of recent studies demonstrating functional CRISPR interference even with multiple mutations disrupting spacer-protospacer complementarity (46, 60). This is a largely unexplored area of research.

Overall, we are only beginning to understand the role of CRISPR self-targeting, and with the diversity of CRISPR-Cas

types and variety of biological consequences of a self-targeting spacer that were already demonstrated, chromosomal-targeting spacers are gaining attention. It will be exciting to see future advances in our understanding of the biological functions and continued applications of CRISPR self-targeting in biotechnology.

ACKNOWLEDGMENTS

This study is supported by National Institutes of Health grants R37 AI83256-06 and R01-AI097307 and NSF grant MCB-9984521to G.A.O.

FUNDING INFORMATION

This work, including the efforts of George O'Toole, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (AI83256-06). This work was funded by National Science Foundation (NSF) (MCB-9984521). This work was funded by HHS | NIH | National Institute of General Medical Sciences (NIGMS) (R01-AI097307).

REFERENCES

1. **Westra ER, Swarts DC, Staals RH, Jore MM, Brouns SJ, van der Oost J.** 2012. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu Rev Genet **46:**311–339. http://dx.doi.org/10.1146/annurev-genet-110711-155447.
2. **Barrangou R, Marraffini LA.** 2014. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Mol Cell **54:**234–244. http://dx.doi.org/10.1016/j.molcel.2014.03.011.
3. **Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P.** 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science **315:**1709–1712. http://dx.doi.org/10.1126/science.1138140.
4. **Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J.** 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science **321:**960–964. http://dx.doi.org/10.1126/science.1159689.
5. **Marraffini LA, Sontheimer EJ.** 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science **322:**1843–1845. http://dx.doi.org/10.1126/science.1165771.
6. **Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH, Horvath P, Moineau S, Mojica FJ, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der Oost J, Backofen R, Koonin EV.** 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol **13:**722–736. http://dx.doi.org/10.1038/nrmicro3569.
7. **Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Minakhin L, Joung J, Konermann S, Severinov K, Zhang F, Koonin EV.** 2015. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol Cell **60:**385–397. http://dx.doi.org/10.1016/j.molcel.2015.10.008.
8. **Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, Vogel J, Sontheimer EJ.** 2013. Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. Mol Cell **50:**488–503. http://dx.doi.org/10.1016/j.molcel.2013.05.001.
9. **Reeks J, Naismith JH, White MF.** 2013. CRISPR interference: a structural perspective. Biochem J **453:**155–166. http://dx.doi.org/10.1042/BJ20130316.
10. **Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJ, van der Oost J, Doudna JA, Nogales E.** 2011. Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature **477:**486–489. http://dx.doi.org/10.1038/nature10402.
11. **Bolotin A, Quinquis B, Sorokin A, Ehrlich SD.** 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology **151:**2551–2561. http://dx.doi.org/10.1099/mic.0.28048-0.
12. **Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E.** 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol **60:**174–182. http://dx.doi.org/10.1007/s00239-004-0046-3.
13. **Pourcel C, Salvignol G, Vergnaud G.** 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology **151:**653–663. http://dx.doi.org/10.1099/mic.0.27437-0.
14. **Brodt A, Lurie-Weinberger MN, Gophna U.** 2011. CRISPR loci reveal networks of gene exchange in archaea. Biol Direct **6:**65. http://dx.doi.org/10.1186/1745-6150-6-65.
15. **Pleckaityte M, Zilnyte M, Zvirbliene A.** 2012. Insights into the CRISPR/Cas system of *Gardnerella vaginalis*. BMC Microbiol **12:**301. http://dx.doi.org/10.1186/1471-2180-12-301.
16. **Briner AE, Lugli GA, Milani C, Duranti S, Turroni F, Gueimonde M, Margolles A, van Sinderen D, Ventura M, Barrangou R.** 2015. Occurrence and diversity of CRISPR-Cas systems in the genus *Bifidobacterium*. PLoS One **10:**e0133661. http://dx.doi.org/10.1371/journal.pone.0133661.
17. **van Belkum A, Soriaga LB, LaFave MC, Akella S, Veyrieras JB, Barbu EM, Shortridge D, Blanc B, Hannum G, Zambardi G, Miller K, Enright MC, Mugnier N, Brami D, Schicklin S, Felderman M, Schwartz AS, Richardson TH, Peterson TC, Hubby B, Cady KC.** 2015. Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant *Pseudomonas aeruginosa*. mBio **6**(6):e01796-15. http://dx.doi.org/10.1128/mBio.01796-15.
18. **Stern A, Keren L, Wurtzel O, Amitai G, Sorek R.** 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? Trends Genet **26:**335–340. http://dx.doi.org/10.1016/j.tig.2010.05.008.
19. **Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R.** 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int J Food Microbiol **131:**62–70. http://dx.doi.org/10.1016/j.ijfoodmicro.2008.05.030.
20. **Grissa I, Vergnaud G, Pourcel C.** 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics **8:**172. http://dx.doi.org/10.1186/1471-2105-8-172.
21. **Westra ER, Buckling A, Fineran PC.** 2014. CRISPR-Cas systems: beyond adaptive immunity. Nat Rev Microbiol **12:**317–326. http://dx.doi.org/10.1038/nrmicro3241.
22. **Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, Clulow JS, Richter C, Przybilski R, Pitman AR, Fineran PC.** 2013. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. PLoS Genet **9:**e1003454. http://dx.doi.org/10.1371/journal.pgen.1003454.
23. **Aklujkar M, Lovley DR.** 2010. Interference with histidyl-tRNA synthetase by a CRISPR spacer sequence as a factor in the evolution of *Pelobacter carbinolicus*. BMC Evol Biol **10:**230. http://dx.doi.org/10.1186/1471-2148-10-230.
24. **Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA.** 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol **31:**233–239. http://dx.doi.org/10.1038/nbt.2508.
25. **Mojica FJ, Díez-Villaseñor C, García-Martínez J, Almendros C.** 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology **155:**733–740. http://dx.doi.org/10.1099/mic.0.023960-0.
26. **Marraffini LA, Sontheimer EJ.** 2010. Self vs. non-self discrimination during CRISPR RNA-directed immunity. Nature **463:**568–571. http://dx.doi.org/10.1038/nature08703.
27. **Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, Edgar R, Qimron U, Sorek R.** 2015. CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature **520:**505–510. http://dx.doi.org/10.1038/nature14302.
28. **Díez-Villaseñor C, Guzman NM, Almendros C, García-Martínez J, Mojica FJ.** 2013. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. RNA Biol **10:**792–802. http://dx.doi.org/10.4161/rna.24023.
29. **Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, Gagarinova A, Pogoutse O, Brown G, Binkowski A, Phanse S, Joachimiak A, Koonin EV, Savchenko A, Emili A, Greenblatt J, Edwards AM, Yakunin AF.** 2011. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Mol Microbiol **79:**484–502. http://dx.doi.org/10.1111/j.1365-2958.2010.07465.x.
30. **Ivančić-Baće I, Cass SD, Wearne SJ, Bolt EL.** 2015. Different genome stability proteins underpin primed and naive adaptation in *E. coli* CRISPR-Cas immunity. Nucleic Acids Res **43:**10821–10830. http://dx.doi.org/10.1093/nar/gkv1213.
31. **Selle K, Klaenhammer TR, Barrangou R.** 2015. CRISPR-based screening of genomic island excision events in bacteria. Proc Natl Acad Sci U S A **112:**8076–8081. http://dx.doi.org/10.1073/pnas.1508525112.
32. **Edgar R, Qimron U.** 2010. The *Escherichia coli* CRISPR system protects

from lambda lysogenization, lysogens, and prophage induction. J Bacteriol **192:**6291–6294. http://dx.doi.org/10.1128/JB.00644-10.
33. **Nozawa T, Furukawa N, Aikawa C, Watanabe T, Haobam B, Kurokawa K, Maruyama F, Nakagawa I.** 2011. CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. PLoS One **6:**e19543. http://dx.doi.org/10.1371/journal.pone.0019543.
34. **DeBardeleben HK, Lysenko ES, Dalia AB, Weiser JN.** 2014. Tolerance of a phage element by *Streptococcus pneumoniae* leads to a fitness defect during colonization. J Bacteriol **196:**2670–2680. http://dx.doi.org/10.1128/JB.01556-14.
35. **Brüssow H, Canchaya C, Hardt WD.** 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev **68:**560–602, table of contents. http://dx.doi.org/10.1128/MMBR.68.3.560-602.2004.
36. **Cumby N, Davidson AR, Maxwell KL.** 2012. The moron comes of age. Bacteriophage **2:**225–228.
37. **Goldberg GW, Jiang W, Bikard D, Marraffini LA.** 2014. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature **514:**633–637. http://dx.doi.org/10.1038/nature13637.
38. **Deng L, Garrett RA, Shah SA, Peng X, She Q.** 2013. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. Mol Microbiol **87:**1088–1099. http://dx.doi.org/10.1111/mmi.12152.
39. **Jiang W, Samai P, Marraffini LA.** 2016. Degradation of phage transcripts by CRISPR-associated RNases enables type III CRISPR-Cas immunity. Cell **164:**710–721. http://dx.doi.org/10.1016/j.cell.2015.12.053.
40. **Goldberg GW, Marraffini LA.** 2015. Resistance and tolerance to foreign elements by prokaryotic immune systems–curating the genome. Nat Rev Immunol **15:**717–724. http://dx.doi.org/10.1038/nri3910.
41. **Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E.** 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat Commun **3:**945. http://dx.doi.org/10.1038/ncomms1937.
42. **Fineran PC, Gerritzen MJ, Suarez-Diez M, Kunne T, Boekhorst J, van Hijum SA, Staals RH, Brouns SJ.** 2014. Degenerate target sites mediate rapid primed CRISPR adaptation. Proc Natl Acad Sci U S A **111:**E1629–E1638. http://dx.doi.org/10.1073/pnas.1400071111.
43. **Richter C, Dy RL, McKenzie RE, Watson BN, Taylor C, Chang JT, McNeil MB, Staals RH, Fineran PC.** 2014. Priming in the type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. Nucleic Acids Res **42:**8516–8526. http://dx.doi.org/10.1093/nar/gku527.
44. **Savitskaya E, Semenova E, Dedkov V, Metlitskaya A, Severinov K.** 2013. High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in *E. coli*. RNA Biol **10:**716–725. http://dx.doi.org/10.4161/rna.24325.
45. **Swarts DC, Mosterd C, van Passel MW, Brouns SJ.** 2012. CRISPR interference directs strand specific spacer acquisition. PLoS One **7:**e35888. http://dx.doi.org/10.1371/journal.pone.0035888.
46. **Xue C, Seetharam AS, Musharova O, Severinov K, J Boruns SJ, Severin AJ, Sashital DG.** 2015. CRISPR interference and priming varies with individual spacer sequences. Nucleic Acids Res **43:**10831–10847. http://dx.doi.org/10.1093/nar/gkv1259.
47. **Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA.** 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. J Bacteriol **191:**210–219. http://dx.doi.org/10.1128/JB.00797-08.
48. **Cady KC, O'Toole GA.** 2011. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. J Bacteriol **193:**3433–3445. http://dx.doi.org/10.1128/JB.01411-10.
49. **Heussler GE, Cady KC, Koeppen K, Bhuju S, Stanton BA, O'Toole GA.** 2015. Clustered regularly interspaced short palindromic repeat-dependent, biofilm-specific death of *Pseudomonas aeruginosa* mediated by increased expression of phage-related genes. mBio **6**(3)**:**e00129-15. http://dx.doi.org/10.1128/mBio.00129-15.
50. **Braid MD, Silhavy JL, Kitts CL, Cano RJ, Howe MM.** 2004. Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*. J Bacteriol **186:**6560–6574. http://dx.doi.org/10.1128/JB.186.19.6560-6574.2004.
51. **Cazares A, Mendoza-Hernández G, Guarneros G.** 2014. Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. BMC Genomics **15:**1146. http://dx.doi.org/10.1186/1471-2164-15-1146.
52. **Terns RM, Terns MP.** 2014. CRISPR-based technologies: prokaryotic defense weapons repurposed. Trends Genet **30:**111–118. http://dx.doi.org/10.1016/j.tig.2014.01.003.
53. **Doudna JA, Charpentier E.** 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science **346:**1258096.
54. **Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F.** 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell **163:**759–771. http://dx.doi.org/10.1016/j.cell.2015.09.038.
55. **Luo ML, Mullis AS, Leenay RT, Beisel CL.** 2015. Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. Nucleic Acids Res **43:**674–681. http://dx.doi.org/10.1093/nar/gku971.
56. **Rath D, Amlinger L, Rath A, Lundgren M.** 2015. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie **117:**119–128. http://dx.doi.org/10.1016/j.biochi.2015.03.025.
57. **Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL.** 2014. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. mBio **5**(1)**:**e00928-13. http://dx.doi.org/10.1128/mBio.00928-13.
58. **Citorik RJ, Mimee M, Lu TK.** 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat Biotechnol **32:**1141–1145. http://dx.doi.org/10.1038/nbt.3011.
59. **Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, Fischetti VA, Marraffini LA.** 2014. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat Biotechnol **32:**1146–1150. http://dx.doi.org/10.1038/nbt.3043.
60. **Maniv I, Jiang W, Bikard D, Marraffini LA.** 2016. Impact of different target sequences on type III CRISPR-Cas immunity. J Bacteriol **198:**941–950. http://dx.doi.org/10.1128/JB.00897-15.

JB Journals.ASM.org

Natural Competence and the Evolution of DNA Uptake Specificity

Joshua Chang Mell, Rosemary J. Redfield

Department of Zoology, University of British Columbia, Vancouver, BC, Canada

Many bacteria are naturally competent, able to actively transport environmental DNA fragments across their cell envelope and into their cytoplasm. Because incoming DNA fragments can recombine with and replace homologous segments of the chromosome, competence provides cells with a potent mechanism of horizontal gene transfer as well as access to the nutrients in extracellular DNA. This review starts with an introductory overview of competence and continues with a detailed consideration of the DNA uptake specificity of competent proteobacteria in the *Pasteurellaceae* and *Neisseriaceae*. Species in these distantly related families exhibit strong preferences for genomic DNA from close relatives, a self-specificity arising from the combined effects of biases in the uptake machinery and genomic overrepresentation of the sequences this machinery prefers. Other competent species tested lack obvious uptake bias or uptake sequences, suggesting that strong convergent evolutionary forces have acted on these two families. Recent results show that uptake sequences have multiple "dialects," with clades within each family preferring distinct sequence variants and having corresponding variants enriched in their genomes. Although the genomic consensus uptake sequences are 12 and 29 to 34 bp, uptake assays have found that only central cores of 3 to 4 bp, conserved across dialects, are crucial for uptake. The other bases, which differ between dialects, make weaker individual contributions but have important cooperative interactions. Together, these results make predictions about the mechanism of DNA uptake across the outer membrane, supporting a model for the evolutionary accumulation and stability of uptake sequences and suggesting that uptake biases may be more widespread than currently thought.

Naturally competent bacteria actively pull DNA fragments from their environment into their cells. These fragments provide nucleotides, but high similarity with the chromosome also allows them to change the cell's genotype by homologous recombination, a process called natural transformation (Fig. 1; reviewed in references 1, 2, 3, 4, and 5). Most competent bacteria that have been tested can take up DNA from any source, but species in two distantly related families of Gram-negative bacteria, the *Pasteurellaceae* and the *Neisseriaceae*, show strong preferences for DNAs containing short sequences that are highly overrepresented in their own genomes, leading to preferential uptake of conspecific DNA (6). This self-specificity both raises questions about and provides a tool for investigating the evolution of competence and the mechanism of DNA uptake.

We begin with a general overview of competence, emphasizing the evolutionary issues. We then describe the two components that together create self-specificity, sequence biases in the DNA uptake machinery and overrepresentation of uptake sequences in the genomes, highlighting recent work on the mechanism and evolution of uptake biases in the two families and an evolutionary model that accounts for uptake sequences as an accidental consequence of these biases. We then consider how uptake biases may themselves result from selection for more-efficient DNA uptake and integrate this into a more unified perspective on natural competence and uptake specificity in all bacteria.

THE MECHANISM AND FUNCTION OF NATURAL COMPETENCE

How is DNA transported into the cell? Gram-negative bacteria take up DNA in two stages, customarily referred to as DNA uptake (across the outer membrane) and DNA translocation (across the inner membrane) (3). After DNA is bound to the cell (step A in Fig. 1), uptake occurs by retraction of cell surface fibers of the type IV pilus family (T4P), which pulls the DNA into the periplasm through secretin pores in the outer membrane (step B). In some species, the same pilus proteins form the long pili used for adhesion and twitching motility (7), but in others, the fibers are inferred to occur only as short competence-specific pseudopili that do not protrude detectably beyond the cell surface, like those used in type II secretion (8, 9). Gram-positive bacteria use similar T4P-related proteins to pull double-stranded DNA across their thick cell walls (1, 10).

Although the main components of the uptake machinery have been identified in multiple species, little is known about how they interact with DNA, and the detailed mechanics of DNA uptake are not well understood in any species (7). DNA uptake presents a particular challenge to Gram-negative bacteria, which must transport stiff and highly charged double-stranded DNA (dsDNA) across their outer membranes. Although uptake could in principle be initiated by threading one dsDNA end through the secretin pore, alongside or at the tip of the pilus/pseudopilus, this has not been demonstrated in any system. Instead, initiation likely occurs at internal sites on DNA fragments, as shown in *Haemophilus influenzae*, where closed circular plasmids are taken up into the periplasm as efficiently as linear DNAs (11). Although the power source for retraction of T4P is usually thought to be ATP-powered disassembly of the pilin subunits by the motor protein PilT (12), several competent species lack *pilT* homologs (13). The ability of cells to take up DNA fragments of >50 kb, much longer than the cell (14–16), creates another difficulty, since pulling these across the very narrow periplasmic space by pseudopilus retraction

Published ahead of print 31 January 2014

Address correspondence to Rosemary J. Redfield, redfield@zoology.ubc.ca.

doi:10.1128/JB.01293-13

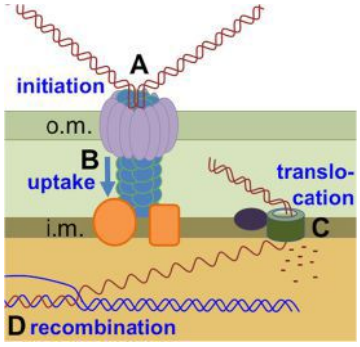


FIG 1 DNA uptake and transformation by competent Gram-negative bacteria. (A) dsDNA is bound at the cell surface. (B) DNA is pulled through the type II secretin pore by retraction of a type 4 pilus (T4P). (C) One strand is translocated intact into the cytoplasm by the Rec2/ComEC protein; the other is degraded. (D) The new strand recombines with a homologous sequence in the chromosome, displacing the resident strand. The abbreviations "o.m." and "i.m." refer to the outer and inner cell membranes, respectively.

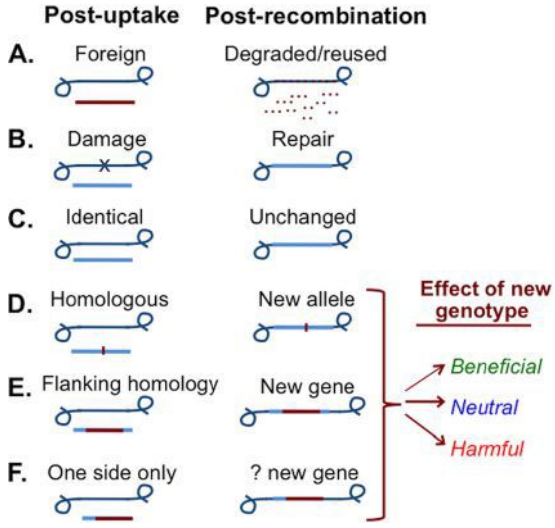


FIG 2 Biochemical and informational consequences of DNA uptake by competent cells.

would require many cycles of pseudopilus elongation, DNA attachment, and retraction.

Once the DNA is in the periplasm (in Gram-negative bacteria) or at the cytoplasmic membrane (in Gram-positive bacteria), one strand is translocated across the membrane into the cytoplasm, with its 3′ end leading (step C). Although Gram-positive bacteria use a cell surface nuclease to cut DNA into smaller fragments before this step (17), Gram-negative bacteria do not. The strand that is not translocated is degraded at the membrane surface to nucleotides, which can then be dephosphorylated and taken up by nucleoside transporters (18). In all competent species, translocation uses a conserved membrane pore encoded by *rec2* or *comEC*. If sequence similarity permits, the translocated strand may then replace a chromosomal strand by homologous recombination (step D); otherwise, it is degraded and its nucleotide subunits are recycled.

What are the consequences of DNA uptake? Because DNA has both biochemical and informational properties, the effects of DNA uptake depend on the nutritional needs of the cell, the presence of DNA damage, the ability of incoming DNA to recombine with chromosomal DNA, and the effects of this recombination on fitness, as summarized by Fig. 2. The most immediate consequence is nutritional (Fig. 2A). DNA is an excellent source of the deoxyribonucleotides needed for replication of the cell's own genome, and *de novo* nucleotide synthesis is expensive both in terms of energy and in terms of the molecular constituents. All cells take up preformed bases and nucleotides where possible, and soil and sediment species often secrete nucleases that allow them to use extracellular DNA as a nutrient source (19). However, uptake of intact DNA is a more efficient way to obtain nucleotides, both because it limits losses due to diffusion and because it avoids the need for nucleoside rephosphosphoryation after uptake (18). Most incoming DNA is degraded even when it is identical to that of the chromosome, and even recombined DNA reduces the cell's need for nucleotides (20, 21).

The genetic consequences of DNA uptake are less predictable. First, they depend on the sequence of the DNA being sufficiently similar to DNA in the cell's genome that it can replace a genomic strand by homologous recombination, catalyzed by the ubiquitous RecA protein. If this replaced segment includes a position with DNA damage (Fig. 2B), the incoming strand could provide a template for DNA repair (22). However, unless the genome is heavily damaged, most recombination occurs at undamaged positions. When the incoming DNA is from cells of the same clonal population (Fig. 2C), recombination does not change the cells' genotypes unless the donor or recipient DNA contains newly arisen mutations. A recent study comparing wild-type and competence-deficient strains of *Streptococcus pneumoniae* found that competence reduced the fixation of new mutations, suggesting that transformation eliminates new mutations from recipients more often than it introduces them from donors (23).

Natural populations may often be mixtures of strains, and transformation then creates new and possibly beneficial combinations of variant alleles and loci (Fig. 2D to F). DNA from one of the common noncompetent strains (see below) (24) may also replace functional competence genes with nonfunctional alleles (24). To the extent that DNA comes from relatives that have died due to deleterious mutations, recombination reduces fitness more often than it increases it (24, 25), but it may also provide cells with locally beneficial alleles if they are invading an established population (26).

Transformation is not limited to simple sequence variants; large insertions and deletions transform moderately well if flanked by sequences of chromosomal homology (Fig. 2E). Homology at only one end of a heterologous segment can be sufficient to promote recombination (Fig. 2F; line thicknesses indicate probable chance of different outcomes) (27), but "illegitimate" recombination with nonhomologous DNA is extremely rare (28). Although transformation's dependence on sequence homology makes it inefficient at introducing novel genes into a species, this is balanced by the high efficiency with which it can spread genes through populations once they have been introduced by such homology-independent processes as specialized transduction or transposition.

Factors that evolved to protect cells against genetic parasites can also limit transformation. Lin et al. (29) found that recombination tracts in *H. pylori* often terminated at restriction sites where the donor DNA was unmethylated. This is unlikely to be due to

action of recipient-specific enzymes in the cytoplasm, since incoming DNA is single-stranded and thus not a target for most restriction enzymes. However, restriction enzymes released by cell lysis could also cut donor DNAs before uptake, and these may be an important limitation on the extent of recombination tracts. Similarly, Bikard et al. (30) showed that a clustered regularly interspaced short palindromic repeat (CRISPR) element engineered to target capsule genes prevented a *S. pneumoniae* strain from acquiring these genes by transformation. However, the likelihood of CRISPR elements acquiring sequences that target functional accessory genes has not been evaluated.

How is competence regulated? Unlike other uptake systems regulated by substrate availability (e.g., the *lac* operon), no bacteria are known to use DNA availability to signal induction of the competence machinery; this may be because environmental DNA is ubiquitous in most bacterial environments, especially biofilms. Instead, in most well-studied species, competence is regulated by other environmental and biochemical cues (4). The exception is *Neisseria*, whose competence appears to be constitutive (31). In these species, genes encoding proteins needed for DNA uptake and translocation are typically coregulated with genes for cytoplasmic proteins. Some of the latter contribute to transformational recombination, but others have no obvious connection to DNA uptake or have no known function at all. Unfortunately, the frequent coregulation of other cellular functions with competence makes it difficult to confidently delineate "competence regulons."

The simplest known regulatory system is that of *H. influenzae*, where cells respond to a lack of phosphotransferase system (PTS) sugars and purine precursors by inducing expression of 25 genes from 12 operons, under the control of the catabolite regulator cyclic AMP receptor protein (CRP) and its competence-specific cofactor Sxy (also called TfoX) (32–36). The functions of the Sxy-regulated genes have been examined using knockout mutations, confirming that most play direct roles in DNA uptake (37). Similar regulons induced by Sxy and CRP are present in other *Pasteurellaceae* and in the related *Vibrionaceae* and *Enterobacteriaceae*. Work in *Vibrio* sp. has shown that competence induction also depends on quorum sensing, pyrimidine precursors, and the presence of chitin breakdown products generated from crustacean exoskeletons (38–42). Although *Escherichia coli*'s competence gene homologs enable cells to use DNA as a sole source of carbon, attempts to demonstrate natural competence in *E. coli* have not been successful (43, 44). Less is known about regulatory mechanisms in other Gram-negative species, since culture conditions that induce competence have not been linked to specific regulators (4, 45). Gram-positive competence regulation is complex, with contributions from overlapping layers of quorum sensing, nutritional signals, and other stress responses (46–49). The competence regulons are also much larger and include many genes whose relationship to competence is not evident (4, 50–52). The extreme variability of competence regulation contrasts with the strong conservation of the DNA uptake machinery and perhaps reflects various benefits of coregulating DNA uptake with other cellular responses.

What is the function of natural competence? Although the action of natural selection with respect to competence genes remains controversial, the immediate consequences of DNA uptake (Fig. 2) provide a framework for thinking about its possible evolutionary function. Competence and transformation have customarily been viewed as adaptations to promote homologous recombination and genetic diversification. However, although transformation can clearly have long-term evolutionary benefits, natural selection acting on individual competent cells cannot foresee these, and the immediate selective advantages of DNA uptake for the cell are less clear (26, 53–56). Most transformation events are expected to be neutral or deleterious (the latter especially if using DNA from selectively killed cells [25]), and more immediate and reliable benefits of DNA uptake arise from DNA repair and nucleotide acquisition (57). A more extreme view is that DNA uptake could also be in part an unselected consequence of the adhesion and motility activities of T4P (58), though the coordinated regulation of the T4P genes responsible for DNA uptake and the non-T4P genes responsible for translocation into the cytoplasm argues against this. In this context, competence-induced cytoplasmic proteins coregulated with T4P genes take on special importance—are they modulations of the cellular response to their immediate environment or specific adaptations to promote transformation?

TABLE 1 Distribution of competence and related traits

Bacterial group(s)[a]	Comp[b]	r/m > 1[c]	T4P[d]	Rec2[e]	DprA[e]	Self[f]
Gammaproteobacteria	+	+/−	+	+	+	+/−
Betaproteobacteria	+	+/−	+	+	+	+
Alphaproteobacteria	+	+/−	−	+	+	ND
Epsilonproteobacteria	+	+	−[g]	+	+	+
Deltaproteobacteria	+	+	+	+	+	ND
Acidobacteria	−	ND	+	+	+	ND
Bacteroidetes/Chlorobi	+	+/−	−	+	+	ND
Spirochaetes	−	−	−	+	+	ND
Chlamydiae/Planctomycetes	−	−	−	+	+	ND
Firmicutes	+	+/−	+	+	+	−
Cyanobacteria	+	+	+	+	+	ND
Chloroflexi	−	ND	−	+	+	ND
Actinobacteria	+	+	−	+	+	−
Aquificae	−	ND	−	+	+	ND
Thermotogae	+	ND	−	+	+	ND
Deinococcus/Thermus	+	ND	+	+	+	−

[a] The bacterial group nomenclature follows Wu and Eisen (130).
[b] One or more species are known to be naturally competent (Comp) (+), or none are known to be naturally competent (−), following references 5 and 127.
[c] One or more species have an inferred recombination-to-mutation rate ratio (r/m) that is greater than 1 (+) or less than 1 (−), group has species with r/m <1 and >1 (+/−), or species in group were not evaluated (ND) (data from reference 66).
[d] T4P machinery is present (+) or is not present (−) for one or more species, as reported by Pelicic (128).
[e] Rec2 and DprA homologs were found in one or more species of all groups by searches of the NCBI protein database.
[f] One or more species show self-specific DNA uptake (+) or no self-specificity (−), or no species were tested (ND) (references in text).
[g] The epsilonproteobacterium *Helicobacter pylori* uses an alternative type IV secretion system for DNA uptake and has no apparent T4P apparatus (129).

THE PHYLOGENETIC DISTRIBUTION OF NATURAL COMPETENCE

Estimates of the distribution of natural competence come from three main approaches, each with specific limitations: (i) direct experimental assays, (ii) inferences from population genetics, and (iii) inferences from the presence of competence genes in sequenced genomes. All suggest that naturally competent species are widely distributed throughout the bacterial tree (Table 1) (5, 59).

Direct assays. Experimental demonstrations of natural competence are limited to only a few dozen species scattered across the

bacterial tree (5, 59); negative results are rarely reported (Table 1, column 2). Assays measuring genetic transformation are highly sensitive, but they can be done only in species where a selectable genetic marker is available (typically an antibiotic-resistance allele) and fail to discover competence in species where DNA uptake rarely leads to recombination. Fewer species have been directly tested for the ability to actively take up DNA; these assays typically use radiolabeled DNA and, although technically straightforward, are relatively insensitive (60).

A bigger problem for both assays is their dependence on prior induction of the postulated DNA uptake machinery; negative results are uninformative since a failure to observe uptake or transformation may simply mean that the appropriate inducing conditions have not been discovered. For example, *Vibrio cholerae* was not experimentally shown to be naturally competent until the inducing role of chitin breakdown products was discovered in 2005 (38). Another reason competence may often be missed is that levels of competence differ greatly even between isolates of the same species. For example, a survey of 34 *H. influenzae* isolates found 10^6-fold variation in the ability to be transformed, with many isolates being completely nontransformable (61, 62). Similar variation in transformability has been found in all other species that have been examined (see citations in reference 61). Thus, a species may be mistakenly thought to lack competence because only noncompetent isolates have been tested.

Inferences from population genetics. Less-direct methods can also be used to detect competence (Table 1, column 3). A number of studies have used population genetic models to estimate the historical frequency of homologous recombination from sequences of a standard set of seven housekeeping genes collected from many isolates of the same species (multilocus sequence typing [MLST]). One such study of pathogenic species found widely differing rates between taxonomic groups but consistently high rates among species known to be competent (63); this was confirmed for *Neisseria* species and *H. influenzae* by more recent studies (64, 65). A comprehensive survey of MLST data found high rates of historical recombination in many species not known to be competent (66), and in principle this might predict transformability.

Whole-genome approaches to detecting historical recombination are becoming increasingly practical as DNA sequencing costs decrease, and these reveal patterns of past recombination events that are not observable by MLST genotyping (67–69). However, these population-based approaches come with several caveats: a high historical recombination frequency may be due to horizontal gene transfer processes other than natural competence, and a competent species might show a low frequency due to scarcity of DNA from other strains, infrequent induction of competence, low recombination rates, or sampling of many nontransformable strains.

Inferences from gene distributions. In species with sequenced genomes, evidence of competence might also come from the presence of homologs of known competence genes (Table 1, columns 4 to 6). While investigations of this characteristic have met with some success, especially in relatives of well-studied species (33, 70), the results are not necessarily informative. The most striking counterexample is *E. coli*. Although it is not naturally transformable in the laboratory (despite extensive efforts) and its populations show little evidence of historical recombination, its genome contains an apparently complete set of competence genes inducible, like *H. influenzae*'s, by CRP and Sxy (71).

One reason the presence of competence gene homologs is not a reliable indicator of competence is that many genes required for DNA uptake have functions outside competence. In particular, the T4P machinery is widely distributed and strongly conserved at least partly because it also provides adhesion and twitching motility (8), and, conversely, the competent species *Helicobacter pylori* does not use T4P to take up DNA (72, 73). The taxonomic distributions of proteins that carry out homologous recombination are particularly uninformative, since these proteins have ubiquitous and conserved functions in DNA replication and repair (74).

Genes with the strongest cases for having functions specific to competence are those for Rec2/ComEC and DprA/Smf. Both proteins act after DNA has been transported across the outer membrane or through the cell wall. The Rec2/ComEC protein is required for translocation of single-stranded DNA into the cytoplasm; mutants are nontransformable. An alternative function for this protein has been recently identified in *Listeria*, in which ComEC (along with T4P genes) is important for infecting cells to escape host cell phagosomes (75). Competent cells defective in the cytoplasmic protein DprA take up and translocate DNA normally but produce few or no transformants; work in *S. pneumoniae* has shown that DprA/Smf is a recombination mediator protein, facilitating loading of RecA onto single-stranded DNA (76, 77). Although this suggests that DprA could function outside transformation, for example, by helping to stabilize stalled DNA replication forks, work in *E. coli* has failed to find a role in DNA repair (78). A GenBank search found 21,389 entries with *dprA* or *smf* homologs, 10,263 with *rec2* or *comEC* homologs, and 3,058 with homologs of both genes, numbers far higher than the 75 verified competent species listed in a recent survey (5). Although the ubiquity of these proteins may indicate that many or even most bacteria are competent, it could also reflect as yet poorly understood cellular functions in a manner independent of DNA uptake.

Summary. Although naturally competent bacteria are common across the bacterial phylogeny, their sporadic distribution both within and above the species level implies frequent losses and/or gains of competence, which obscure its evolutionary origins (66, 79). The pleiotropic functions of the T4P machinery suggest that it might be repeatedly co-opted to function in DNA uptake, with the genes themselves maintained by selection for adhesion and motility functions. Another possibility is that competence is an ancient ancestral trait that is frequently lost but that the long-term consequence of lacking a genetic recombination system is an evolutionary dead end that ultimately leads to extinction (80).

THE MECHANISM AND DIVERSITY OF DNA UPTAKE SPECIFICITY

Convergent evolution of uptake specificity. Self-specific DNA uptake was first discovered in two competent Gram-negative species, *H. influenzae* and *Neisseria gonorrhoeae*. Although most competent bacteria that have been tested appear to take up any available DNA (Table 1, column 7) (81–86), competent cells of these species take up genomic DNA from their own species dramatically better than DNA of *E. coli* or other distant relatives (87, 88). Subsequent work showed that this self-specificity is not due to a preference for "self" DNA *per se* but to the genome of each species

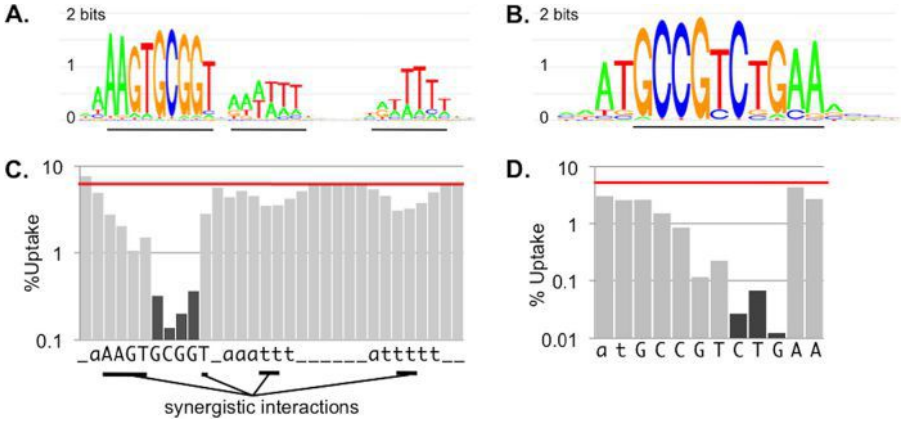


FIG 3 Effect of mismatches on uptake. (A and B) Logos of the *H. influenzae* and *N. meningitidis* genomic uptake sequences (*Hin*-USS and AT-DUS, respectively) derived by Gibbs recursive sampling of the complete genome sequences (adapted from *Genetics* [91] with permission of the publisher; permission conveyed through Copyright Clearance Center, Inc.). (C and D) Effects of mutations of the *Hin*-USS and AT-DUS consensus sequences on DNA uptake by *H. influenzae* and *N. meningitidis*, respectively. Consensus is shown on the *x* axis, with lowercase letters indicating positions with low information content and absence of a letter indicating no information content. Average percent DNA uptake of mutated constructs is shown on the *y* axis, with effects of mismatched core bases shown as dark gray bars. A red horizontal line indicates the average uptake of fragments carrying the consensus. Data were adapted from *Nucleic Acids Research* (112), by permission of Oxford University Press; also see Fig. 3 in reference 92. For panel C, the horizontal bars below the sequence indicate positions whose synergistic interactions increase their contributions to uptake. See references 91 and 92 for logos and consensus sequences of the alternative dialects in the members of the *Pasteurellaceae* and *Neisseriaceae*. Effects of mismatches on DUS uptake closely parallel those of ComP binding (93).

being highly enriched for a sequence motif preferred by its own uptake machinery.

Before genome sequences became available, sequencing of short preferentially taken-up fragments identified the *H. influenzae* uptake signal sequence (USS) as the 9-mer AAGTGCGGT, and the *N. gonorrhoeae* DNA uptake sequence (DUS) as the 10-mer GCCGTCTGAA, with initial estimates of each of at least several hundred occurrences per genome and an uptake bias of at least 100-fold (89, 90). Later sequencing of both genomes found about one occurrence per kb and revealed the more nuanced motif models shown in Fig. 3A and B (6, 91). Such motif models found that the USS has two helically phased flanking segments of A/T bases, a feature absent from the DUS.

Subsequent experimental work and genomic inference have found that uptake self-specificity is the norm throughout the *Pasteurellaceae* and *Neisseriaceae* families (80, 92, 93). Within the two families, the abundance of genomic uptake sequences suggests that many species are descended from recent competent ancestors even where their competence has not been demonstrated in laboratory cultures. This analysis cannot readily be extended beyond these families (using the presence of uptake-sequence-like repeats as indicators of competence), since motifs with properties similar to those of uptake sequences may occur for other reasons, and motifs whose properties are less dramatic may be difficult to detect without experimental data. The *Pasteurellaceae* and *Neisseriaceae* families are related only distantly, in the gamma and beta branches of the proteobacteria, respectively. Since the USS and DUS are distinct sequences and uptake sequences have not been identified in other groups, the self-specificities of the two families are likely to be of independent evolutionary origins. The only other bacterial species with demonstrated self-specificity are within the epsilonproteobacteria, *Campylobacter* and *Helicobacter*, but these have no strongly overrepresented genomic motif (94–96). The many similarities between the USS and DUS described below suggest that they are convergent products of the same evolutionary forces.

The genomic distribution of uptake sequences. In both *H. influenzae* and *Neisseria meningitidis* (and its close relative *N. gonorrhoeae*), the density of uptake sequences is about 1/kb—more than 100-fold higher than expected by chance—and they collectively comprise ~1% of each genome (6). Both the USS and DUS are enriched in the relatively permissive regions of the genome (intergenic regions, poorly conserved genes, and poorly conserved parts of genes), presumably to minimize interference with protein-coding functions (97). In addition, both the USS and DUS in intergenic regions frequently occur as inverted repeats that act as *rho*-independent transcriptional terminators (90, 98). Although DUS were reported to be enriched in "genome maintenance" genes, reanalysis of the data found no significant correlation (97, 99).

Uptake sequences are not insertions arising from a copy-paste or other duplicative mechanism. Instead, alignment of *H. influenzae* USS-containing genes with homologs from relatives without USSs shows that they have evolved mainly by simple point mutations (80); DUSs in coding sequences show the same pattern (100). Once they have arisen at specific positions, uptake sequences are stable; many USSs are in homologous positions in *H. influenzae* and its relative *Pasteurella multocida* despite the hundreds of millions of years since their divergence (101). Comparison of DUS locations in three closely related species of *Neisseria* also found strong conservation (100); more divergent genomes are now available for analysis (68).

Genes that have undergone horizontal gene transfer provide additional evidence of how uptake sequences accumulate. When genome sequences of multiple isolates of *H. influenzae* and of *Neisseria* species are compared, uptake sequences have a higher density in genes present in all isolates (core genes) than in genes found in only some isolates (accessory genes) (100, 102). This would be expected if uptake sequences were to accumulate slowly, since recently acquired accessory genes would not yet have accumulated them. Slow accumulation is also consistent with the absence of the USS from the Mu-like prophage recently acquired by

the Rd strain of *H. influenzae* (103). Independent estimates of the timing of accessory gene acquisition could permit accurate estimates of the rate of uptake sequence accumulation, which would in turn aid in interpreting other phenomena such as the high frequency of USSs in the genome of *H. influenzae* phage HP1 (90).

Uptake sequences have also changed the frequencies of specific peptide signatures in proteomes. A total of 70% of *H. influenzae* USSs and 35% of N. meningitidis DUSs are in open reading frames, where they encode specific tripeptides. Analysis of tripeptide frequencies reveals that these occur twice as often in the proteomes of their own species as in the proteomes of species lacking that uptake sequence (e.g., *E. coli*). This pattern can be explained by the combined effects on mutational divergence of selection for their coding function and of molecular drive arising from their effects on DNA uptake. Because many, often most, of these tripeptides are encoded by sequences that no longer match the uptake sequence consensus, they provide stark historical evidence of the power of biased DNA uptake to affect genome evolution and of the strong similarities of these evolutionary forces in the two families (97).

Mechanism of uptake specificity. Since only DNA fragments with uptake sequences become protected from experimentally added nucleases, uptake sequences must determine whether DNAs are transported across the outer membrane (104, 105). And because a single USS or DUS is sufficient to mobilize uptake of both short and long DNA fragments (106, 107), uptake specificity likely acts only at the initiation step, not continuously, during uptake of long molecules. The *H. influenzae* protein or proteins responsible for USS specificity have still not been identified, but the *N. meningitidis* ComP minor pilus protein has recently been shown to specifically bind DUSs (93, 108). *H. influenzae* lacks a ComP ortholog, but this role might instead be filled by a minor pilin encoded by the competence-regulated *pilNOPQ* operon, which is conserved across the *Pasteurellaceae* and *Enterobacteriaceae* (33).

We propose that strong binding of the uptake machinery to DNA is likely needed to overcome the physical difficulties posed by transport of dsDNA across the outer membrane pore, an obstacle not faced by Gram-positive bacteria. One difficulty is the 6-nm diameter of the secretin pore. Initiation of uptake from internal sites requires that DNA be sharply kinked back on itself at the point of initiation to pass through this pore, which is made difficult by DNA's 50-nm persistence length. The secretin pore is also just wide enough to accommodate either the type 4 pilus (109) or the pair of DNA double helices on both sides of the initiation site, but not both, suggesting that the tip of the pilus/pseudopilus may lead the DNA through the pore. After initiation, tight DNA-protein interactions may also be required to transmit the extremely strong (up to 40 pN) DNA uptake forces that have been recorded during pilus retraction (7, 72, 110); we propose that such strong interactions are unlikely to be achieved without at least some degree of sequence specificity. In a later section, we discuss the implications of this for the many species that do not demonstrate self-specific uptake.

Recent work has found that only 3 to 4 "core" bases within the USS and DUS make strong contributions to DNA uptake, with the remaining bases in each uptake motif individually making smaller contributions to uptake specificity (91–93, 111, 112). In *H. influenzae*, changing the consensus USS within a DNA fragment dramatically reduces uptake if the change affects one of the four bases GCGG, while mutations at other USS positions have substantially smaller effects (Fig. 3C). Similarly, in *N. meningitidis* the 3 bases CTG dramatically reduce DNA uptake, while other bases make smaller individual contributions (Fig. 3D). These data suggest that the core bases of the USS and DUS make strong sequence-specific contacts that tightly bind DNA during the retraction that initiates uptake. Initial assays of DUS binding by ComP did not distinguish between core and noncore binding, since they compared a perfect 12-bp DUS to a control construct mismatched at every other base (108), but later assays found that mutations to the core had much larger impacts on binding than mutations to noncore bases, in a pattern closely matching the *in vivo* DNA uptake results shown in Fig. 3D (92, 93).

These data also raise the issue of why noncore bases show such strong signatures in the accumulated genomic uptake sequence motif, despite making apparently minor contributions to uptake (Fig. 3). A detailed characterization of the *H. influenzae* bias was accomplished using deep sequencing of DNA fragments containing a degenerate USS; a pool of DNA fragments were recovered from competent cells' periplasms, and $\sim 10^7$ fragments of both this periplasmic pool and the input pool were sequenced and compared. This corroborated the previous analysis of singly mutated USS constructs, but because of the high degeneracy and high sequence yield, that study was also able to show the effects of interactions ("positional dependencies") between noncore bases, with doubly mutated constructs having on average ~5-fold-less-efficient uptake than that predicted from singly mutated constructs (indicated by the "synergistic interactions" shown at the bottom of Fig. 3C) (112). Thus, noncore bases still make important contributions to DNA uptake; although their individual effects are small, their collective effects are substantial. Similar interaction effects are suggested by the decreased uptake seen for some multiply mismatched DUS variants (92), but no systematic survey has yet been undertaken. Such interaction effects could arise because noncore bases weakly but cooperatively bind the uptake machinery, but they could also arise from cooperative contributions to DNA bending or kinking.

Subclade-specific uptake sequence dialects. Uptake sequences have often been assumed to be species-specific mate recognition signals, but several lines of evidence show both sharing and divergence of uptake sequences across related species. Analysis of heterospecific transformation in *H. influenzae* found that DNAs from many other *Haemophilus* species are readily taken up although they are otherwise too divergent to produce recombinants (113). Sharing of the same uptake sequence across diverged species can also promote harmful uptake of foreign DNA. In *H. influenzae*, this has been shown to kill cells by inducing the SOS response, which activates a resident prophage (114, 115).

Later work analyzing eight *Pasteurellaceae* genomes identified two monophyletic subclades with divergent USS types (dialects), *Hin*-USS and *Apl*-USS; Fig. 4 shows the number of each USS type in the genome of each species (80). The two USS dialects share the GCGG core sequence and the helically phase T tracts, but several other bases do not align and the *Apl*-USS extends beyond the second T tract. The densities of the 9-mer consensus *Hin*-USS and *Apl*-USS within each subclade range from 117 to 836 per Mb, and bacteria in each subclade exhibit a strong preference for their USS dialect. More recent work in the *Neisseriaceae* has identified eight distinct DUS dialects, with each dialect highly enriched in the genomes of subclades within the *Neisseriaceae* phylogenetic tree

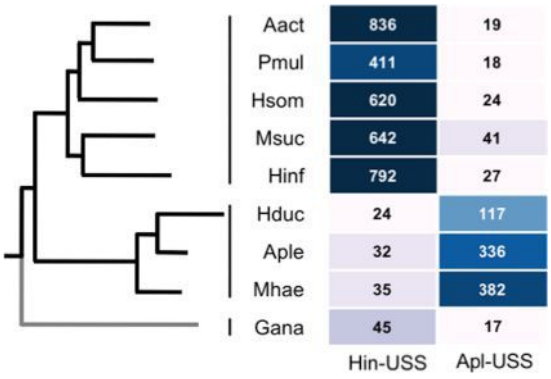


FIG 4 Clade-specific USS variants in 9 *Pasteurellaceae* species. Each cell shows the density of that USS type per Mb, with corresponding color-coding on a log scale. The phylogenetic tree to the left was created based on data presented in reference 80, with an added *G. anatis* branch based on data presented in references 117 and 118. Taxa listed from top to bottom are as follows: *Aggregatibacter actinomycetemcomitans* D11S-1, *Pasteurella multocida* Pm70, "*Haemophilus somnus*" 129PT, *Mannheimia succiniciproducens* MBEL55E, *H. influenzae* 86-028NP, *H. ducreyi* 35000HP, *Actinobacillus pleuropneumoniae* serovar 5b sp. strain L20, *Mannheimia haemolytica* USDA-ARS-SAM-185, and *Gallibacterium anatis* UMN179.

(Fig. 5) (92, 93). As previously seen for USS dialects, the densities of 12-mer consensus DUS of different dialects differed dramatically, ranging from 705 to 1,754 per Mb, and measurements of uptake biases and transformation frequencies for several species also indicate preferential uptake of the corresponding genomic consensus.

Variation between dialects in both families is found outside the core bases that are particularly crucial for DNA uptake; i.e., the core DUS bases CTG, like the USS core bases GCGG, are invariant between dialects. This suggests a remarkable similarity between the families in both the mechanism of DNA uptake and the coevolutionary constraints operating between uptake specificity and genomic accumulation of uptake sequences. We reason that changes to uptake specificity that affect binding to a crucial core base are highly constrained, since these would also drastically reduce DNA uptake. Changes in uptake specificity affecting binding to a noncore base would have comparatively weaker effects on DNA uptake, allowing for less-constrained coevolutionary changes of the specificity machinery and the uptake sequences accumulated in the genome. However, the strong specificities of the uptake machineries for the different dialects and the deep-

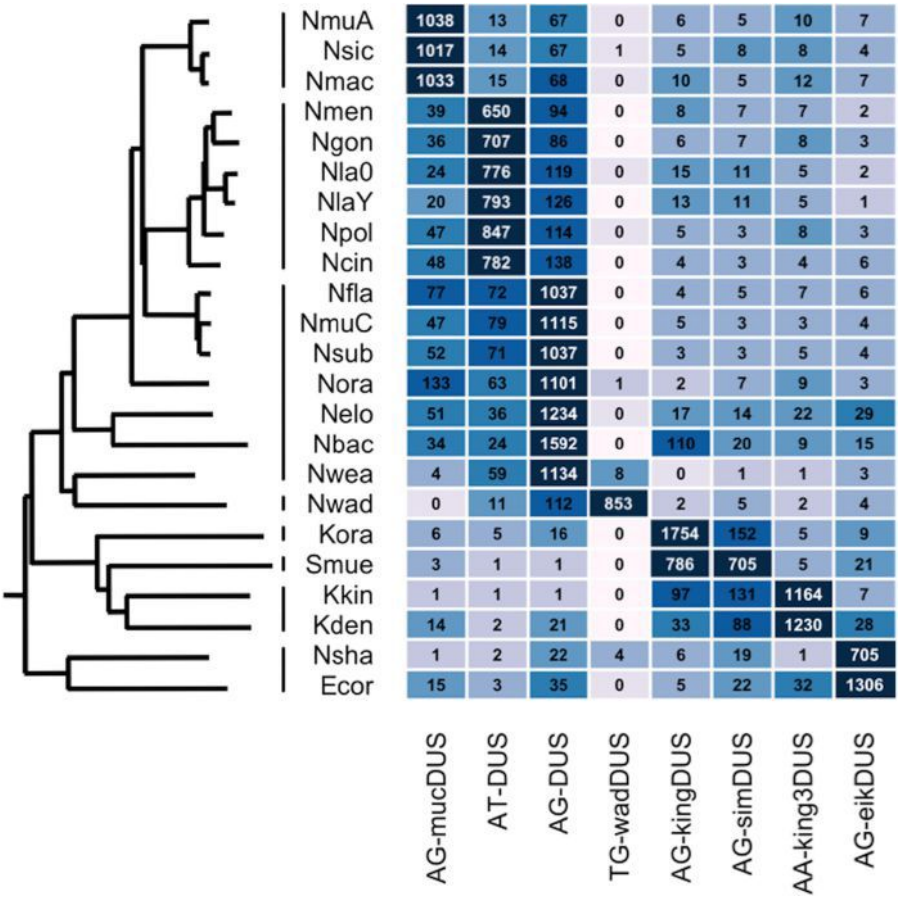


FIG 5 Clade-specific uptake sequence dialects in 23 *Neisseriaceae* species. Each cell shows the density of that DUS type per Mb, with corresponding color coding on a log scale. The sequences of the eight 12-mer *Neisseriaceae* dialects and the phylogenetic tree were created based on data presented in Fig. 1 and 2 of reference 92. Taxa listed from top to bottom are follows: *Neisseria mucosa* ATCC 25996, *N. sicca* ATCC 29256, *N. macacae* ATCC 33926, *N. meningitidis* MC58, *N. gonorrhoeae* FA1090, *N. lactamica* 020-06, *N. lactamica* Y92-1009, *N. polysaccharea* ATCC 43768, *N. cinerea* ATCC 14685, *N. flavescens* SK114, *N. mucosa* C102, *N. subflava* NJ9703, *N.* oral taxon 014-F0314 (*Neisseria*-like isolate taken from an oral microbiome), *N. elongata* subsp. *glycolytica* ATCC 29315, *N. bacilliformis* ATCC-BAA120, *N. weaveri* ATCC 51223, *N. wadsworthii* 971, *Kingella oralis* ATCC 51147, *Simonsiella muelleri* ATCC 29453, *Kingella kingae* ATCC 23330, *Kingella denitrificans* ATCC 33394, *N. shayeganii* 871, and *Eikenella corrodens* ATCC 23834.

sequencing results from *H. influenzae* suggest that sequence-specific DNA-protein interactions are still important at noncore positions but that several different combinations of bases and protein residues can provide for these.

Exceptional species within each family may be particularly useful for better understanding of the coevolution of uptake specificity and uptake sequence accumulation: analysis of the *Pasteurellaceae* species *Gallibacterium anatis* found self-specificity, but the densities of *Hin*-USS and *Apl*-USS are substantially lower than for other species in the family (Fig. 4) (73, 116). *G. anatis* represents an outgroup of the two main subclades (117, 118), suggesting that this species would be an ideal system for investigation of the origins of uptake bias and uptake sequence accumulation. In contrast, *Haemophilus ducreyi* (*Apl*-USS subclade) has a substantially lower density of USS than other *Pasteurellaceae*, suggesting either that it has lost competence or that its uptake specificity is changing (Fig. 4). Notably, one of the *Neisseriaceae* genomes analyzed, *Simonsiella muelleri*, contained evidence of two distinct dialects, both its own AG-simDUS and the AG-kingDUS found in related *Kingella oralis*, in roughly equal proportions (Fig. 5) (92). It remains unclear whether this is the result of two distinct uptake biases in the same organism or whether one of the DUS dialects is a remnant from a time before the uptake bias changed. Further experimental and genomic study of these species is clearly in order.

A MODEL FOR THE EVOLUTION OF SELF-SPECIFICITY

Molecular drive can explain the accumulation of uptake sequences in genomes. A satisfactory explanation of self-specificity must account for both the sequence bias of the DNA uptake machinery and the accumulation of its preferred sequences in the genome. The commonly assumed mate recognition function requires that both components evolve simultaneously by natural selection operating on their combined effects, but relaxing this assumption greatly simplifies the evolutionary steps. Here we first consider how—regardless of whether or not recombination is a selected function of competence—uptake bias causes a "molecular drive" leading to accumulation of uptake sequences in genomes, in the same way that biased gene conversion causes a molecular drive leading to allele fixation in sexual eukaryotes (91, 101, 119). We then consider how weak uptake biases could be amplified over time.

Consider an ancestral species that is competent but has few or no uptake sequences. This species is likely to have had some uptake bias, for reasons discussed above and below. Provided some fraction of the available DNA is conspecific (or otherwise able to recombine with the chromosome), the combination of random mutation, biased DNA uptake, and subsequent recombination will enrich the genomes of this species' descendants with the preferred sequences, as follows (Fig. 6): at any position in the genome whose sequence is not already well matched to the uptake bias, random mutation will sometimes create variants that better fit this bias, and cell death will create a pool of environmental DNA that includes this variation. Competent cells will then preferentially take up those DNA variants that better match their bias, and recombination of these with their chromosomal homologs will transfer these preferences into the genome (91). The preferred sequences need not be functionally beneficial in any way; their accumulation will inevitably continue until it is limited by mutational degradation of uptake sequences and by selection against uptake sequences that conflict with genetic functions. In simulations, the mutation and transformation rates determine the final density of uptake sequences, the uptake bias determines their sequence distribution, the sizes of available DNA fragments determine their spacing, and natural selection determines their relationships to protein coding and other genome functions (91). Molecular drive is also predicted to prevent accumulation of uptake-reducing mutations once strong uptake sequences have arisen.

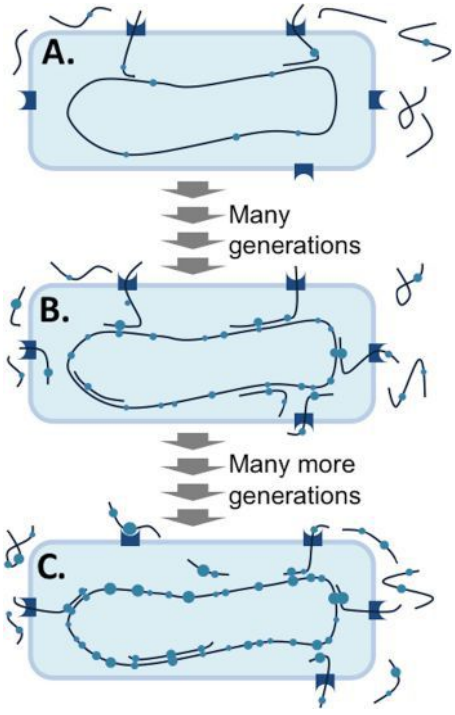


FIG 6 Evolution of uptake sequences by molecular drive (adapted from *Genetics* [91] with permission of the publisher; permission conveyed through Copyright Clearance Center, Inc.). Large circles, strongly preferred sequences; small circles, weakly preferred sequences. (A) The ancestral cell has biased DNA uptake machinery, but preferred sequences occur only at the locations of random expectation. (B) After many generations of biased uptake, recombination has enriched the genome for moderately preferred sequences, creating moderate self-specificity. (C) The increased availability of the preferred DNA leads to further enrichment of increasingly preferred sequences and stronger self-specificity.

Notably, the molecular drive model does not preclude the possibility that competence is in part maintained for its genetic/informational consequences (26, 53–56); rather, it stipulates only that the existence of self-specificity does not necessarily mean that uptake sequences are there to screen for homologous DNA. Consistent with this, the only known specificity factor, ComP from *Neisseria*, is not an "add-on" filter for DUS-containing DNA but is an intrinsic part of the uptake mechanism, since knockouts have strong defects in DNA uptake (92, 93).

Causes of uptake biases. The evolutionary model described above begs the question of why DNA uptake would have a preexisting sequence bias, but uptake biases are likely to be more common than is suggested by the known occurrences of self-specificity. Sequence biases are typical of DNA-binding proteins in general, even those whose function is sequence independent, with tighter binding typically associated with stronger bias (120, 121).

As discussed above, the strong forces and topological deformations needed for DNA uptake, especially by Gram-negative bacteria, require that DNA be tightly bound to the uptake machinery. Thus, organisms with no reported self-specificity are predicted to nevertheless exhibit some degree of sequence bias in a manner independent of any enrichment in the genome.

Development of self-specificity. If the above conditions are met, the genome will gradually and inevitably become enriched for sequences preferred by the DNA uptake machinery. This enrichment then sets the stage for selection on the genes encoding this machinery, amplifying the bias to promote more-efficient uptake or better exclusion of harmful foreign DNAs. Such selection could occur regardless of whether the main benefit of the presence of the DNA is its nutrient content or its genetic information.

Why then do most tested groups have no self-specificity? Two factors are important here—whether preferred sequences recombine often enough to create a signature in the genome and whether this signature is strong enough to cause self-specific uptake. If conspecific DNA is rare in the local microenvironment, or if incoming DNA rarely recombines with the chromosome, even strong uptake biases would not give rise to a genomic signature. And even a strong signature does not cause strong self-specificity when it also occurs frequently in other DNAs; e.g., if cells had an absolute requirement for a simple motif such as AGTC, they would still take up many foreign DNAs as well as their own.

Is there something special about *Pasteurellaceae* and *Neisseriaceae* that predisposed them to evolve strong uptake biases and accumulate genomic uptake sequences? Species in both families predominantly live in respiratory tracts and other mucosal environments rich in host DNA, but other competent Gram-negative species lacking uptake sequences coexist with these (see, for example, reference 122).

It has been suggested that beneficial genetic recombination is optimized by quorum-sensing regulation as well as by self-specific uptake, since quorum sensing would ensure that cells become competent only when surrounded by other members of their species (4). However, the links between quorum sensing and competence are often incomplete or indirect. For example, only a subset of *V. cholerae*'s competence genes are regulated by a secreted autoinducer (39), and in *S. pneumoniae* and *Bacillus subtilis*, competence is part of a much larger set of genes and processes influenced by secreted autoinducers (39, 50, 51). Interpretation of this regulation is further complicated by uncertainties about whether secreted autoinducers exist primarily for sensing cell densities or for sensing the physical properties of the cell's microenvironment (123). Further comparative studies into factors that limit DNA uptake to close relatives, such as uptake specificity and quorum sensing, are warranted.

OUTSTANDING ISSUES

The finding that subclades of the *Pasteurellaeceae* and *Neisseriaceae* have distinct dialects is consistent with mutational divergence of specificity components of the uptake machinery as the subclades diverged. The resulting changes in uptake bias would then lead to corresponding changes in the population of uptake sequences in the genome by the same combination of random mutation, biased uptake, and recombination that caused their original accumulation and maintenance. Within the *Hin*-USS subclade, the positions of uptake sequence have remained quite stable (101), but their stability over the deeper evolutionary time separating the two subclades is unknown. A more systematic examination of synteny between uptake sequences of distinct dialects in both families is clearly in order.

The divergence of uptake specificity within each family could also give insights into the molecular basis of uptake specificity. Do changes in specific amino acid residues in components of the uptake machinery cooccur with changes in specificity? For example, if ComP interacts with dialect-specific bases in the DUS, mutations at specific surface residues in its positively charged channel could be responsible for shifts in specificity and the subsequent turnover of DUS dialects observed in the different *Neisseriaceae* species. Recent results show that single mutations to noncore bases in DUS have only weak effects on ComP binding (93), suggesting either that ComP does not directly interact with these positions or that effects are detected only when multiple noncore bases are changed. Experiments with ComP protein purified from species with different dialects will help sort this out. A comparative approach could also potentially also identify and validate candidates for the still unknown genes responsible for uptake specificity in the *Pasteurellaceae*.

The drive model predicts that the uptake sequences in modern genomes will have lower divergence rates than the parts of the genome that are under otherwise similar constraints, a pattern we have observed in preliminary studies of *H. influenzae* genomes (J. C. Mell and R. J. Redfield, unpublished data). This pattern could be informative about the function of competence, since molecular drive predicts reduced variation only at uptake sequences whereas selection for a mate-choice function predicts that reduced variation would also affect flanking sequences containing beneficial alleles. Since only a subset of genes in the USS-poor species *H. ducreyi* and *G. anatis* have nearby uptake sequences, similar population genetic analysis results at these loci could reveal biases in recombination to these loci. Finally, since species with particular dialects often contain many uptake sequences matching other dialects (Fig. 4 and 5), cooccurrence could cause genes close to these to exhibit a higher probability of cross-clade recombination. These and other population-based studies are becoming increasingly feasible as the number of species with many sequenced isolates increases.

We argue above that uptake biases are expected to be the norm rather than the exception among competent bacteria, which predicts that species with no reported self-specificity nevertheless would still have some degree of uptake bias. Although identifying weak biases was previously impractical, this can now be done using a modification of the deep-sequencing approach used to characterize the uptake bias of *H. influenzae* (112), in which periplasmic DNA taken up by competent cells was purified away from their chromosomes. The best candidates for these tests are the competent Gram-negative species demonstrated to lack self-specificity, *Acinetobacter baylyi*, *Thermus thermophilus*, *Vibrio cholerae*, and *Pseudomonas stutzeri* (82–84), but the well-studied competence models in the Gram-positive bacteria *Bacillus subtilis* and *Streptococcus pneumoniae* should also be examined (85, 86). Finding that uptake biases are still present even in the absence of self-specificity would confirm that intrinsic sequence constraints play important roles in DNA uptake.

The epsilonproteobacteria *Helicobacter pylori* and *Campylobacter jejuni* have been shown to have uptake specificity, but the mechanism is unclear since no uptake sequences have been identified in their genomes (95, 96). Although this may simply reflect

lower enrichment or less-specific biases that are not as readily detected as those of the *Pasteurellaceae* and *Neisseriaceae*, these species could also require specific DNA modifications for efficient uptake. Glucosylation of DNA has been shown to affect DNA binding by *B. subtilis* (124), and ethylation of specific USS positions can either reduce or enhance their uptake (89). Since *Helicobacter* and *Campylobacter* use a type IV secretion system for DNA uptake instead of T4P (114), insights into uptake specificity in this independently evolved system will provide powerful evidence of the forces responsible for uptake specificity.

That there is a lack of consensus on the evolutionary function of natural competence is not because we lack evidence of how natural selection acts on genetic recombination. Many microbiologists are probably unaware of the extensive body of theoretical and experimental work on the evolution of genes causing sexual (meiotic) recombination in eukaryotes (reviewed in reference 125). This work has shown that such genes and alleles are not favored by natural selection except in restricted circumstances, and even in these circumstances selection for recombination is readily swamped by molecular drive forces (126). Lack of selection for recombination is not inconsistent with the strongly beneficial recombination events observed in bacterial genomes, since rare beneficial genotypes will be preserved while harmful recombinants are swept under the rug of evolutionary history.

ACKNOWLEDGMENTS

This work was supported by grants to R.J.R. from the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council of Canada and by a postdoctoral fellowship to J.C.M. from the National Institutes of Health.

We thank three anonymous reviewers for their detailed and helpful comments.

REFERENCES

1. **Chen I, Christie PJ, Dubnau D.** 2005. The ins and outs of DNA transfer in bacteria. Science **310:**1456–1460. http://dx.doi.org/10.1126/science.1114021.
2. **Krüger NJ, Stingl K.** 2011. Two steps away from novelty—principles of bacterial DNA uptake. Mol. Microbiol. **80:**860–867. http://dx.doi.org/10.1111/j.1365-2958.2011.07647.x.
3. **Chen I, Dubnau D.** 2004. DNA uptake during bacterial transformation. Nat. Rev. Microbiol. **2:**241–249. http://dx.doi.org/10.1038/nrmicro844.
4. **Seitz P, Blokesch M.** 2013. Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiol. Rev. **37:**336–363. http://dx.doi.org/10.1111/j.1574-6976.2012.00353.x.
5. **Johnsborg O, Eldholm V, Håvarstein LS.** 2007. Natural genetic transformation: prevalence, mechanisms and function. Res. Microbiol. **158:**767–778. http://dx.doi.org/10.1016/j.resmic.2007.09.004.
6. **Smith HO, Gwinn ML, Salzberg SL.** 1999. DNA uptake signal sequences in naturally transformable bacteria. Res. Microbiol. **150:**603–616. http://dx.doi.org/10.1016/S0923-2508(99)00130-8.
7. **Clausen M, Jakovljevic V, Søgaard-Andersen L, Maier B.** 2009. High-force generation is a conserved property of type IV pilus systems. J. Bacteriol. **191:**4633–4638. http://dx.doi.org/10.1128/JB.00396-09.
8. **Burrows LL.** 2005. Weapons of mass retraction. Mol. Microbiol. **57:**878–888. http://dx.doi.org/10.1111/j.1365-2958.2005.04703.x.
9. **Craig L, Li J.** 2008. Type IV pili: paradoxes in form and function. Curr. Opin. Struct. Biol. **18:**267–277. http://dx.doi.org/10.1016/j.sbi.2007.12.009.
10. **Laurenceau R, Péhau-Arnaudet G, Baconnais S, Gault J, Malosse C, Dujeancourt A, Campo N, Chamot-Rooke J, Le Cam E, Claverys J-P, Fronzes R.** 2013. A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. PLoS Pathog. **9:**e1003473. http://dx.doi.org/10.1371/journal.ppat.1003473.
11. **Kahn ME, Maul G, Goodgal SH.** 1982. Possible mechanism for donor DNA binding and transport in *Haemophilus*. Proc. Natl. Acad. Sci. U. S. A. **79:**6370–6374. http://dx.doi.org/10.1073/pnas.79.20.6370.
12. **Bulyha I, Schmidt C, Lenz P, Jakovljevic V, Höne A, Maier B, Hoppert M, Søgaard-Andersen L.** 2009. Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. Mol. Microbiol. **74:**691–706. http://dx.doi.org/10.1111/j.1365-2958.2009.06891.x.
13. **Peabody CR, Chung YJ, Yen MR, Vidal-Ingigliardi D, Pugsley AP, Saier MH, Jr.** 2003. Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. Microbiology **149**(Pt 11)**:**3051–3072. http://dx.doi.org/10.1099/mic.0.26364-0.
14. **Mell JC, Shumilina S, Hall IM, Redfield RJ.** 2011. Transformation of natural genetic variation into *Haemophilus influenzae* genomes. PLoS Pathog. **7:**e1002151. http://dx.doi.org/10.1371/journal.ppat.1002151.
15. **Goodgal SH.** 1982. DNA uptake in *Haemophilus* transformation. Annu. Rev. Genet. **16:**169–192. http://dx.doi.org/10.1146/annurev.ge.16.120182.001125.
16. **Hamilton HL, Dillard JP.** 2006. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. Mol. Microbiol. **59:**376–385. http://dx.doi.org/10.1111/j.1365-2958.2005.04964.x.
17. **Lacks S, Neuberger M.** 1975. Membrane location of a deoxyribonuclease implicated in the genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **124:**1321–1329.
18. **Neuhard J, Nygaard P.** 1987. Pyrimidines and purines, p 445–473. *In* Neidhardt FC (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 1st ed. American Society for Microbiology, Washington, DC.
19. **Pietramellara G, Ascher J, Borgogni F, Ceccherini MT, Guerri G, Nannipieri P.** 2008. Extracellular DNA in soil and sediment: fate and ecological relevance. Biol. Fertil. Soils **45:**219–235. http://dx.doi.org/10.1007/s00374-008-0345-8.
20. **Barouki R, Smith HO.** 1985. Reexamination of phenotypic defects in rec-1 and rec-2 mutants of *Haemophilus influenzae* Rd. J. Bacteriol. **163:**629–634.
21. **Redfield RJ.** 1993. Genes for breakfast: the have-your-cake-and-eat-it-too of bacterial transformation. J. Hered. **84:**400–404.
22. **Michod RE, Bernstein H, Nedelcu AM.** 2008. Adaptive value of sex in microbial pathogens. Infect. Genet. Evol. **8:**267–285. http://dx.doi.org/10.1016/j.meegid.2008.01.002.
23. **Engelmoer DJP, Donaldson I, Rozen DE.** 2013. Conservative sex and the benefits of transformation in *Streptococcus pneumoniae*. PLoS Pathog. **9:**e1003758. http://dx.doi.org/10.1371/journal.ppat.1003758.
24. **Redfield RJ, Schrag MR, Dean AM.** 1997. The evolution of bacterial transformation: sex with poor relations. Genetics **146:**27–38.
25. **Redfield RJ.** 1988. Evolution of bacterial transformation: is sex with dead cells ever better than no sex at all? Genetics **119:**213–221.
26. **Moradigaravand D, Engelstädter J.** 2013. The evolution of natural competence: disentangling costs and benefits of sex in bacteria. Am. Nat. **182:**E112–E126. http://dx.doi.org/10.1086/671909.
27. **de Vries J, Wackernagel W.** 2002. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. Proc. Natl. Acad. Sci. U. S. A. **99:**2094–2099. http://dx.doi.org/10.1073/pnas.042263399.
28. **Hülter N, Wackernagel W.** 2008. Double illegitimate recombination events integrate DNA segments through two different mechanisms during natural transformation of *Acinetobacter baylyi*. Mol. Microbiol. **67:**984–995. http://dx.doi.org/10.1111/j.1365-2958.2007.06096.x.
29. **Lin EA, Zhang XS, Levine SM, Gill SR, Falush D, Blaser MJ.** 2009. Natural transformation of *Helicobacter pylori* involves the integration of short DNA fragments interrupted by gaps of variable size. PLoS Pathog. **5:**e1000337. http://dx.doi.org/10.1371/journal.ppat.1000337.
30. **Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA.** 2012. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. Cell Host Microbe **12:**177–186. http://dx.doi.org/10.1016/j.chom.2012.06.003.
31. **Sparling PF.** 1966. Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. J. Bacteriol. **92:**1364–1371.
32. **Cameron ADS, Volar M, Bannister LA, Redfield RJ.** 2008. RNA secondary structure regulates the translation of *sxy* and competence development in *Haemophilus influenzae*. Nucleic Acids Res. **36:**10–20. http://dx.doi.org/10.1093/nar/gkm915.
33. **Cameron ADS, Redfield RJ.** 2006. Non-canonical CRP sites control competence regulons in *Escherichia coli* and many other gamma-

proteobacteria. Nucleic Acids Res. **34:**6001–6014. http://dx.doi.org/10.1093/nar/gkl734.
34. **Redfield RJ, Cameron ADS, Qian Q, Hinds J, Ali TR, Kroll JS, Langford PR.** 2005. A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. J. Mol. Biol. **347:**735–747. http://dx.doi.org/10.1016/j.jmb.2005.01.012.
35. **MacFadyen LP, Chen D, Vo HC, Liao D, Sinotte R, Redfield RJ.** 2001. Competence development by *Haemophilus influenzae* is regulated by the availability of nucleic acid precursors. Mol. Microbiol. **40:**700–707. http://dx.doi.org/10.1046/j.1365-2958.2001.02419.x.
36. **Sinha S, Mell J, Redfield R.** 2013. The availability of purine nucleotides regulates natural competence by controlling translation of the competence activator Sxy. Mol. Microbiol. **88:**1106–1119. http://dx.doi.org/10.1111/mmi.12245.
37. **Sinha S, Mell JC, Redfield RJ.** 2012. Seventeen Sxy-dependent cyclic AMP receptor protein site-regulated genes are needed for natural transformation in *Haemophilus influenzae*. J. Bacteriol. **194:**5245–5254. http://dx.doi.org/10.1128/JB.00671-12.
38. **Meibom KL, Blokesch M, Dolganov NA, Wu C-Y, Schoolnik GK.** 2005. Chitin induces natural competence in *Vibrio cholerae*. Science **310:**1824–1827. http://dx.doi.org/10.1126/science.1120096.
39. **Lo Scrudato M, Blokesch M.** 2012. The regulatory network of natural competence and transformation of *Vibrio cholerae*. PLoS Genet. **8:**e1002778. http://dx.doi.org/10.1371/journal.pgen.1002778.
40. **Yamamoto S, Morita M, Izumiya H, Watanabe H.** 2010. Chitin disaccharide (GlcNAc)2 induces natural competence in *Vibrio cholerae* through transcriptional and translational activation of a positive regulatory gene *tfoXVC*. Gene **457:**42–49. http://dx.doi.org/10.1016/j.gene.2010.03.003.
41. **Yamamoto S, Izumiya H, Mitobe J, Morita M, Arakawa E, Ohnishi M, Watanabe H.** 2011. Identification of a chitin-induced small RNA that regulates translation of the *tfoX* gene, encoding a positive regulator of natural competence in *Vibrio cholerae*. J. Bacteriol. **193:**1953–1965. http://dx.doi.org/10.1128/JB.01340-10.
42. **Sun Y, Bernardy EE, Hammer BK, Miyashiro T.** 2013. Competence and natural transformation in vibrios. Mol. Microbiol. **89:**583–595. http://dx.doi.org/10.1111/mmi.12307.
43. **Palchevskiy V, Finkel SE.** 2006. *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. J. Bacteriol. **188:**3902–3910. http://dx.doi.org/10.1128/JB.01974-05.
44. **Sinha S, Redfield RJ.** 2012. Natural DNA uptake by *Escherichia coli*. PLoS One **7:**e35620. http://dx.doi.org/10.1371/journal.pone.0035620.
45. **Vegge CS, Brøndsted L, Ligowska-Marzêta M, Ingmer H.** 2012. Natural transformation of *Campylobacter jejuni* occurs beyond limits of growth. PLoS One **7:**e45467. http://dx.doi.org/10.1371/journal.pone.0045467.
46. **Claverys J-P, Prudhomme M, Martin B.** 2006. Induction of competence regulons as a general response to stress in gram-positive bacteria. Annu. Rev. Microbiol. **60:**451–475. http://dx.doi.org/10.1146/annurev.micro.60.080805.142139.
47. **Martin B, Quentin Y, Fichant G, Claverys J-P.** 2006. Independent evolution of competence regulatory cascades in streptococci? Trends Microbiol. **14:**339–345. http://dx.doi.org/10.1016/j.tim.2006.06.007.
48. **Solomon JM, Grossman AD.** 1996. Who's competent and when: regulation of natural genetic competence in bacteria. Trends Genet. **12:**150–155. http://dx.doi.org/10.1016/0168-9525(96)10014-7.
49. **Johnsborg O, Håvarstein LS.** 2009. Regulation of natural genetic transformation and acquisition of transforming DNA in *Streptococcus pneumoniae*. FEMS Microbiol. Rev. **33:**627–642. http://dx.doi.org/10.1111/j.1574-6976.2009.00167.x.
50. **Berka RM, Hahn J, Albano M, Draskovic I, Persuh M, Cui X, Sloma A, Widner W, Dubnau D.** 2002. Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. Mol. Microbiol. **43:**1331–1345. http://dx.doi.org/10.1046/j.1365-2958.2002.02833.x.
51. **Peterson SN, Sung CK, Cline R, Desai BV, Snesrud EC, Luo P, Walling J, Li H, Mintz M, Tsegaye G, Burr PC, Do Y, Ahn S, Gilbert J, Fleischmann RD, Morrison DA.** 2004. Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. Mol. Microbiol. **51:**1051–1070. http://dx.doi.org/10.1046/j.1365-2958.2003.03907.x.
52. **Ogura M, Yamaguchi H, Kobayashi K, Ogasawara N, Fujita Y, Tanaka T.** 2002. Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. J. Bacteriol. **184:**2344–2351. http://dx.doi.org/10.1128/JB.184.9.2344-2351.2002.
53. **Levin BR, Cornejo OE.** 2009. The population and evolutionary dynamics of homologous gene recombination in bacterial populations. PLoS Genet. **5:**e1000601. http://dx.doi.org/10.1371/journal.pgen.1000601.
54. **Johnsen PJ, Dubnau D, Levin BR.** 2009. Episodic selection and the maintenance of competence and natural transformation in *Bacillus subtilis*. Genetics **181:**1521–1533. http://dx.doi.org/10.1534/genetics.108.099523.
55. **Cornejo OE, Rozen DE, May RM, Levin BR.** 2009. Oscillations in continuous culture populations of *Streptococcus pneumoniae*: population dynamics and the evolution of clonal suicide. Proc. Biol. Sci. **276:**999–1008. http://dx.doi.org/10.1098/rspb.2008.1415.
56. **Wylie CS, Trout AD, Kessler DA, Levine H.** 2010. Optimal strategy for competence differentiation in bacteria. PLoS Genet. **6:**e1001108. http://dx.doi.org/10.1371/journal.pgen.1001108.
57. **Redfield RJ.** 2001. Do bacteria have sex? Nat. Rev. Genet. **2:**634–639. http://dx.doi.org/10.1038/35084593.
58. **Bakkali M.** 2013. Could DNA uptake be a side effect of bacterial adhesion and twitching motility? Arch. Microbiol. **195:**279–289. http://dx.doi.org/10.1007/s00203-013-0870-1.
59. **Lorenz MG, Wackernagel W.** 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. **58:**563–602.
60. **Poje G, Redfield RJ.** 2003. Transformation of *Haemophilus influenzae*. Methods Mol. Med. **71:**57–70.
61. **Maughan H, Redfield RJ.** 2009. Extensive variation in natural competence in *Haemophilus influenzae*. Evolution **63:**1852–1866. http://dx.doi.org/10.1111/j.1558-5646.2009.00658.x.
62. **Evans BA, Rozen DE.** 2013. Significant variation in transformation frequency in *Streptococcus pneumoniae*. ISME J. **7:**791–799. http://dx.doi.org/10.1038/ismej.2012.170.
63. **Pérez-Losada M, Browne EB, Madsen A, Wirth T, Viscidi RP, Crandall KA.** 2006. Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. Infect. Genet. Evol. **6:**97–112. http://dx.doi.org/10.1016/j.meegid.2005.02.003.
64. **Marri PR, Paniscus M, Weyand NJ, Rendón MA, Calton CM, Hernández DR, Higashi DL, Sodergren E, Weinstock GM, Rounsley SD, So M.** 2010. Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species. PLoS One **5:**e11835. http://dx.doi.org/10.1371/journal.pone.0011835.
65. **Lacross NC, Marrs CF, Patel M, Sandstedt SA, Gilsdorf JR.** 2008. High genetic diversity of nontypeable *Haemophilus influenzae* isolates from two children attending a day care center. J. Clin. Microbiol. **46:**3817–3821. http://dx.doi.org/10.1128/JCM.00940-08.
66. **Vos M, Didelot X.** 2009. A comparison of homologous recombination rates in bacteria and archaea. ISME J. **3:**199–208. http://dx.doi.org/10.1038/ismej.2008.93.
67. **Kong Y, Ma JH, Warren K, Tsang RSW, Low DE, Jamieson FB, Alexander DC, Hao W.** 2013. Homologous recombination drives both sequence diversity and gene content variation in *Neisseria meningitidis*. Genome Biol. Evol. **5:**1611–1627. http://dx.doi.org/10.1093/gbe/evt116.
68. **Didelot X, Lawson D, Darling A, Falush D.** 2010. Inference of homologous recombination in bacteria using whole-genome sequences. Genetics **186:**1435–1449. http://dx.doi.org/10.1534/genetics.110.120121.
69. **Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, McGee L, von Gottberg A, Song JH, Ko KS, Pichon B, Baker S, Parry CM, Lambertsen LM, Shahinas D, Pillai DR, Mitchell TJ, Dougan G, Tomasz A, Klugman KP, Parkhill J, Hanage WP, Bentley SD.** 2011. Rapid pneumococcal evolution in response to clinical interventions. Science **331:**430–434. http://dx.doi.org/10.1126/science.1198545.
70. **Claverys JP, Martin B.** 2003. Bacterial "competence" genes: signatures of active transformation, or only remnants? Trends Microbiol. **11:**161–165. http://dx.doi.org/10.1016/S0966-842X(03)00064-7.
71. **Sinha S, Cameron AD, Redfield RJ.** 2009. Sxy induces a CRP-S regulon in *Escherichia coli*. J. Bacteriol. **191:**5180–5195. http://dx.doi.org/10.1128/JB.00476-09.
72. **Stingl K, Muller S, Scheidgen-Kleyboldt G, Clausen M, Maier B.** 2010. Composite system mediates two-step DNA uptake into *Helicobacter pylori*. Proc. Natl. Acad. Sci. U. S. A. **107:**1184–1189. http://dx.doi.org/10.1073/pnas.0909955107.
73. **Kristensen BM, Sinha S, Boyce JD, Bojesen AM, Mell JC, Redfield RJ.**

2012. Natural transformation of *Gallibacterium anatis*. Appl. Environ. Microbiol. **78:**4914–4922. http://dx.doi.org/10.1128/AEM.00412-12.
74. **Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ, Marians KJ.** 2000. The importance of repairing stalled replication forks. Nature **404:**37–41. http://dx.doi.org/10.1038/35003501.
75. **Rabinovich L, Sigal N, Borovok I, Nir-Paz R, Herskovits AA.** 2012. Prophage excision activates *Listeria* competence genes that promote phagosomal escape and virulence. Cell **150:**792–802. http://dx.doi.org/10.1016/j.cell.2012.06.036.
76. **Mortier-Barrière I, Velten M, Dupaigne P, Mirouze N, Piétrement O, McGovern S, Fichant G, Martin B, Noirot P, Le Cam E, Polard P, Claverys JP.** 2007. A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. Cell **130:**824–836. http://dx.doi.org/10.1016/j.cell.2007.07.038.
77. **Quevillon-Cheruel S, Campo N, Mirouze N, Mortier-Barrière I, Brooks MA, Boudes M, Durand D, Soulet A-L, Lisboa J, Noirot P, Martin B, van Tilbeurgh H, Noirot-Gros M-F, Claverys J-P, Polard P.** 2012. Structure-function analysis of pneumococcal DprA protein reveals that dimerization is crucial for loading RecA recombinase onto DNA during transformation. Proc. Natl. Acad. Sci. U. S. A. **109:**E2466–E2475. http://dx.doi.org/10.1073/pnas.1205638109.
78. **Smeets LC, Becker SC, Barcak GJ, Vandenbroucke-Grauls CM, Bitter W, Goosen N.** 2006. Functional characterization of the competence protein DprA/Smf in *Escherichia coli*. FEMS Microbiol. Lett. **263:**223–228. http://dx.doi.org/10.1111/j.1574-6968.2006.00423.x.
79. **Johnson PL, Slatkin M.** 2009. Inference of microbial recombination rates from metagenomic data. PLoS Genet. **5:**e1000674. http://dx.doi.org/10.1371/journal.pgen.1000674.
80. **Redfield RJ, Findlay WA, Bosse J, Kroll JS, Cameron AD, Nash JH.** 2006. Evolution of competence and DNA uptake specificity in the Pasteurellaceae. BMC Evol. Biol. **6:**82. http://dx.doi.org/10.1186/1471-2148-6-82.
81. **Suckow G, Seitz P, Blokesch M.** 2011. Quorum sensing contributes to natural transformation of *Vibrio cholerae* in a species-specific manner. J. Bacteriol. **193:**4914–4924. http://dx.doi.org/10.1128/JB.05396-11.
82. **Friedrich A, Hartsch T, Averhoff B.** 2001. Natural transformation in mesophilic and thermophilic bacteria: identification and characterization of novel, closely related competence genes in *Acinetobacter* sp. strain BD413 and *Thermus thermophilus* HB27. Appl. Environ. Microbiol. **67:**3140–3148. http://dx.doi.org/10.1128/AEM.67.7.3140-3148.2001.
83. **Graupner S, Frey V, Hashemi R, Lorenz MG, Brandes G, Wackernagel W.** 2000. Type IV pilus genes pilA and pilC of *Pseudomonas stutzeri* are required for natural genetic transformation, and *pilA* can be replaced by corresponding genes from nontransformable species. J. Bacteriol. **182:**2184–2190. http://dx.doi.org/10.1128/JB.182.8.2184-2190.2000.
84. **Stone BJ, Kwaik YA.** 1999. Natural competence for DNA transformation by *Legionella pneumophila* and its association with expression of type IV pili. J. Bacteriol. **181:**1395–1402.
85. **Lerman LS, Tolmach LJ.** 1957. Genetic transformation. I. Cellular incorporation of DNA accompanying transformation in *Pneumococcus*. Biochim. Biophys. Acta **26:**68–82.
86. **Soltyk A, Shugar D, Piechowska M.** 1975. Heterologous deoxyribonucleic acid uptake and complexing with cellular constituents in competent *Bacillus subtilis*. J. Bacteriol. **124:**1429–1438.
87. **Dougherty TJ, Asmus A, Tomasz A.** 1979. Specificity of DNA uptake in genetic transformation of gonococci. Biochem. Biophys. Res. Commun. **86:**97–104. http://dx.doi.org/10.1016/0006-291X(79)90386-3.
88. **Scocca JJ, Poland RL, Zoon KC.** 1974. Specificity in deoxyribonucleic acid uptake by transformable *Haemophilus influenzae*. J. Bacteriol. **118:**369–373.
89. **Danner DB, Deich RA, Sisco KL, Smith HO.** 1980. An eleven-base-pair sequence determines the specificity of DNA uptake in *Haemophilus* transformation. Gene **11:**311–318. http://dx.doi.org/10.1016/0378-1119(80)90071-2.
90. **Goodman SD, Scocca JJ.** 1988. Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. Proc. Natl. Acad. Sci. U. S. A. **85:**6982–6986. http://dx.doi.org/10.1073/pnas.85.18.6982.
91. **Maughan H, Wilson LA, Redfield RJ.** 2010. Bacterial DNA uptake sequences can accumulate by molecular drive alone. Genetics **186:**613–627. http://dx.doi.org/10.1534/genetics.110.119438.
92. **Frye SA, Nilsen M, Tønjum T, Ambur OH.** 2013. Dialects of the DNA uptake sequence in Neisseriaceae. PLoS Genet. **9:**e1003458. http://dx.doi.org/10.1371/journal.pgen.1003458.
93. **Berry J-L, Cehovin A, McDowell MA, Lea SM, Pelicic V.** 2013. Functional analysis of the interdependence between DNA uptake sequence and its cognate ComP receptor during natural transformation in Neisseria species. PLoS Genet. **9:**e1004014. http://dx.doi.org/10.1371/journal.pgen.1004014.
94. **Wang Y, Taylor DE.** 1990. Natural transformation in *Campylobacter* species. J. Bacteriol. **172:**949–955.
95. **Levine SM, Lin EA, Emara W, Kang J, DiBenedetto M, Ando T, Falush D, Blaser MJ.** 2007. Plastic cells and populations: DNA substrate characteristics in *Helicobacter pylori* transformation define a flexible but conservative system for genomic variation. FASEB J. **21:**3458–3467. http://dx.doi.org/10.1096/fj.07-8501com.
96. **Saunders NJ, Peden JF, Moxon ER.** 1999. Absence in *Helicobacter pylori* of an uptake sequence for enhancing uptake of homospecific DNA during transformation. Microbiology **145**(Pt 12)**:**3523–3528.
97. **Findlay WA, Redfield RJ.** 2009. Coevolution of uptake sequences and bacterial proteomes. Genome Biol. Evol. **1:**45–55. http://dx.doi.org/10.1093/gbe/evp005.
98. **Kingsford CL, Ayanbule K, Salzberg SL.** 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol. **8:**R22. http://dx.doi.org/10.1186/gb-2007-8-2-r22.
99. **Davidsen T, Rødland EA, Lagesen K, Seeberg E, Rognes T, Tønjum T.** 2004. Biased distribution of DNA uptake sequences towards genome maintenance genes. Nucleic Acids Res. **32:**1050–1058. http://dx.doi.org/10.1093/nar/gkh255.
100. **Treangen TJ, Ambur OH, Tonjum T, Rocha EP.** 2008. The impact of the neisserial DNA uptake sequences on genome evolution and stability. Genome Biol. **9:**R60. http://dx.doi.org/10.1186/gb-2008-9-3-r60.
101. **Bakkali M, Chen TY, Lee HC, Redfield RJ.** 2004. Evolutionary stability of DNA uptake signal sequences in the Pasteurellaceae. Proc. Natl. Acad. Sci. U. S. A. **101:**4513–4518. http://dx.doi.org/10.1073/pnas.0306366101.
102. **Hogg JS, Hu FZ, Janto B, Boissy R, Hayes J, Keefe R, Post JC, Ehrlich GD.** 2007. Characterization and modeling of the *Haemophilus influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains. Genome Biol. **8:**R103. http://dx.doi.org/10.1186/gb-2007-8-6-r103.
103. **Smith HO, Tomb JF, Dougherty BA, Fleischmann RD, Venter JC.** 1995. Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. Science **269:**538–540. http://dx.doi.org/10.1126/science.7542802.
104. **Graves JF, Biswas GD, Sparling PF.** 1982. Sequence-specific DNA uptake in transformation of *Neisseria gonorrhoeae*. J. Bacteriol. **152:**1071–1077.
105. **Sisco KL, Smith HO.** 1979. Sequence-specific DNA uptake in *Haemophilus* transformation. Proc. Natl. Acad. Sci. U. S. A. **76:**972–976. http://dx.doi.org/10.1073/pnas.76.2.972.
106. **Goodgal SH, Mitchell MA.** 1990. Sequence and uptake specificity of cloned sonicated fragments of *Haemophilus influenzae* DNA. J. Bacteriol. **172:**5924–5928.
107. **Ambur OH, Frye SA, Tonjum T.** 2007. New functional identity for the DNA uptake sequence in transformation and its presence in transcriptional terminators. J. Bacteriol. **189:**2077–2085. http://dx.doi.org/10.1128/JB.01408-06.
108. **Cehovin A, Simpson PJ, McDowell MA, Brown DR, Noschese R, Pallett M, Brady J, Baldwin GS, Lea SM, Matthews SJ, Pelicic V.** 2013. Specific DNA recognition mediated by a type IV pilin. Proc. Natl. Acad. Sci. U. S. A. **110:**3065–3070. http://dx.doi.org/10.1073/pnas.1218832110.
109. **Burkhardt J, Vonck J, Averhoff B.** 2011. Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. J. Biol. Chem. **286:**9977–9984. http://dx.doi.org/10.1074/jbc.M110.212688.
110. **Maier B, Chen I, Dubnau D, Sheetz MP.** 2004. DNA transport into *Bacillus subtilis* requires proton motive force to generate large molecular forces. Nat. Struct. Mol. Biol. **11:**643–649. http://dx.doi.org/10.1038/nsmb783.
111. **Bakkali M.** 2007. Genome dynamics of short oligonucleotides: the example of bacterial DNA uptake enhancing sequences. PLoS One **2:**e741. http://dx.doi.org/10.1371/journal.pone.0000741.
112. **Mell JC, Hall IM, Redfield RJ.** 2012. Defining the DNA uptake speci-

ficity of naturally competent *Haemophilus influenzae* cells. Nucleic Acids Res. **40:**8536–8549. http://dx.doi.org/10.1093/nar/gks640.
113. **Albritton WL, Setlow JK, Thomas ML, Sottnek FO.** 1986. Relatedness within the family Pasteurellaceae as determined by genetic transformation. Int. J. Syst. Bacteriol. **36:**103–106. http://dx.doi.org/10.1099/00207713-36-1-103.
114. **Setlow JK, Boling ME, Allison DP, Beattie KL.** 1973. Relationship between prophage induction and transformation in *Haemophilus influenzae*. J. Bacteriol. **115:**153–161.
115. **Albritton WL, Setlow JK, Thomas M, Sottnek F, Steigerwalt AG.** 1984. Heterospecific transformation in the genus *Haemophilus*. Mol. Gen. Genet. **193:**358–363. http://dx.doi.org/10.1007/BF00330693.
116. **Johnson TJ, Fernandez-Alarcon C, Bojesen AM, Nolan LK, Trampel DW, Seemann T.** 2011. Complete genome sequence of *Gallibacterium anatis* strain UMN179, isolated from a laying hen with peritonitis. J. Bacteriol. **193:**3676–3677. http://dx.doi.org/10.1128/JB.05177-11.
117. **Kuhnert P, Korczak BM.** 2006. Prediction of whole-genome DNA-DNA similarity, determination of G+C content and phylogenetic analysis within the family Pasteurellaceae by multilocus sequence analysis (MLSA). Microbiology **152:**2537–2548. http://dx.doi.org/10.1099/mic.0.28991-0.
118. **Kuhnert P, Scholten E, Haefner S, Mayor D, Frey J.** 2010. *Basfia succiniciproducens* gen. nov., sp. nov., a new member of the family Pasteurellaceae isolated from bovine rumen. Int. J. Syst. Evol. Microbiol. **60**(Pt 1)**:**44–50. http://dx.doi.org/10.1099/ijs.0.011809-0.
119. **Dover GA.** 1986. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends Genet. **2:**159–165. http://dx.doi.org/10.1016/0168-9525(86)90211-8.
120. **Berg OG, von Hippel PH.** 1988. Selection of DNA binding sites by regulatory proteins. J. Mol. Biol. **200:**709–723. http://dx.doi.org/10.1016/0022-2836(88)90482-2.
121. **Linxweller W, Hörz W.** 1985. Reconstitution experiments show that sequence-specific histone-DNA interactions are the basis for nucleosome phasing on mouse satellite DNA. Cell **42:**281–290. http://dx.doi.org/10.1016/S0092-8674(85)80123-9.
122. **Bogaert D, Keijser B, Huse S, Rossen J, Veenhoven R, van Gils E, Bruin J, Montijn R, Bonten M, Sanders E.** 2011. Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. PLoS One **6:**e17035. http://dx.doi.org/10.1371/journal.pone.0017035.
123. **Redfield RJ.** 2002. Is quorum sensing a side effect of diffusion sensing? Trends Microbiol. **10:**365–370. http://dx.doi.org/10.1016/S0966-842X(02)02400-9.
124. **Lopez P, Perez Ureña MT, Garcia E, Espinosa M.** 1980. Interactions of homologous and heterologous deoxyribonucleic acids and competent *Bacillus subtilis* cells. J. Bacteriol. **142:**229–235.
125. **Otto SP, Gerstein AC.** 2006. Why have sex? The population genetics of sex and recombination. Biochem. Soc. Trans. **34**(Pt 4)**:**519–522.
126. **Pineda-Krch M, Redfield RJ.** 2005. Persistence and loss of meiotic recombination hotspots. Genetics **169:**2319–2333. http://dx.doi.org/10.1534/genetics.104.034363.
127. **Wang J, Hu W, Lux R, He X, Li Y, Shi W.** 2011. Natural transformation of *Myxococcus xanthus*. J. Bacteriol. **193:**2122–2132. http://dx.doi.org/10.1128/JB.00041-11.
128. **Pelicic V.** 2008. Type IV pili: *e pluribus unum*? Mol. Microbiol. **68:**827–837. http://dx.doi.org/10.1111/j.1365-2958.2008.06197.x.
129. **Hofreuter D, Odenbreit S, Haas R.** 2001. Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. Mol. Microbiol. **41:**379–391. http://dx.doi.org/10.1046/j.1365-2958.2001.02502.x.
130. **Wu M, Eisen JA.** 2008. A simple, fast, and accurate method of phylogenomic inference. Genome Biol. **9:**R151. http://dx.doi.org/10.1186/gb-2008-9-10-r151.

JB Journals.ASM.org

MINIREVIEW

Regulation of DNA Replication Initiation by Chromosome Structure

David Magnan,[a]* David Bates[a,b]

Integrative Molecular and Biomedical Sciences[a] and Department of Molecular and Human Genetics,[b] Baylor College of Medicine, Houston, Texas, USA

Recent advancements in fluorescence imaging have shown that the bacterial nucleoid is surprisingly dynamic in terms of both behavior (movement and organization) and structure (density and supercoiling). Links between chromosome structure and replication initiation have been made in a number of species, and it is universally accepted that favorable chromosome structure is required for initiation in all cells. However, almost nothing is known about whether cells use changes in chromosome structure as a regulatory mechanism for initiation. Such changes could occur during natural cell cycle or growth phase transitions, or they could be manufactured through genetic switches of topoisomerase and nucleoid structure genes. In this review, we explore the relationship between chromosome structure and replication initiation and highlight recent work implicating structure as a regulatory mechanism. A three-component origin activation model is proposed in which thermal and topological structural elements are balanced with *trans*-acting control elements (DnaA) to allow efficient initiation control under a variety of nutritional and environmental conditions. Selective imbalances in these components allow cells to block replication in response to cell cycle impasse, override once-per-cell-cycle programming during growth phase transitions, and promote reinitiation when replication forks fail to complete.

Regulation of the timing and number of replication events is critical for genomic stability and evolutionary fitness in all cells. Normally, all chromosomes in a cell replicate exactly once per division cycle and in a timely manner to allow successful chromosome segregation. Even subtle deviations from this formula can have severe consequences for cell viability, including increased mutation rate and DNA repair stress (1, 2) and increased rates of missegregation, leading to aneuploidy—a major driver of genetic disease, including cancer (3). Precise replication timing is even more critical in bacteria, which have strong evolutionary pressure to replicate and divide as rapidly as possible. Additionally, as most bacteria utilize a single replication origin to replicate their chromosome, origins must fire with 100% efficiency to keep pace. Replication timing precision is illustrated by the extraordinarily low variability in cell mass at the time of replication initiation (coefficient of variation, 9% [4]). During fast growth, all copies of the origin present on a multiforked chromosome (usually 4 or 8) fire simultaneously, with <5% of wild-type cells exhibiting a nonsynchronous initiation (5).

In the majority of cases, once a replication fork is started, it progresses at a relatively constant rate to the terminus. When forks stall (and they frequently do), dedicated and highly conserved mechanisms exist to restart the fork at the site of failure (6). Thus, in all cells replication is controlled at the step of initiation. Regulation of initiation is often considered a binary relationship between the origin (the replicator) and the *trans*-acting protein that catalyzes DNA duplex opening (the initiator). This model, known as the replicon hypothesis, was first proposed by Jacob and colleagues in 1963 (7), and the root principles have been confirmed in all domains of life (8). However, the replicator/initiator relationship is only one component of a larger initiation regulatory system; there is also strong dependence on chromosome structure, loosely measured in terms of supercoiling density (below), both at the origin and globally. For example, the selection and timing of origin firing in eukaryotes are largely dependent on local chromatin structure, with origins in the decondensed regions initiating first (9). Such dependence is not typically associated with bacterial origins, although this view is beginning to change. For instance, binding of the bacterial initiator protein, DnaA, to the bipartite origin of *Helicobacter pylori* is supercoiling dependent (10). Also, replication initiation in *Caulobacter crescentus* is regulated through cell cycle changes in chromosome structure and position (11). It is well established that chromosome condensation in early stationary phase of bacterial growth is highly refractive to initiation of replication and transcription (12), both of which require duplex melting, and there is emerging evidence that initiation in *Escherichia coli* is sensitive to chromosome structure changes in the cell cycle (13). In this review, we outline the key determinants of chromosome structure in bacteria and discuss the role of DNA structure in regulating replication initiation.

INITIATION IS A THERMODYNAMIC PROCESS GOVERNED BY FACTORS THAT INFLUENCE DUPLEX MELTING

The dependency of bacterial replication initiation on favorable DNA topology has been known for nearly as long as the requirement for DnaA (14–16); however, the understanding of topology's role as a regulatory mechanism has developed more slowly. This is in part due to the inherent differences in DNA structures of the various experimental systems (*in vitro*, *in vivo*, plasmid, or chromosome) as well as a lack of tools to measure DNA structure.

Accepted manuscript posted online 17 August 2015

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

Editor: W. Margolin

Address correspondence to David Bates, bates@bcm.edu.

* Present address: David Magnan, Biosciences, Rice University, Houston, Texas, USA.

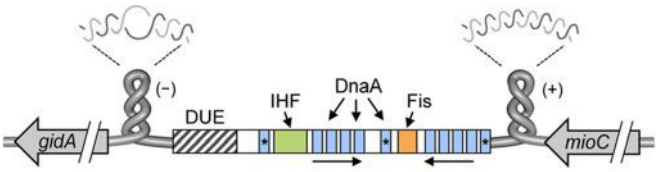


FIG 1 The *E. coli* origin of replication. The 245-bp *oriC* sequence composed of an AT-rich DNA-unwinding element (DUE) and binding sites for DnaA, IHF, and Fis is shown. High-affinity DnaA binding sites (asterisks) and low-affinity DnaA binding site arrays (horizontal arrows) are indicated (20). The transcription direction of the *oriC*-flanking genes, *mioC* and *gidA* (large arrows), with predicted topological effects on DNA supercoiling and duplex twist (thick and thin helices, respectively), is shown.

Replication origins in all cells are thermodynamically unstable AT-rich elements that become single stranded upon the supply of sufficient duplex underwinding (reduced twist in units of base pairs per helical turn). DNA in this state generally forms compensatory negative supercoils, and underwound segments are commonly referred to as negatively supercoiled. Although most natural chromosomes are maintained with a net negative supercoiling, DNA topology fluctuates strongly both along the chromosome and during different phases of growth (17, 18).

DnaA: ONE PART OF THE PUZZLE

DnaA promotes strand opening at *oriC* by modulating nearby DNA topology. In *E. coli*, DnaA is bound to *oriC* for much of the cell cycle at three high-affinity binding sites; then, through a complex maneuver involving an exchange of Fis and integration host factor (IHF) binding (19; see also below), DnaA binds sequentially along two arrays of low-affinity binding sites (20) (Fig. 1). Rozgaja and colleagues (20) propose that oligomerization of DnaA between the two arrays, which are out of helical phase, may introduce torsional strain on the DNA duplex, resulting in strand opening at the adjacent DNA-unwinding element (DUE). After DUE unwinding, DnaA remains bound and may stabilize short-lived unwound structures (21). Most other bacteria have similarly organized origins, with a series of high- and low-affinity DnaA binding sites 50 to 100 bp from an AT-rich DUE (reviewed in reference 22).

Several lines of evidence indicate that DnaA alone is insufficient to drive initiation. First, the ability of DnaA to catalyze strand opening is highly dependent on global DNA context. DNA footprinting in stationary-phase cells permeabilized with ethanol shows that origins have a protein binding signature identical to that of growing cells at the time of initiation (23), suggesting that some other non-origin-binding component is repressing initiation. Given the reduced supercoiling status of stationary-phase chromosomes (17), it is a fair conclusion that origin firing is prevented in these cells by an insufficient level of free negative supercoiling. Similarly, in growing cells the amount of DnaA binding required for initiation varies significantly with growth rate (24), supercoiling status (14), and whether initiation occurs on the chromosome or a plasmid (reviewed in reference 25), which have very different supercoiling buffering capacities. Second, overproduction of DnaA has a limited effect on initiation. High overexpression of wild-type DnaA triggers a rapid initiation event, but subsequent initiations occur at normal once-per-division-cycle intervals (e.g., see reference 26) with only slightly upset initiation synchrony between multiple origins in the same cell (27). It is possible that the excess DnaA molecules are inactivated by the Hda-mediated RIDA (regulatory inactivation of DnaA) system, which hydrolyzes bound ATP on DnaA (28). Supporting this theory, oversupply of a DnaA variant that is RIDA insensitive (DnaA-cos) is lethal, presumably due to overinitiation and subsequent replication fork collapse (1). However, it was very recently shown that subtle (50%) overexpression of ATP-DnaA caused no change in the cell cycle timing of initiation under a wide range of growth conditions (29). On the whole, it appears that DnaA is necessary but not sufficient for replication initiation and that origin function is ultimately dependent on other factors besides DnaA.

OTHER FACTORS THAT AFFECT *oriC* STRUCTURE

In addition to DnaA, several accessory DNA structure-modifying proteins bind *E. coli oriC*, including (but not limited to) IHF, Fis, HU, and SeqA (30). Similarly to DnaA, these proteins affect DNA topology and may regulate initiation by generating favorable or unfavorable torsional strain at the DUE (31–34). None of the accessory proteins are essential, but null mutants exhibit severely asynchronous initiations and grow poorly in rich medium (e.g., see references 34 and 35), suggesting that they are important for initiation timing during multiforked replication. SeqA is a particularly potent negative regulator of initiation, and immediately after initiation, *oriC* is strongly and specifically bound by SeqA protein, which precludes origin firing in a process known as sequestration (36, 37). SeqA binds preferentially to hemimethylated GATC sequences, normally remethylated by DNA adenine methylase ~5 or so min after passage of the replication fork but extended to 10 to 20 min at a few chromosomal loci, including *oriC* (37). Enhanced SeqA binding at the origin may affect initiation by occluding DnaA binding, either at specific DnaA boxes where SeqA and DnaA are juxtaposed (38) or over a broader region by a SeqA-promoted association of *oriC* with the inner membrane (39–42; see also below). In addition, SeqA may directly inhibit DUE melting by forming a RecA-like filament along the GATC-rich origin (33, 43, 44), which has been shown to reduce available free negative supercoiling and block open complex formation on *oriC* plasmids (31, 33). Whatever the exact mechanism of SeqA, independent cycles of *oriC* sequestration and DnaA control are key elements of *E. coli*'s precise and synchronous initiation system (45). The replication origins of *C. crescentus* and *Bacillus subtilis* are also DnaA binding centers; however, unlike *E. coli*, these organisms also utilize master response regulators to modulate structure at the DUE. In *Caulobacter*, the transcription regulator CtrA, which is at the center of a comprehensive cell cycle control network (46), binds the origin (*Cori*) and represses initiation through DnaA occlusion and/or modulating transcription within the origin (47; below). Similarly, Spo0A, originally discovered as a sporulation regulator in *B. subtilis*, also inhibits replication initiation through direct binding to *oriC* (48).

DNA topology is also affected by active transcription complexes, and promoters in and around origins have a stimulatory effect on replication initiation in a number of bacteria, including *E. coli* and *C. crescentus* (e.g., see references 16 and 47). Activating transcription does not prime DNA synthesis, as transcripts lacking a 3′-OH group are fully capable of driving *oriC* initiation *in vitro* and DnaG primase is essential even in the presence of the transcription (16). Instead, it appears that transcription disrupts base pairing at the DUE either by creating a stable R-loop (16) or by the introduction of DNA supercoils from the migrating RNA poly-

merase complex (49). By far, the best-understood relationship between transcription and replication initiation is at the *E. coli* origin, which is flanked by two well-conserved genes, *gidA* and *mioC* (Fig. 1). Given their orientations about *oriC*, the twin-domain supercoiling model (49) predicts that *mioC* transcription introduces duplex overtwist (positive supercoils) into the DUE and *gidA* introduces duplex undertwist (negative supercoils), although some *mioC* transcripts progress completely through *oriC* (50), thus possibly having the opposite effect. Supporting this model, their transcription is strongly cell cycle regulated, with maximal *gidA* (activating) transcription before initiation and maximal *mioC* (inhibiting) transcription immediately after initiation (51). Also, transcription from at least one of the two genes is required for initiation of an *E. coli* extragenic *oriC* replicon, or minichromosome (15, 52), and can be replaced with an antibiotic marker oriented away from the DUE (53). Surprisingly, however, *gidA* and *mioC* are completely dispensable on the chromosome, and in fact, double promoter deletion mutants show no measurable change in growth, initiation rate, or synchrony under a variety of growth conditions (35, 54). This discrepancy may be due to differences in supercoiling buffering capacity between plasmids and the chromosome (25, 54). Why then are these genes and their positions so highly conserved among enterobacteria? One possibility is that their transcriptions help drive initiations under suboptimal conditions or at times outside the normal cell cycle initiation window. Supporting this idea, cells initiating asynchronously via a partial *oriC* deletion required either *mioC* or *gidA* for viability (54). Also, severe overinitiation leading to fork breakage and cell death after thymine starvation is prevented by inactivation of the *gidA* and *mioC* promoters (55). This result suggests that these transcriptions may be part of a (sometimes pathological) response pathway to reinitiate replication on chromosomes with stalled forks. In a greater context, cells may utilize *gidA* and *mioC* to trigger other "nonstandard" initiations, such as those that occur during entry and exit from multiforked (fast growth) replication, which requires division-less initiations and initiation-less divisions, respectively. Another reason that these genes may be so well conserved is that their gene products have an apparent role in cell division (35). As the name implies, *gidA* (glucose-inhibited division) mutants, and to a lesser extent *mioC* mutants, exhibit a delayed cell division phenotype that is exacerbated in rich medium (35, 56). It is possible that replication-dependent expression of these genes, by promoter remodeling at initiation, provides an activating signal to the cell division machinery (35).

FACTORS THAT AFFECT GLOBAL CHROMOSOME STRUCTURE

Superhelical tension along the chromosome is mainly a product of the DNA-unwinding activities of replication and transcription, constraint of free supercoils by nucleoid-associated proteins, and enzymatic control of supercoiling by topoisomerases (57). DNA gyrase and topoisomerase I (Topo I), which introduce and remove negative supercoils, respectively, have strong genetic interactions with DnaA. For instance, deletion of *topA* (Topo I) causes increased negative supercoiling and suppresses the temperature sensitivity of *dnaA46* mutants (58). Conversely, partial loss-of-function mutations in *gyrA* and *gyrB* (gyrase) cause reduced negative supercoiling and enhance the replication defects of *dnaA46* (14). Topoisomerase mutations also disrupt initiation synchrony (59), implying poor initiation control. Supercoiling density is also strongly affected by the nucleoid-associated proteins (60), which can bind and constrain negative supercoils from driving strand-opening reactions. Among these proteins, HU is probably the most important and conserved, with mutants exhibiting severely decondensed nucleoids and reduced supercoiling (61). Conversely, overproduction of HU apparently has the opposite effect, as it suppresses the temperature sensitivity of *dnaA46* (62). Another abundant DNA-binding protein with significant effects on global DNA topology is the *B. subtilis* DnaD protein, which is essential for replication initiation (63). Similarly, SMC in *B. subtilis* and the SMC-like MukB protein in *E. coli* contribute to nucleoid condensation, and mutants have reduced plasmid and chromosome supercoiling and exhibit initiation defects (64, 65). Additionally, *mukB* null mutants are hypersensitive to the gyrase inhibitor novobiocin and are suppressed by a *topA* mutation (66), demonstrating their strong effect on chromosome topology. Importantly, biochemical evidence indicates that DnaA binding to the origin is not supercoiling dependent (67, 68), signifying that the above-observed suppression of DnaA deficiency by supercoiling is not likely caused by increased DnaA binding.

Another factor affecting chromosome supercoiling is transcription. Although duplex unwinding by RNA polymerase generates both positive and negative supercoiling (in front of and behind the transcribing complex), a collective topoisomerase bias toward removal of positive supercoils likely results in a net increase in negative supercoiling (57, 69). Treatment of cells with the RNA polymerase (RNAP) inhibitor rifampin causes immediate decondensation of the nucleoid with reduced supercoiling, presumably resulting from a sudden lack of active RNAP-generated supercoiling (70, 71). The rRNA genes, which account for >80% of all transcription activity in rapidly growing *E. coli* (72), may account for the bulk of the supercoiling effects, as blocking rRNA transcription specifically (by the stringent response) causes nucleoid decondensation (73). Additionally, many highly transcribed genes, including 5 of 7 rRNA genes, are positioned near the origin, in an ~1-Mb zone known as the Ori macrodomain (74). This region displays unique cellular localization (74) and significantly elevated negative supercoiling (57). Inhibiting transcription globally with rifampin (75) or at rRNA operons by the stringent response (73) causes an immediate block to replication initiation.

It has also been shown that various environmental signals such as temperature and osmolarity can greatly affect the levels of chromosome supercoiling, which also have significant effects on replication initiation. Thermal energy promotes DNA duplex denaturation by lengthening hydrogen bonds, which results in decreased bond strength between base pairs. Increasing the temperature of exponentially growing *E. coli* cells by >10°C induces an immediate "round" of DNA replication at all existing origins (76). This so-called heat-induced replication is dependent on a fully intact DUE (77) and probably triggers initiation by decreasing the activation energy of open complex formation. Since only a single round of replication is triggered by an increase in temperature, topological changes are likely quickly compensated for by adjustments to expression of gyrase and Topo I (78), implying that net origin energy status is under homeostatic control (below). Similarly, rapidly increasing osmotic levels (to ~0.5 M NaCl), which results in an immediate but temporary increase in negative supercoiling (79), induces replication initiation in *dnaA46* mutant cells at a restrictive temperature (80) and also in cells blocked for

replication initiation by a chromosome-membrane tether (13; see also below).

CHROMOSOME STRUCTURE CHANGES DURING THE CELL CYCLE

Do chromosome structure changes that might regulate initiation occur predictably during the cell cycle? Both the aforementioned *gidA-mioC* transcription switch and origin sequestration are chromosome structure-modifying events triggered by replication of the *oriC* sequence. Remodeling of DnaA and SeqA at the *gidA* and *mioC* promoters triggers a switch from an initiation-promoting *gidA*-on/*mioC*-off state to an initiation-repressing *gidA*-off/*mioC*-on state (35, 51; see also above). At the same time, strong binding by SeqA protein at hemimethylated *oriC* could restrain negative supercoils through formation of an extended filament (33; see also above). Thus, cell cycle-specific protein remodeling at the time of initiation may induce a local topological state that is incompatible with further DUE opening.

Another source of chromosome structure change that occurs during the cell cycle is the replisome itself, which generates superhelical torque at the fork and leads to nucleoid expansion and reorganization as new material is added and segregated. Sufficient positive supercoiling is generated at the fork that it evidently migrates backwards, wrapping newly replicated daughter DNA duplexes together in what is known as a precatenane (81). Precatenanes for most of the chromosome are estimated to be removed in <10 min (82–85), but several key loci have prolonged entanglement (cohesion), including *oriC*, *ter*, and a right-arm multilocus region (82, 84, 85). Delayed release of these regions correlates precisely with the timing of observed jumps in nucleoid size (length and volume) as measured by HU-mCherry fluorescence in *E. coli* (86). The cause-effect relationship between nucleoid expansion events and the release of cohesion linkages is unknown, but expansion appears to be fueled by rapid wave-like nucleoid density oscillations that migrate back and forth across the nucleoid in the time frame of a few seconds (86). Given the magnitude of nucleoid growth seen during the peak of each expansion event (~15 nm in length per min), there are potentially significant consequences for replication initiation, and further studies are needed to explore this new aspect of chromosome behavior.

TETHERING AND OTHER DRAMATIC CHANGES TO CHROMOSOME STRUCTURE

Some less subtle nucleoid changes seen in growing cells or after drug treatment have unambiguous effects on initiation. We previously observed a period late in the cell cycle in which the nucleoid and chromosomal loci (*oriC* and *ter*) remained relatively motionless (87). After cell birth, an increase in mobility preceded replication initiation, and we speculated that this mobility shift reflected a structural change that licensed a round of replication initiation (87). Both this *ter*-mediated immobility period (88, 89) and the origin sequestration period (39–42) involve specific attachments of the chromosome to the cell membrane. Association of *oriC* with acidic phospholipids in the cell membrane stimulates turnover of bound nucleotide on DnaA, resulting in rejuvenation of the active ATP-DnaA form (reviewed in reference 90), and also sequesters the origin from Dam methylase for an extended period, which results in continued SeqA binding and *oriC* repression (39–41). However, the mechanism by which a *ter*-membrane connection could affect initiation is less clear.

We recently tested whether chromosome-membrane attachments in general can inhibit initiation by artificially tethering the chromosome via a transmembrane-Tet repressor fusion protein and chromosomally inserted *tetO* array (13). This study showed that tethering any chromosomal locus caused a rapid initiation block without affecting replication elongation or any known metabolic or cell cycle response. As tethers placed far (>1 Mb) from *oriC* were no less effective, it is unlikely that the blockage resulted from an increased association of origin-bound DnaA with the inner membrane. Furthermore, initiation blocking could not be suppressed by manipulation of any *trans*-acting initiation factor (including DnaA overexpression), and untethered *oriC* minichromosome replication was unaffected when the host cell chromosome was tethered, indicating that the blocking mechanism operated in *cis*. The only discernible physical effect of tethering was a dramatic decondensation of the nucleoid and global reduction in supercoiling, which may have directly prevented open complex formation at the DUE.

Strikingly, tethering of the chromosome blocked initiation with kinetics nearly identical to those of rifampin treatment, which targets RNA polymerase. Why replication initiation is sensitive to rifampin is a long and unsettled question in bacterial genetics, but the mechanism does not involve production of an essential protein (75) or transcription of the origin-flanking gene *gidA* or *mioC* (54). Like tethering, rifampin treatment causes nucleoid decondensation and reduced chromosome supercoiling (70, 71), and we expect that rifampin and tethering block initiation by the same supercoiling mechanism. Supporting this view, initiation in tethered cells was temporarily restored after treatment with high concentrations of salt, suggesting that a rapid influx of negative supercoiling (above) activated the blocked origins. Together, these findings demonstrate the unconditional requirement for negative supercoiling in replication initiation and point to possible routes for controlling initiation through natural supercoiling transitions (below).

THREE-COMPONENT ENERGY MODEL LINKS REPLICATION INITIATION TO CELL PHYSIOLOGY

Replication initiation is dependent on three major energy components: (i) unregulatable DUE parameters that dictate relaxed DNA hydrogen bonding strength (base composition, temperature, and ionic strength), (ii) *trans*-acting DNA-binding proteins that torque DNA (most notably DnaA and SeqA), and (iii) negative supercoiling, which provides general DNA undertwist (Fig. 2). To maintain matched rates of replication and cell division (balanced growth) under a variety of growth conditions, the sum of these three components must be maintained at a near-constant level. Supporting this model, it is well established that supercoiling levels adapt rapidly to an array of environmental changes such as temperature (78, 91), pH (91), osmolarity (92), and oxygen availability (93). Also, species or mutants with altered levels of supercoiling are more or less tolerant of thermal and ionic extremes (e.g., see references 94 and 95). As described above, supercoiling changes can be either localized at the DUE or global and can result from a number of mechanisms, including altered expression of Topo I and DNA gyrase (reviewed in reference 91), altered topoisomerase function caused by a change in the cell energy (~ATP/ADP ratio) status (92, 93), changes in transcriptional activity, or changing the availability of free supercoils (constraint) by nucleoid binding proteins (79).

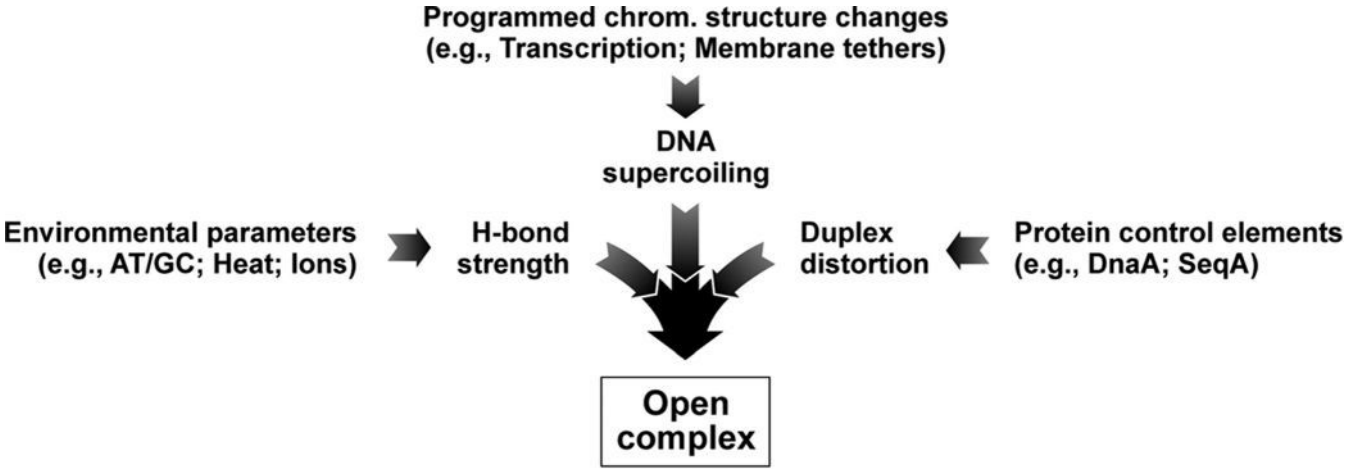


FIG 2 Thermophysical dynamics of replication initiation. Open complex formation at the replication origin is dependent on the cumulative effects of DNA torsional strain, provided by negative supercoiling and protein-mediated duplex distortion, and environmental hydrogen bonding parameters. Independent control of DNA supercoiling and protein binding at *oriC* enables cells to nimbly coordinate initiation with the cell cycle, growth phase, environmental conditions, and stress responses (see the text).

The three-component energy model (Fig. 2) predicts that a change in temperature, supercoiling, or DnaA will result in an immediate but short-lived effect on replication initiation and a slower but stable compensatory adjustment of another energy component. Indeed, this appears to be the case. For example, rapidly reducing negative supercoiling by novobiocin treatment or upshift of a temperature-sensitive gyrase mutant disrupts initiation synchrony (59), enhances the temperature sensitivity of a *dnaA46* mutant (14), and leads to an increase in DnaA expression (96). Also, a temperature upshift of more than 10°C induces a single round of replication initiation (76), followed by a reduction in negative supercoiling via altered expression of gyrase and Topo I (76, 78). A sudden increase in temperature can even induce a round of initiation in the presence of rifampin (76, 97), suggesting that thermal activation can compensate for a gross deficit in negative supercoiling. This kind of homeostatic control of origin energy status presumably allows *E. coli* cells to initiate replication at the proper cell age and mass to achieve balanced growth over a range of temperatures of about 35°C. Such a control feature may explain different requirements for *oriC* depending on its setting: chromosome, plasmid, or *in vitro*. For example, *E. coli* cells can grow at temperatures below 10°C, while open complex formation does not occur below 28°C *in vitro* (98). Or, deletion of roughly half the DnaA binding sites is permissible in chromosomal *oriC* without loss of function, while *oriC* plasmids, which have much lower supercoiling capacity, cannot tolerate deletion of a single binding site (24, 99). We envision that programmed changes to DNA topology, for example, those occurring when the chromosome terminus is attached to the division septum (87), could act as checkpoints to reset chromosome structure to an initiation-competent state and thus ensure a once-per-cell-cycle relationship between replication and division (13). Additionally, cells could create temporary imbalances in supercoiling to change initiation frequency during growth phase changes or in response to replication elongation problems. Both of these latter deviations require breaking the standard rule of a 1:1 ratio of initiation to division.

Of course, the effects of chromosome structure changes are not limited to replication initiation but include all DNA metabolic processes involving strand separation, most notably transcription. These effects are well documented (e.g., see references 12 and 18). Modification of origin supercoiling independently of the rest of the chromosome, such as occurs in thymine-starved cells, which promote hyperinitiation by regulated *gidA* transcription (55), might enable cells to change initiation rate without affecting global supercoiling and thus transcription rates (100).

ACKNOWLEDGMENTS

We thank S. M. Rosenberg, C. Herman, M. C. Joshi, A. K. Barker, B. J. Visser, and P. J. Chen for discussion and helpful comments.

Research was supported by National Institutes of Health (NIH) grant R01GM102679 to D.B.

REFERENCES

1. **Simmons LA, Breier AM, Cozzarelli NR, Kaguni JM.** 2004. Hyperinitiation of DNA replication in *Escherichia coli* leads to replication fork collapse and inviability. Mol Microbiol **51:**349–358. http://dx.doi.org/10.1046/j.1365-2958.2003.03842.x.
2. **Sima J, Gilbert DM.** 2014. Complex correlations: replication timing and mutational landscapes during cancer and genome evolution. Curr Opin Genet Dev **25:**93–100. http://dx.doi.org/10.1016/j.gde.2013.11.022.
3. **Rajagopalan H, Lengauer C.** 2004. Aneuploidy and cancer. Nature **432:**338–341. http://dx.doi.org/10.1038/nature03099.
4. **Boye E, Stokke T, Kleckner N, Skarstad K.** 1996. Coordinating DNA replication initiation with cell growth: differential roles for DnaA and SeqA proteins. Proc Natl Acad Sci U S A **93:**12206–12211. http://dx.doi.org/10.1073/pnas.93.22.12206.
5. **Skarstad K, von Meyenburg K, Hansen FG, Boye E.** 1988. Coordination of chromosome replication initiation in *Escherichia coli*: effects of different *dnaA* alleles. J Bacteriol **170:**852–858.
6. **Yeeles JT, Poli J, Marians KJ, Pasero P.** 2013. Rescuing stalled or damaged replication forks. Cold Spring Harb Perspect Biol **5:**a012815. http://dx.doi.org/10.1101/cshperspect.a012815.
7. **Jacob F, Brenner S, Cuzin F.** 1963. On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon. Cold Spring Harb Symp Quant Biol **23:**329.
8. **O'Donnell M, Langston L, Stillman B.** 2013. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring

Harb Perspect Biol **5:**a010108. http://dx.doi.org/10.1101/cshperspect.a010108.
9. **Ding Q, MacAlpine DM.** 2011. Defining the replication program through the chromatin landscape. Crit Rev Biochem Mol Biol **46:**165–179. http://dx.doi.org/10.3109/10409238.2011.560139.
10. **Donczew R, Mielke T, Jaworski P, Zakrzewska-Czerwinska J, Zawilak-Pawlik A.** 2014. Assembly of *Helicobacter pylori* initiation complex is determined by sequence-specific and topology-sensitive DnaA-*oriC* interactions. J Mol Biol **426:**2769–2782. http://dx.doi.org/10.1016/j.jmb.2014.05.018.
11. **Goley ED, Toro E, McAdams HH, Shapiro L.** 2009. Dynamic chromosome organization and protein localization coordinate the regulatory circuitry that drives the bacterial cell cycle. Cold Spring Harb Symp Quant Biol **74:**55–64. http://dx.doi.org/10.1101/sqb.2009.74.005.
12. **Dorman CJ.** 2013. Genome architecture and global gene regulation in bacteria: making progress towards a unified model? Nat Rev Microbiol **11:**349–355. http://dx.doi.org/10.1038/nrmicro3007.
13. **Magnan D, Joshi MC, Barker AK, Visser BJ, Bates D.** 2015. DNA replication initiation is blocked by a distant chromosome-membrane attachment. Curr Biol **25:**2143–2149. http://dx.doi.org/10.1016/j.cub.2015.06.058.
14. **Filutowicz M.** 1980. Requirement of DNA gyrase for the initiation of chromosome replication in *Escherichia coli* K-12. Mol Gen Genet **177:**301–309. http://dx.doi.org/10.1007/BF00267443.
15. **Leonard AC, Whitford WG, Helmstetter CE.** 1985. Involvement of DNA superhelicity in minichromosome maintenance in *Escherichia coli*. J Bacteriol **161:**687–695.
16. **Baker TA, Kornberg A.** 1988. Transcriptional activation of initiation of replication from the *E. coli* chromosomal origin: an RNA-DNA hybrid near *oriC*. Cell **55:**113–123. http://dx.doi.org/10.1016/0092-8674(88)90014-1.
17. **Postow L, Hardy CD, Arsuaga J, Cozzarelli NR.** 2004. Topological domain structure of the *Escherichia coli* chromosome. Genes Dev **18:**1766–1779. http://dx.doi.org/10.1101/gad.1207504.
18. **Sobetzko P, Travers A, Muskhelishvili G.** 2012. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. Proc Natl Acad Sci U S A **109:**E42–E50. http://dx.doi.org/10.1073/pnas.1108229109.
19. **Ryan VT, Grimwade JE, Camara JE, Crooke E, Leonard AC.** 2004. *Escherichia coli* prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. Mol Microbiol **51:**1347–1359. http://dx.doi.org/10.1046/j.1365-2958.2003.03906.x.
20. **Rozgaja TA, Grimwade JE, Iqbal M, Czerwonka C, Vora M, Leonard AC.** 2011. Two oppositely oriented arrays of low-affinity recognition sites in *oriC* guide progressive binding of DnaA during *Escherichia coli* pre-RC assembly. Mol Microbiol **82:**475–488. http://dx.doi.org/10.1111/j.1365-2958.2011.07827.x.
21. **Speck C, Messer W.** 2001. Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. EMBO J **20:**1469–1476. http://dx.doi.org/10.1093/emboj/20.6.1469.
22. **Leonard AC, Mechali M.** 2013. DNA replication origins. Cold Spring Harb Perspect Biol **5:**a010116. http://dx.doi.org/10.1101/cshperspect.a010116.
23. **Cassler MR, Grimwade JE, McGarry KC, Mott RT, Leonard AC.** 1999. Drunken-cell footprints: nuclease treatment of ethanol-permeabilized bacteria reveals an initiation-like nucleoprotein complex in stationary phase replication origins. Nucleic Acids Res **27:**4570–4576. http://dx.doi.org/10.1093/nar/27.23.4570.
24. **Stepankiw N, Kaidow A, Boye E, Bates DB.** 2009. The right half of the *Escherichia coli* replication origin is not essential for viability, but facilitates multi-forked replication. Mol Microbiol **74:**467–479. http://dx.doi.org/10.1111/j.1365-2958.2009.06877.x.
25. **Asai T, Bates DB, Boye E, Kogoma T.** 1998. Are minichromosomes valid model systems for DNA replication control? Lessons learned from *Escherichia coli*. Mol Microbiol **29:**671–675.
26. **Atlung T, Hansen FG.** 1993. Three distinct chromosome replication states are induced by increasing concentrations of DnaA protein in *Escherichia coli*. J Bacteriol **175:**6537–6545.
27. **Skarstad K, Lobner-Olesen A, Atlung T, von Meyenburg K, Boye E.** 1989. Initiation of DNA replication in *Escherichia coli* after overproduction of the DnaA protein. Mol Gen Genet **218:**50–56. http://dx.doi.org/10.1007/BF00330564.
28. **Kato J, Katayama T.** 2001. Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*. EMBO J **20:**4253–4262. http://dx.doi.org/10.1093/emboj/20.15.4253.
29. **Flatten I, Fossum-Raunehaug S, Taipale R, Martinsen S, Skarstad K.** 2015. The DnaA protein is not the limiting factor for initiation of replication in *Escherichia coli*. PLoS Genet **11:**e1005276. http://dx.doi.org/10.1371/journal.pgen.1005276.
30. **Wolanski M, Donczew R, Zawilak-Pawlik A, Zakrzewska-Czerwinska J.** 2014. *oriC*-encoded instructions for the initiation of bacterial chromosome replication. Front Microbiol **5:**735. http://dx.doi.org/10.3389/fmicb.2014.00735.
31. **Torheim NK, Skarstad K.** 1999. *Escherichia coli* SeqA protein affects DNA topology and inhibits open complex formation at *oriC*. EMBO J **18:**4882–4888. http://dx.doi.org/10.1093/emboj/18.17.4882.
32. **Ryan VT, Grimwade JE, Nievera CJ, Leonard AC.** 2002. IHF and HU stimulate assembly of pre-replication complexes at *Escherichia coli oriC* by two different mechanisms. Mol Microbiol **46:**113–124. http://dx.doi.org/10.1046/j.1365-2958.2002.03129.x.
33. **Odsbu I, Klungsoyr HK, Fossum S, Skarstad K.** 2005. Specific N-terminal interactions of the *Escherichia coli* SeqA protein are required to form multimers that restrain negative supercoils and form foci. Genes Cells **10:**1039–1049. http://dx.doi.org/10.1111/j.1365-2443.2005.00898.x.
34. **Flatten I, Skarstad K.** 2013. The Fis protein has a stimulating role in initiation of replication in *Escherichia coli in vivo*. PLoS One **8:**e83562. http://dx.doi.org/10.1371/journal.pone.0083562.
35. **Lies M, Visser BJ, Joshi MC, Magnan D, Bates D.** 2015. MioC and GidA proteins promote cell division in *E. coli*. Front Microbiol **6:**516. http://dx.doi.org/10.3389/fmicb.2015.00516.
36. **Lu M, Campbell JL, Boye E, Kleckner N.** 1994. SeqA: a negative modulator of replication initiation in *E. coli*. Cell **77:**413–426. http://dx.doi.org/10.1016/0092-8674(94)90156-2.
37. **Slater S, Wold S, Lu M, Boye E, Skarstad K, Kleckner N.** 1995. *E. coli* SeqA protein binds *oriC* in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. Cell **82:**927–936. http://dx.doi.org/10.1016/0092-8674(95)90272-4.
38. **Nievera C, Torgue JJ, Grimwade JE, Leonard AC.** 2006. SeqA blocking of DnaA-oriC interactions ensures staged assembly of the *E. coli* pre-RC. Mol Cell **24:**581–592. http://dx.doi.org/10.1016/j.molcel.2006.09.016.
39. **Ogden GB, Pratt MJ, Schaechter M.** 1988. The replicative origin of the *E. coli* chromosome binds to cell membranes only when hemimethylated. Cell **54:**127–135. http://dx.doi.org/10.1016/0092-8674(88)90186-9.
40. **Chakraborti A, Gunji S, Shakibai N, Cubeddu J, Rothfield L.** 1992. Characterization of the *Escherichia coli* membrane domain responsible for binding *oriC* DNA. J Bacteriol **174:**7202–7206.
41. **Boeneman K, Crooke E.** 2005. Chromosomal replication and the cell membrane. Curr Opin Microbiol **8:**143–148. http://dx.doi.org/10.1016/j.mib.2005.02.006.
42. **Rotman E, Bratcher P, Kuzminov A.** 2009. Reduced lipopolysaccharide phosphorylation in *Escherichia coli* lowers the elevated *ori*/*ter* ratio in *seqA* mutants. Mol Microbiol **72:**1273–1292. http://dx.doi.org/10.1111/j.1365-2958.2009.06725.x.
43. **Guarne A, Brendler T, Zhao Q, Ghirlando R, Austin S, Yang W.** 2005. Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. EMBO J **24:**1502–1511. http://dx.doi.org/10.1038/sj.emboj.7600634.
44. **Leonard AC, Grimwade JE.** 2010. Regulating DnaA complex assembly: it is time to fill the gaps. Curr Opin Microbiol **13:**766–772. http://dx.doi.org/10.1016/j.mib.2010.10.001.
45. **Skarstad K, Katayama T.** 2013. Regulating DNA replication in bacteria. Cold Spring Harb Perspect Biol **5:**a012922. http://dx.doi.org/10.1101/cshperspect.a012922.
46. **McAdams HH, Shapiro L.** 2003. A bacterial cell-cycle regulatory network operating in time and space. Science **301:**1874–1877. http://dx.doi.org/10.1126/science.1087694.
47. **Marczynski GT, Lentine K, Shapiro L.** 1995. A developmentally regulated chromosomal origin of replication uses essential transcription elements. Genes Dev **9:**1543–1557. http://dx.doi.org/10.1101/gad.9.12.1543.
48. **Boonstra M, de Jong IG, Scholefield G, Murray H, Kuipers OP, Veening JW.** 2013. Spo0A regulates chromosome copy number during sporulation by directly binding to the origin of replication in

Bacillus subtilis. Mol Microbiol **87:**925–938. http://dx.doi.org/10.1111/mmi.12141.

49. **Liu LF, Wang JC.** 1987. Supercoiling of the DNA template during transcription. Proc Natl Acad Sci U S A **84:**7024–7027. http://dx.doi.org/10.1073/pnas.84.20.7024.
50. **Rokeach LA, Zyskind JW.** 1986. RNA terminating within the *E. coli* origin of replication: stringent regulation and control by DnaA protein. Cell **46:**763–771. http://dx.doi.org/10.1016/0092-8674(86)90352-1.
51. **Theisen PW, Grimwade JE, Leonard AC, Bogan JA, Helmstetter CE.** 1993. Correlation of gene transcription with the time of initiation of chromosome replication in *Escherichia coli*. Mol Microbiol **10:**575–584. http://dx.doi.org/10.1111/j.1365-2958.1993.tb00929.x.
52. **Ogawa T, Okazaki T.** 1991. Concurrent transcription from the *gid* and *mioC* promoters activates replication of an *Escherichia coli* minichromosome. Mol Gen Genet **230:**193–200. http://dx.doi.org/10.1007/BF00290668.
53. **Asai T, Takanami M, Imai M.** 1990. The AT richness and *gid* transcription determine the left border of the replication origin of the *E. coli* chromosome. EMBO J **9:**4065–4072.
54. **Bates DB, Boye E, Asai T, Kogoma T.** 1997. The absence of effect of *gid* or *mioC* transcription on the initiation of chromosomal replication in *Escherichia coli*. Proc Natl Acad Sci U S A **94:**12497–12502. http://dx.doi.org/10.1073/pnas.94.23.12497.
55. **Martín CM, Viguera E, Guzmán EC.** 2014. Rifampicin suppresses thymineless death by blocking the transcription-dependent step of chromosome initiation. DNA Repair (Amst) **18:**10–17. http://dx.doi.org/10.1016/j.dnarep.2014.03.004.
56. **von Meyenburg K, Hansen FG.** 1980. The origin of replication, *oriC*, of the *Escherichia coli* chromosome: genes near *oriC* and construction of *oriC* deletion mutations. ICN-UCLA Symp Mol Cell Biol **19:**137–159.
57. **Rovinskiy N, Agbleke AA, Chesnokova O, Pang Z, Higgins NP.** 2012. Rates of gyrase supercoiling and transcription elongation control supercoil density in a bacterial chromosome. PLoS Genet **8:**e1002845. http://dx.doi.org/10.1371/journal.pgen.1002845.
58. **Louarn J, Bouche JP, Patte J, Louarn JM.** 1984. Genetic inactivation of topoisomerase I suppresses a defect in initiation of chromosome replication in *Escherichia coli*. Mol Gen Genet **195:**170–174. http://dx.doi.org/10.1007/BF00332741.
59. **von Freiesleben U, Rasmussen KV.** 1992. The level of supercoiling affects the regulation of DNA replication in *Escherichia coli*. Res Microbiol **143:**655–663. http://dx.doi.org/10.1016/0923-2508(92)90060-2.
60. **Dillon SC, Dorman CJ.** 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat Rev Microbiol **8:**185–195. http://dx.doi.org/10.1038/nrmicro2261.
61. **Broyles SS, Pettijohn DE.** 1986. Interaction of the *Escherichia coli* HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch. J Mol Biol **187:**47–60.
62. **Bahloul A, Boubrik F, Rouviere-Yaniv J.** 2001. Roles of *Escherichia coli* histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of *dnaA46ts*. Biochimie **83:**219–229. http://dx.doi.org/10.1016/S0300-9084(01)01246-9.
63. **Zhang W, Allen S, Roberts CJ, Soultanas P.** 2006. The *Bacillus subtilis* primosomal protein DnaD untwists supercoiled DNA. J Bacteriol **188:**5487–5493. http://dx.doi.org/10.1128/JB.00339-06.
64. **Weitao T, Nordstrom K, Dasgupta S.** 2000. *Escherichia coli* cell cycle control genes affect chromosome superhelicity. EMBO Rep **1:**494–499. http://dx.doi.org/10.1093/embo-reports/kvd106.
65. **Lindow JC, Britton RA, Grossman AD.** 2002. Structural maintenance of chromosomes protein of *Bacillus subtilis* affects supercoiling *in vivo*. J Bacteriol **184:**5317–5322. http://dx.doi.org/10.1128/JB.184.19.5317-5322.2002.
66. **Sawitzke JA, Austin SJ.** 2000. Suppression of chromosome segregation defects of *Escherichia coli* muk mutants by mutations in topoisomerase I. Proc Natl Acad Sci U S A **97:**1671–1676. http://dx.doi.org/10.1073/pnas.030528397.
67. **Margulies C, Kaguni JM.** 1996. Ordered and sequential binding of DnaA protein to *oriC*, the chromosomal origin of *Escherichia coli*. J Biol Chem **271:**17035–17040. http://dx.doi.org/10.1074/jbc.271.29.17035.
68. **Weigel C, Schmidt A, Ruckert B, Lurz R, Messer W.** 1997. DnaA protein binding to individual DnaA boxes in the *Escherichia coli* replication origin, *oriC*. EMBO J **16:**6574–6583. http://dx.doi.org/10.1093/emboj/16.21.6574.
69. **Zechiedrich EL, Khodursky AB, Cozzarelli NR.** 1997. Topoisomerase IV, not gyrase, decatenates products of site-specific recombination in *Escherichia coli*. Genes Dev **11:**2580–2592. http://dx.doi.org/10.1101/gad.11.19.2580.
70. **Dworsky P, Schaechter M.** 1973. Effect of rifampin on the structure and membrane attachment of the nucleoid of *Escherichia coli*. J Bacteriol **116:**1364–1374.
71. **Pettijohn DE, Hecht R.** 1974. RNA molecules bound to the folded bacterial genome stabilize DNA folds and segregate domains of supercoiling. Cold Spring Harb Symp Quant Biol **38:**31–41. http://dx.doi.org/10.1101/SQB.1974.038.01.006.
72. **Bremer H, Dennis PP.** 1996. Modulation of chemical composition and other parameters of the cell by growth rate, p 1553–1569. *In* Neidhardt FC, Curtiss R III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed, vol 2. ASM Press, Washington, DC.
73. **Cabrera JE, Cagliero C, Quan S, Squires CL, Jin DJ.** 2009. Active transcription of rRNA operons condenses the nucleoid in *Escherichia coli*: examining the effect of transcription on nucleoid structure in the absence of transertion. J Bacteriol **191:**4180–4185. http://dx.doi.org/10.1128/JB.01707-08.
74. **Espeli O, Mercier R, Boccard F.** 2008. DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. Mol Microbiol **68:**1418–1427. http://dx.doi.org/10.1111/j.1365-2958.2008.06239.x.
75. **Lark KG.** 1972. Evidence for the direct involvement of RNA in the initiation of DNA replication in *Escherichia coli* 15T. J Mol Biol **64:**47–60. http://dx.doi.org/10.1016/0022-2836(72)90320-8.
76. **Botello E, Jimenez-Sanchez A.** 1997. A temperature upshift induces initiation of replication at *oriC* on the *Escherichia coli* chromosome. Mol Microbiol **26:**133–144. http://dx.doi.org/10.1046/j.1365-2958.1997.5621924.x.
77. **Gonzalez-Soltero R, Botello E, Jimenez-Sanchez A.** 2006. Initiation of heat-induced replication requires DnaA and the L-13-mer of *oriC*. J Bacteriol **188:**8294–8298. http://dx.doi.org/10.1128/JB.00814-06.
78. **Mizushima T, Natori S, Sekimizu K.** 1993. Relaxation of supercoiled DNA associated with induction of heat shock proteins in *Escherichia coli*. Mol Gen Genet **238:**1–5.
79. **Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, Bremer E.** 1988. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. Cell **52:**569–584. http://dx.doi.org/10.1016/0092-8674(88)90470-9.
80. **Kondo T, Mima S, Fukuma N, Sekimizu K, Tsuchiya T, Mizushima T.** 2000. Suppression of temperature-sensitivity of a *dnaA46* mutant by excessive DNA supercoiling. Biochem J **348:**375–379.
81. **Postow L, Crisona NJ, Peter BJ, Hardy CD, Cozzarelli NR.** 2001. Topological challenges to DNA replication: conformations at the fork. Proc Natl Acad Sci U S A **98:**8219–8226. http://dx.doi.org/10.1073/pnas.111006998.
82. **Nielsen HJ, Li Y, Youngren B, Hansen FG, Austin S.** 2006. Progressive segregation of the *Escherichia coli* chromosome. Mol Microbiol **61:**383–393. http://dx.doi.org/10.1111/j.1365-2958.2006.05245.x.
83. **Nielsen HJ, Youngren B, Hansen FG, Austin S.** 2007. Dynamics of *Escherichia coli* chromosome segregation during multifork replication. J Bacteriol **189:**8660–8666. http://dx.doi.org/10.1128/JB.01212-07.
84. **Joshi MC, Bourniquel A, Fisher J, Ho BT, Magnan D, Kleckner N, Bates D.** 2011. *Escherichia coli* sister chromosome separation includes an abrupt global transition with concomitant release of late-splitting intersister snaps. Proc Natl Acad Sci U S A **108:**2765–2770. http://dx.doi.org/10.1073/pnas.1019593108.
85. **Joshi MC, Magnan D, Montminy TP, Lies M, Stepankiw N, Bates D.** 2013. Regulation of sister chromosome cohesion by the replication fork tracking protein SeqA. PLoS Genet **9:**e1003673. http://dx.doi.org/10.1371/journal.pgen.1003673.
86. **Fisher JK, Bourniquel A, Witz G, Weiner B, Prentiss M, Kleckner N.** 2013. Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. Cell **153:**882–895. http://dx.doi.org/10.1016/j.cell.2013.04.006.
87. **Bates D, Kleckner N.** 2005. Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. Cell **121:**899–911. http://dx.doi.org/10.1016/j.cell.2005.04.013.
88. **Espeli O, Borne R, Dupaigne P, Thiel A, Gigant E, Mercier R, Boccard F.** 2012. A MatP-divisome interaction coordinates chromosome segre-

gation with cell division in *E. coli*. EMBO J **31:**3198–3211. http://dx.doi.org/10.1038/emboj.2012.128.
89. **Stouf M, Meile JC, Cornet F.** 2013. FtsK actively segregates sister chromosomes in *Escherichia coli*. Proc Natl Acad Sci U S A **110:**11157–11162. http://dx.doi.org/10.1073/pnas.1304080110.
90. **Crooke E.** 2001. *Escherichia coli* DnaA protein–phospholipid interactions: *in vitro* and *in vivo*. Biochimie **83:**19–23. http://dx.doi.org/10.1016/S0300-9084(00)01224-4.
91. **Rui S, Tse-Dinh YC.** 2003. Topoisomerase function during bacterial responses to environmental challenge. Front Biosci **8:**d256–d263. http://dx.doi.org/10.2741/984.
92. **Hsieh LS, Rouviere-Yaniv J, Drlica K.** 1991. Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. J Bacteriol **173:**3914–3917.
93. **Hsieh LS, Burger RM, Drlica K.** 1991. Bacterial DNA supercoiling and [ATP]/[ADP]. Changes associated with a transition to anaerobic growth. J Mol Biol **219:**443–450.
94. **Friedman SM, Malik M, Drlica K.** 1995. DNA supercoiling in a thermotolerant mutant of *Escherichia coli*. Mol Gen Genet **248:**417–422. http://dx.doi.org/10.1007/BF02191641.
95. **Jamroze A, Perugino G, Valenti A, Rashid N, Rossi M, Akhtar M, Ciaramella M.** 2014. The reverse gyrase from *Pyrobaculum calidifontis*, a novel extremely thermophilic DNA topoisomerase endowed with DNA unwinding and annealing activities. J Biol Chem **289:**3231–3243. http://dx.doi.org/10.1074/jbc.M113.517649.
96. **Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR.** 2004. Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. Genome Biol **5:**R87. http://dx.doi.org/10.1186/gb-2004-5-11-r87.
97. **Atlung T, Hansen FG.** 1999. Low-temperature-induced DnaA protein synthesis does not change initiation mass in *Escherichia coli* K-12. J Bacteriol **181:**5557–5562.
98. **van der Ende A, Baker TA, Ogawa T, Kornberg A.** 1985. Initiation of enzymatic replication at the origin of the *Escherichia coli* chromosome: primase as the sole priming enzyme. Proc Natl Acad Sci U S A **82:**3954–3958. http://dx.doi.org/10.1073/pnas.82.12.3954.
99. **Bates DB, Asai T, Cao Y, Chambers MW, Cadwell GW, Boye E, Kogoma T.** 1995. The DnaA box R4 in the minimal *oriC* is dispensable for initiation of *Escherichia coli* chromosome replication. Nucleic Acids Res **23:**3119–3125. http://dx.doi.org/10.1093/nar/23.16.3119.
100. **Ostrer L, Hamann BL, Khodursky A.** 2015. Perturbed states of the bacterial chromosome: a thymineless death case study. Front Microbiol **6:**363. http://dx.doi.org/10.3389/fmicb.2015.00363.

JB Journals.ASM.org

MINIREVIEW

The Precarious Prokaryotic Chromosome

Andrei Kuzminov

Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Evolutionary selection for optimal genome preservation, replication, and expression should yield similar chromosome organizations in any type of cells. And yet, the chromosome organization is surprisingly different between eukaryotes and prokaryotes. The nuclear versus cytoplasmic accommodation of genetic material accounts for the distinct eukaryotic and prokaryotic modes of genome evolution, but it falls short of explaining the differences in the chromosome organization. I propose that the two distinct ways to organize chromosomes are driven by the differences between the global-consecutive chromosome cycle of eukaryotes and the local-concurrent chromosome cycle of prokaryotes. Specifically, progressive chromosome segregation in prokaryotes demands a single duplicon per chromosome, while other "precarious" features of the prokaryotic chromosomes can be viewed as compensations for this severe restriction.

Cells are what their genomes instruct them to be. The observed uniformity, continuity, and robustness of specific life forms reflect how securely their genomes are preserved, how faithfully they are replicated, and how reliably they are expressed to yield specific cellular phenotypes. Formally, the genome is a set of trait-encoding entities (genes) irrespective of how the information is coded, organized, or read. Since the three main functions of genetic information (preservation, replication, and expression) transcend the cell types, one could conservatively expect that, while genome evolution modes might be different between the eukaryotic and prokaryotic cell types (1), the genome organizations would be similar. Indeed, there are basic features of genome organization common to any type of cells: (i) genetic material is always duplex DNA that always replicates semiconservatively; (ii) genetic material is "quantal" in that genes are separate from each other, each gene occupying its own designated stretch of duplex DNA; (iii) besides the gene-encoding DNA, genomes always have noncoding DNA (for example, regulatory regions of genes), as well as selfish elements that use the genome as a habitat in which to multiply; and (iv) major genome changes are either internal rearrangements (usually by DNA repeats) or acquisitions of foreign DNA carrying new genes from the environment (horizontal gene transfer). (Note that, in contrast to cellular life forms, viruses are cell-supported life forms: they are dead outside the cells but can organize their own metabolism and genome replication once inside the host cell. Chemically, viral genomes can be based on RNA or DNA, and either biopolymer can be either single stranded or double stranded. Viral genome organization is diverse and is not covered in this minireview.) There are also differences in genome organization and evolution between eukaryotic cells, which keep their genome in a special compartment called the nucleus, and prokaryotic cells, which keep their genome free-floating in the cytoplasm (1) (with the exception of the membrane-wrapped nucleoid of planctomycetes [2]).

The gene content of the eukaryotic genomes correlates poorly with the genome size (3). There is a lot of noncoding DNA between eukaryotic genes, and the coding sequences of genes themselves are interrupted by introns, both short and long (4). But it is not the random DNA from the environment that inflates the eukaryotic genomes. In fact, eukaryotic genome evolution is not much influenced by horizontal gene transfer, as it is difficult for the unprotected exogenous DNA to reach the nucleus through the cytoplasm, due to the cytoplasmic DNases (5–7) and the cytoplasmic DNA routing that specifically avoids the nucleus (8, 9). The major type of exogenous DNA that has a significant chance of inserting into the eukaryotic genome is the cDNA of retroviruses, single-stranded RNA (ssRNA) viruses that replicate only in the nucleus via the duplex cDNA intermediates integrated into the host genome (10), making retroviral infections a major driver of the eukaryotic genome evolution. The small sizes of the retroviral genomes, the one-enzyme mechanism of retroviral cDNA formation (11), and the rampant recombination during cDNA synthesis (12) breed ever-changing families of simplistic mobile retroelements that infest eukaryotic genomes with thousands of repeats each. These retroelements and the layers of their decaying remnants comprise the bulk of noncoding DNA in the eukaryotic genomes (13–16). The retroelement-derived repeats in eukaryotic genomes facilitate peculiar karyotype fluidity: eukaryotic chromosomes keep exchanging arms with each other, fuse together, or split apart (17, 18). As a result of this constant karyotype reshuffling, even evolutionarily closely related organisms (such as mouse and human) have different numbers of chromosomes and no common genome frame (18, 19). At the same time, "naked" genes from the environment rarely make it into the genomes of higher eukaryotes (20, 21), although horizontal gene transfer does contribute to the genome evolution in unicellular eukaryotes (22).

In contrast, prokaryotic genomes are jam-packed with genes (with minimal intragenic regions and almost no repeats, the genome size becomes an accurate reflection of the gene content) (3), while the very few introns in the prokaryotic genomes are always big, coding for selfish elements (23). In further contrast, prokaryotic genome evolution is dominated by horizontal gene transfer (24, 25), where relatively long uninterrupted chunks of foreign DNA are internalized for food (25) but end up being inserted into the chromosome, becoming part of the genome. Horizontal gene transfer is further enhanced by the "mobilome" (24)—the collection of genes on the extrachromosomal elements (plasmids and phages) staying for a few, or a few thousand, generations within

Published ahead of print 14 March 2014

Address correspondence to kuzminov@life.illinois.edu.

This minireview is dedicated to the memory of Rolf Bernander.

doi:10.1128/JB.00022-14

prokaryotic cells. The efficient horizontal gene transfer and the mobilome allow any particular prokaryote to move into any environment compatible with the general metabolism of the newcomer cell: the habitat-specific genes are supplied later by aboriginal neighbors. In yet another stark difference from eukaryotes, prokaryotic genomes have a few active mobile elements (26), and these few are always tightly controlled, as element jumping or repeat-induced recombination in the gene-packed prokaryotic genomes always reduces adaptation and is often lethal. The low activity of mobile elements is a major contributor to the evolutionarily stable common frames in prokaryotic genomes (related prokaryotes show a high degree of syntheny [27]); another major contributor to the genome frame stability has been recently recognized as the spatiotemporal pattern of nucleoid condensation and regulation of gene expression relative to the origin-terminus axis (28, 29). Finally, in yet another contrast to the ever-inflating eukaryotic genomes, prokaryotic genomes strongly prefer to delete rather than insert DNA; this preference, apparently, drives their unrelenting space crunch (30, 31).

DIFFERENCES BETWEEN EUKARYOTIC AND PROKARYOTIC CHROMOSOMES

There are at least four more specific, structural genome organization features common to both eukaryotic and prokaryotic cells: (i) genes are always arranged as unidimensional chains, like beads on a string (genomic DNA is never branched or star-shaped, for example); (ii) these genomic DNA chains, called chromosomes, are always long, comprising hundreds and thousands of genes; (iii) since the length of the chromosomes is always 100 to 1,000× greater than that of the cell or cellular compartment in which these chromosomes are housed (32–35), chromosomes are always highly compacted, in a local fold-back pattern resembling rosettes of radial loops (36, 37); and (iv) while some "genes from the environment" arrive on chromosomal fragments that would be lost unless incorporated into the host chromosomes, some other environmental genes arrive on small autonomously replicating and segregating extrachromosomal elements, called plasmids. Chromosomes, as specific molecular structures performing certain functions and undergoing certain transitions, are practical representations of the cell's vision of how to best organize preservation, replication, and expression of its genetic information. After billions of years of evolution, the specific chemical way to code information (DNA) and the cell's way to organize the genome (chromosomes) must reflect the winning strategy, evolutionarily optimized over an uncountable number of generations. From this perspective, the major details of the chromosome structure or function are also expected to be similar among all cell types. Surprisingly, beyond the four basic structural aspects mentioned above, the chromosome structures and functions are dramatically different between eukaryotes and prokaryotes, the nuclear versus anuclear organization of genetic material having little relevance to this difference. Indeed, both the prokaryotic chromosome organization and the eukaryotic chromosome organization "rules" allow numerous exceptions of the opposite type, suggesting that at the chromosomal level the dichotomy is maintained by a different kind of selection. The structure/function differences between eukaryotes and prokaryotes in the chromosome organization are compared below (for a different view on the dichotomy, see reference 38). As the "opinions" presented at the end of each section argue, this comparison makes it clear that one of the two ways to organize chromosomes is more precarious than the other.

STRUCTURAL DIFFERENCES

The structural differences between eukaryotic and prokaryotic chromosomes are so dramatically obvious that they, together with the presence or absence of the nucleus itself, were offered to secure the concept of the prokaryotic cell some 50 years ago (1).

Eukaryotic cells have multiple chromosomes per karyotype (complete chromosome set), with a typical diploid number of between 10 and 100 (39, 40). The two reported exceptions with a single chromosome per haploid set are the nematode *Parascaris equorum univalens* (41) and the ant *Mirmecia pilosula* (42), but they are truly unique, because even their closest relatives are multichromosomal. In contrast, bacteria usually have a single chromosome. A few bacteria, such as *Vibrio* (43) or *Brucella* (44) (and a few others [38]), have two chromosomes. All archaea with characterized genomes have a single chromosome (45); *Haloarcula marismortui*, with two, is the only known exception (46).

Plasmids are extrachromosomal DNA molecules with their own replicons/segregons that carry no *sensu stricto* essential genes. Plasmids are rare in the eukaryotic genomes (restricted to lower eukaryotes and fungi); many of them are mitochondrial (compartment of the prokaryotic origin), and all of them are small and adaptationally neutral (47, 48). The "instability" of plasmids in the nucleus is likely due to rapid invasion by retroelements, facilitating their terminal integration into one of the chromosomes via repeat-mediated exchanges. In contrast, the unique prokaryotic chromosome is frequently accompanied by one or a few plasmids. Moreover, prokaryotic plasmids tend to carry genes increasing adaptation of their host cells to specific environments, so they are frequently not adaptationally neutral. In fact, a small fraction of bacterial plasmids, carrying niche-specific essential genes and having chromosome-like GC content and codon usage, are now classified as "chromids" (49, 50) (basically, a part of the genome on an auxiliary replicon). Even though they contribute genetically and readily fuse with the chromosome by repeats provided by mobile elements (recall the famous HFR strains of *Escherichia coli* [51]), for some unclear reason prokaryotic plasmids are not allowed to stay within the chromosome for evolutionarily relevant periods of time, even if their copy number is low.

Eukaryotic chromosomes are always linear. Circular chromosomes can be engineered in eukaryotes but are unstable (52, 53), as there is no mechanism to resolve chromosomal dimers. In contrast, prokaryotic chromosomes are almost always circular (52); at the same time, there are sporadic lineages with linear chromosomes (and/or plasmids) (54). Moreover, circular prokaryotic chromosomes, once made linear using hairpin telomeres, remain stable and fully functional (55). Archaeal chromosomes are always circular (45).

Eukaryotic chromosomes are always equipped with centromeres (either single or multiple ones)—places of attachment of the segregation spindle (56). In contrast, prokaryotic chromosomes are either completely devoid of centromeres or carry the so-called "plasmid centromeres" which are not essential (with a few exceptions, such as *Caulobacter*) (57–60).

Opinion. Multiple chromosomes are better than a single chromosome as the gene storage option (to avoid putting all eggs in a single basket), and linear chromosomes are obviously better than the circular ones, because they avoid the potentially lethal prob-

lems of chromosome dimerization and catenation. To their credit, prokaryotes have successfully solved both problems (61), but the rationale behind such a precarious chromosomal format as the single circular chromosome without a centromere is unclear. Perhaps the single chromosome in prokaryotes facilitates segregation? The eukaryotic response, having protein-mediated long-lasting sister-chromatid cohesion and allocating at least one centromere per chromosome, guarantees faithful segregation during cell division.

FUNCTIONAL DIFFERENCES

DNA condensation and packing. Eukaryotic DNA is wrapped around protein nucleosomes and is further organized by histones and other proteins into a toroidal coil of "30-nm fibers" (39, 62–64), bringing the mass ratio of basic proteins to DNA in the eukaryotic chromatin to ~1 (65). In contrast to this eukaryotic DNA wrapping on spiral rows of histone "bobbins," prokaryotic DNA appears naked in that the isolated nucleoids look like a collection of wire loops, loosely held together by a proteinaceous core (36, 66, 67). To give these disorganized loops some order, prokaryotes make them braid with the help of unique topoisomerases capable of introducing unconstrained DNA superhelicity. Mesophilic prokaryotes employ DNA gyrase to introduce negative supercoils; thermophilic prokaryotes similarly employ reverse gyrase to introduce positive supercoils (68). No topoisomerase capable of introducing unconstrained supercoiling operates in the eukaryotic nucleus (69), only in prokaryote-descendant mitochondria and chloroplasts (68). Prokaryotic DNA is not "naked" in the strict sense, being complexed by thousands of molecules of the nucleoid-associated proteins and transcription factors, and yet the mass ratio of basic protein to DNA in prokaryotic chromosomes is only ~0.02, in line with histoneless chromosomes of dinoflagellates (65). Besides this dynoflagellate exception to the eukaryotic histone packaging rule, there is an opposite exception to the prokaryotic "naked DNA" rule: of the two archeael groups, euryarchaea actually use minimalistic histones to pack their DNA (70).

Opinion. While the eukaryotic DNA looks significantly more secure, the naked prokaryotic DNA is easier to replicate and transcribe.

Replication organization and regulation. Eukaryotic chromosomes have multiple and alternative replication origins (ARSes), generating up to hundreds of replication bubbles per chromosome (Fig. 1A) (71). There are a few preferred origins that tend to fire every replication round, but most origins fire in only a fraction of replication rounds, and if replication is behind schedule in a particular chromosomal region, "*ad hoc*" origins fire in the region, accelerating local replication. In contrast, prokaryotic chromosomes typically have a single, unique replication origin that initiates a single replication bubble per chromosome (Fig. 1A) (72). There are examples of archaeal chromosomes with three or even four origins, though (73, 74).

Termination zones where converging replication forks meet are not defined in the eukaryotic chromosomes (Fig. 1B), even though there may be slow-replication zones, revealed in the S-phase checkpoint mutants in yeast (75) and explained by decreased availability of deoxynucleoside triphosphates (dNTPs) (76). The lack of dedicated termination zones is expected, since replication origin usage differs among replication rounds, shifting the location of replication fork fusion. In contrast, prokaryotic chromosomes, with their unique replication origin, have a defined zone, called the terminus, where converging replication forks fuse (Fig. 1B). Unidirectional termination sites bracket this chromosomal zone to form a "replication fork trap" into which replication forks can enter but from which they cannot escape (77, 78).

Notwithstanding the regulation complexity of multiple replication origins, eukaryotic chromosomes always undergo a single replication round at a time, so that during the S phase the ratio of maximally replicated DNA to unreplicated DNA is always 2:1 (Fig. 1C) (79). In contrast, the prokaryotic chromosomes can have several replication rounds in the same chromosome, so that the ratio of maximally replicated DNA to unreplicated DNA (which in prokaryotes can be expressed as the origin/terminus [ori/ter] ratio) can reach 8:1 (Fig. 1C) (80, 81).

Eukaryotic replication origins fire during the whole S phase, so the "early" ARSes fire at the beginning of S, while the "late" ARSes fire toward the end of S (Fig. 1D) (71). In contrast, in prokaryotes, all replication origins in the same cell always fire at once (synchronously) (Fig. 1D) (82, 83).

Opinion. By all these replication parameters, the eukaryotic organization of replication looks natural, while the prokaryotic way again looks precarious. Why limit the number of replication origins to one? Why insist on a specific termination zone? Why allow the logistical nightmare of several replication rounds in the same cell? And why then demand that they initiate synchronously? With all these arbitrary-looking features, prokaryotes must be experiencing significant stresses in their replication system for an unclear payoff.

SCC. Sister-chromatid cohesion (SCC) is the postreplication state throughout which separation of the sister chromatids is suppressed, so they appear as a single chromosome (84, 85). SCC guards the critical period of maturation of nascent DNA, during which at least four important tasks must be accomplished: (i) introduction of the regular coiling into the newly synthesized duplexes that emerge from the replisomes essentially paranemic, without coils; (ii) linking of Okazaki fragments together (86); (iii) repairing of persistent single-strand gaps and double-strand breaks (87); and (iv) removing the precatenanes that always accumulate behind replication forks (88, 89). In eukaryotic chromosomes, sister-chromatid cohesion is protein-(cohesin)-mediated and lasts several hours, encompassing the whole S, the whole G_2, and part of the M phase until chromatid separation occurs (90). Completely replicated chromosomes do retain a low level of catenation, but it is not responsible for holding sister chromatids together (91). In contrast, in the prokaryotic chromosomes, the duration of sister-chromatid cohesion is short (only 6 min in the rapidly growing *E. coli* bacterium [92]) and the process is mostly mediated by precatenanes (93, 94) (even though sister chromatids may be held together at late-segregating loci by a special protein [95]).

Opinion. It is not clear why the prokaryotic cells have to rush through this critical stage of nascent DNA maturation, especially given that prokaryotic DNA has to immediately undergo the stresses of segregation (as described next).

Segregation. In eukaryotes, chromosomes are segregated once their replication is complete, after additional condensation, by the mitotic spindle, all at once (ensemble segregation), pulled toward the opposite cell poles by microtubules attached to their centromeres (96, 97) (Fig. 2A). In contrast, prokaryotic nucleoids were always known to segregate continuously, as they replicate, and without additional condensation (1, 98–100) (Fig. 2A). These

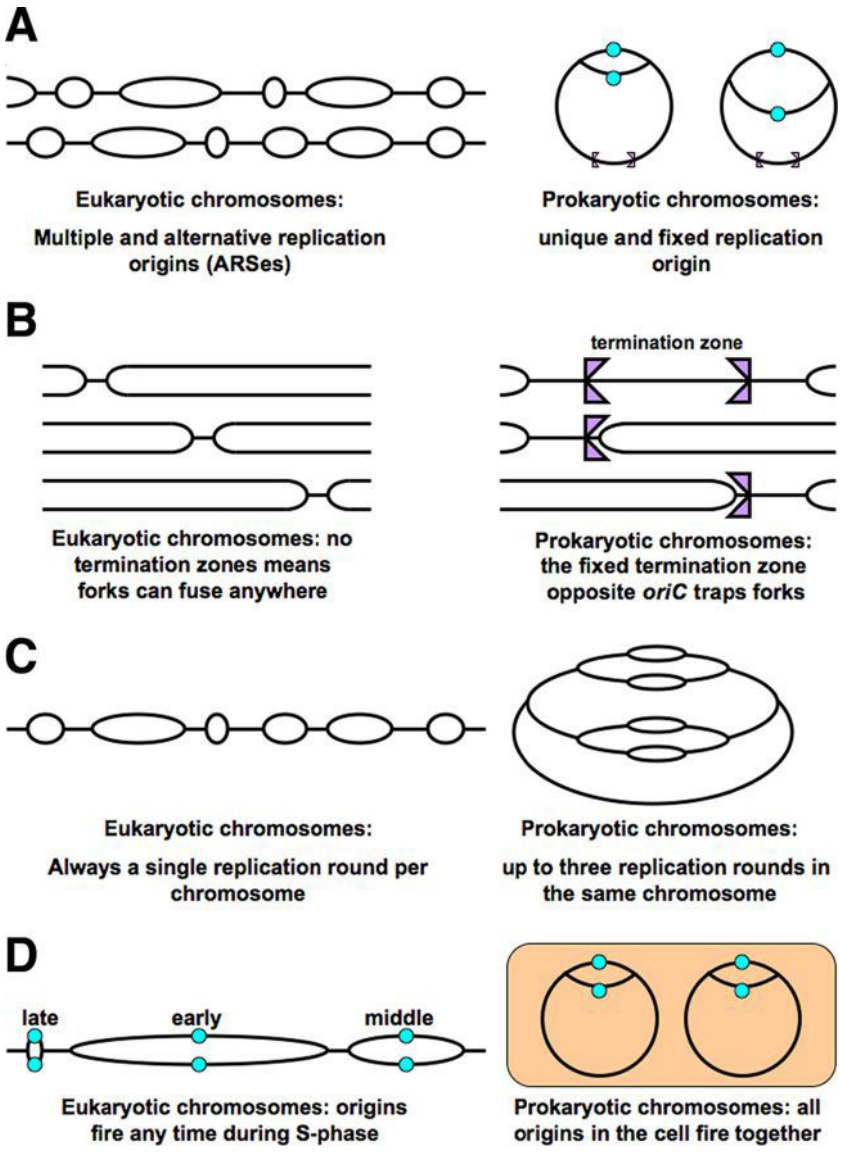


FIG 1 The differences between eukaryotic and prokaryotic chromosomes in organization and regulation of DNA replication. (A) Replication origins. (B) Replication termini. (C) Replication rounds. (D) Timing of origin firing.

days, we say that prokaryotic chromosomes are segregated locally (maybe even at the level of naked DNA), progressively (via the segregation forks [101] moving along the replicating chromosome), and concurrently with replication, once the short SCC is resolved by decatenation (92, 102–104). Nothing is known about the actual mechanism of prokaryotic chromosome segregation, with several possibilities discussed (61, 67, 96).

Opinion. If confirmed, chromosomal segregation at the DNA level in prokaryotes is a precarious molecular manipulation requiring pulling on the naked DNA with forces strong enough to break it.

Replication-transcription conflict. In eukaryotes, genome-average DNA replication rates (the size of genome over the replication time × the calculated number of replication forks), at 40 to 100 bp/s in yeast (105) and 10 to 60 bp/s in human cells (106), are similar to the mRNA transcription rates, at 30 to 60 nucleotides (nt)/s (107, 108) (Fig. 2B), so there is no conflict between the two processes except maybe in the rRNA gene arrays, where the rate is regulated by the replication fork barriers (78) and is further reduced due to the extra replication origins. In contrast, in the prokaryotic chromosomes, the genome-average replication rate during fast growth is at least an order of magnitude higher than the transcription rate (in *E. coli*, 600 to 900 nt/s for DNA synthesis versus 40 to 60 nt/s for transcription [109]) (Fig. 2B), making the conflict between the two processes unavoidable.

The acute reality of this conflict is reflected in the spectacular coorientation of the actively transcribed genes with the direction of replication in prokaryotic chromosomes (110, 111) (Fig. 2C). There are bacterial genomes with more than 80% of all genes cooriented with replication (112). The conflict can be demonstrated experimentally, by inversion of part of the replichore (113–115). In contrast, even though some degree of replication/transcription coorientation was proposed for human chromosomes on the basis of *in silico* analysis (116), an essentially random

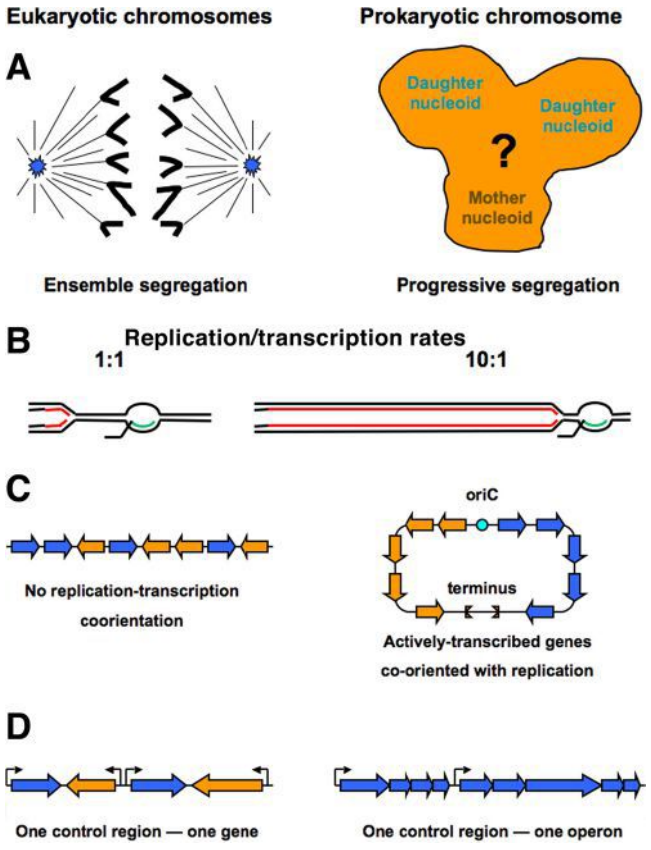


FIG 2 The differences between segregation of eukaryotic and prokaryotic chromosomes and the severity of the replication-transcription conflict. (A) Spindle-driven ensemble segregation in eukaryotes versus the unknown mechanism of progressive segregation in prokaryotes. (B) Schematic differences in rates of replication (red lines) versus transcription (green lines). (C) Gene coorientation with replication through the region. (D) The ratio of control regions to genes.

orientation of genes was found around experimentally identified replication origins (117) (Fig. 2C). In fact, the bidirectional nature of transcription from the strong eukaryotic promoters (118, 119) makes coorientation of genes with replication in eukaryotic chromosomes irrelevant.

The absence of the replication-transcription conflict in the eukaryotic chromosomes is corroborated by the fact that essentially all eukaryotic genes have their own promoters ("one control region = one gene") (Fig. 2D) (120). The promoter recognition algorithms predict a promoter consensus in every kbp of DNA in higher eukaryotes, which is too frequent for primary selection of transcription-initiation sites. Instead, transcription-initiation sites are apparently selected because of a particular nucleosome modification; the distance between stretches of such modified nucleosomes corresponds to the observed 30-to-40-kbp distance between experimentally confirmed transcription-initiation sites (121), while the apparently ubiquitous promoters in these nucleosome-depleted DNA regions then initiate transcription (122). Even eukaryotic genes assembled in functional clusters, analogous to prokaryotic operons, still retain their own individual promoters (123). In contrast to this pattern, the number of recognizable promoters in prokaryotic chromosomes is significantly lower than the number of genes (Fig. 2D): the experimentally identified promoter-to-open-reading-frame (ORF) ratio is ~1:10 for *E. coli* (124) and ~1:3 for *Bacillus subtilis* (111, 125). Chromatin immunoprecipitation with microarray technology (ChIP-chip) analysis of genome sites associated with initiation-poised RNA polymerases in *E. coli* (the "actual promoters") brings this ratio in line with the 1:3 ratio of *B. subtilis* (126). Because of the lower number of available promoters, most prokaryotic genes are assembled into cotranscribed groups called operons, so in prokaryotes, "one control region = one operon."

The organization of prokaryotic genes into operons is often attributed to frequent horizontal gene transfer, which does play a leading role in prokaryotic genome evolution (see above). Indeed, the several-gene limit of a typical horizontally transferred piece promotes clustering of all the genes required for a particular function: when transferred as a cluster, the new genes instantly provide the recipient cell with a useful function, driving selection for clustering (127). However, horizontal gene transfer explains only the physical proximity of genes (clustering itself) (128) and fails to provide selection for coorientation of the genes in the cluster, let alone for their coregulation via promoter sharing. The few evolutionarily stable "superoperons" in bacteria contain multiple genes involved in the same pathway and may have to be cotranscribed not only because the genes need to be coregulated (as originally proposed [129]) but also because the resulting proteins form a complex and need to be coproduced in a particular order for the complex to have the full activity (130). At the same time, most bacterial operons are evolutionary unstable (131, 132) and the majority of recently formed cotranscribed clusters combine genes coding for proteins of unrelated functions (133–135), suggesting that the main evolutionary drive behind combining genes into operons is to reduce the number of transcription-regulation points. Prokaryotes may need to reduce the number of these points because of their gene regulation logic: whereas the nucleosome packing of eukaryotic DNA automatically maintains a transcriptionally restrictive ground state, the naked DNA in prokaryotes is available for transcription at any time, necessitating multiple repressors to hold in check promoter-bound and initiation-poised RNA polymerases (136), sometimes organized in elaborately looped repressosomes (137, 138). I argue that, having initiation-poised promoter-bound RNA polymerases on their DNA, prokaryotic chromosomes evolved to use fewer transcription-control points to reduce the impediment for replication forks.

Opinion. Why are prokaryotes forced to race the high-speed trains on the tracks built for horse-drawn carriages?

THE TWO CHROMOSOME CYCLES

In summary, compared to the organization of eukaryotic chromosomes, the organization of prokaryotic chromosomes seems unnecessarily constrained and precarious, raising the questions about its causes and benefits. At the same time, the nuclear organization of the genetic material in eukaryotes, while defining the eukaryotic mode of genome evolution, cannot explain the "safe" state of the eukaryotic chromosomes compared to the precarious state of the prokaryotic chromosomes, as is abundantly demonstrated by various exceptions. Perhaps the variety of differences between the two chromosome organizations hides one primary difference that constrains the system, necessitating compensations that make prokaryotic chromosomes precarious? To discover what is responsible for constraining prokaryotic chromo-

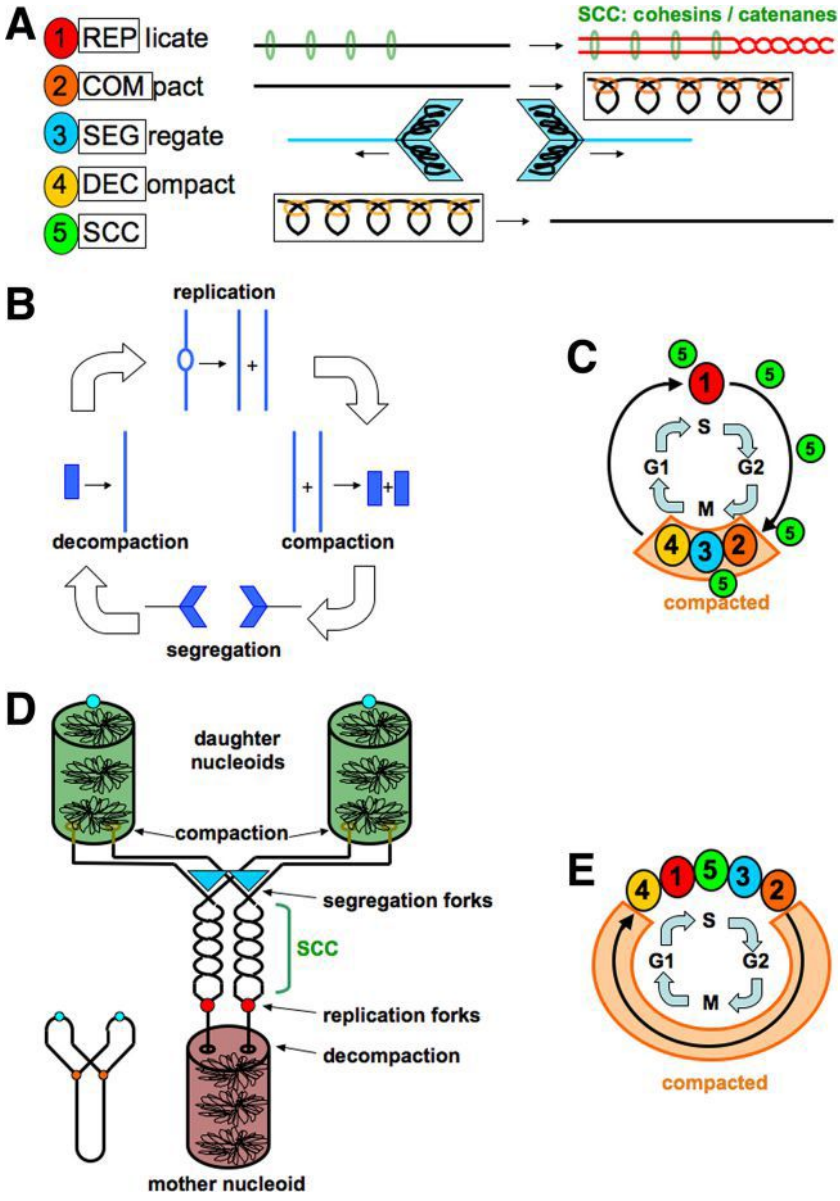


FIG 3 Chromosome transactions and cycles. (A) The five standard chromosome transactions, color coded to correspond to the data in the schemes in panels C and E. (B) The eukaryotic chromosome cycle. (C) Individual transactions of the eukaryotic chromosome cycle over the standard cell cycle grid. (D) The prokaryotic chromosome cycle. (E) Individual transactions of the prokaryotic chromosome cycle over the standard cell cycle grid.

somes, let us consider the differences between eukaryotes and prokaryotes in their chromosome cycles.

The chromosome cycle is defined as a set of chromosomal transactions following a particular order within the cell cycle (101, 139, 140). The cell cycle with the DNA content per cell as a readout [→G_1→S→G_2→D(M)→] is invariant across kingdoms, with the two critical stages, S and D(M), corresponding to two critical chromosomal transactions, replication and segregation (141–143). Because of the necessary ~1,000× degree of intracellular compaction (reviewed in reference 101), chromosomal DNA has to decompact and recompact in order to replicate and segregate. Thus, the full complement of the chromosomal transactions during the cell cycle invariably includes the following steps (Fig. 3A): (i) replication (Rep); (ii) compaction (Com); (iii) segregation (Seg); and (iv) decompaction (Dec) (101, 139, 140). As already mentioned, the fifth transaction, sister-chromatid cohesion (SCC), is a postreplication chromosomal condition that has the DNA component (catenanes) and protein component (cohesins) (Fig. 3A).

The eukaryotic cell cycle is driven by the cyclin-dependent kinase (CDK) engine (144). The eukaryotic chromosome cycle is →Rep→Com→Seg→Dec→ (Fig. 3B) (139, 140) and is driven by the same CDK engine (145, 146). If laid over the cell cycle grid for reference, Rep corresponds to S, G_2 has no chromosomal transactions, and then the Com→Seg→Dec transition happens during M, while G_1 is again devoid of chromosomal transactions (Fig. 3C). Sister-chromatid cohesion is a chromosome condition in eukaryotes that starts before S and ends by the end of M, overlapping

exactly with the cyclin-regulated part of the eukaryotic cell cycle (Fig. 3C) (145, 146). The eukaryotic chromosome cycle is global and consecutive in that the entire set of chromosomes proceeds through a particular stage or transition together before moving to the next stage or transition (Fig. 3B). Also of notice, relative to the maximal degree of compaction during mitosis, eukaryotic chromosomes stay globally decompacted (still locally compacted) most of the cell cycle (Fig. 3C).

The prokaryotic chromosome cycle is based on the version of the Cairns model of theta replication that emphasizes segregation (147) (Fig. 3D, lower left corner) and features a brief period of sister-chromatid cohesion (Fig. 3D). Its sequence is distinct from the one of the eukaryotic chromosome cycle and goes →Dec→Rep→SCC→Seg→Com→ (101). When laid over the invariant cell cycle grid for reference, a single chromosome cycle transition is revealed that comprises all chromosomal transactions, squeezed together into the same "S phase" of the cell cycle, while no chromosomal transitions happen during the G_2, D, and G_1 phases of the prokaryotic cell cycle (Fig. 3E). In particular, prokaryotic SCC is a short stage, sandwiched between Rep and Seg (Fig. 3E). In contrast to the global and consecutive chromosome cycle of eukaryotes, the prokaryotic chromosome cycle is local and concurrent in that, at any given time, only a particular and limited part of the chromosome undergoes all the transactions of the chromosome cycle, while all other parts of the chromosome stay compacted (Fig. 3D). The concurrent nature of the prokaryotic chromosome cycle, in which all chromosomal events roll in a single succession once the replication is initiated, is likely why the prokaryotic cell cycle requires no CDK-like engine and is simply driven by replication initiation (148). In contrast to the eukaryotic chromosomes, prokaryotic chromosomes stay maximally compacted for most of the cell cycle (Fig. 3E), but they do not undergo additional condensation.

PROGRESSIVE CHROMOSOME SEGREGATION OBVIATES PRESORTING AND LOGISTIC NEGOTIATION

Comparison of the two chromosome cycles (Fig. 3B and C versus D and E) suggests that the selection for the precarious prokaryotic chromosome organization is driven by the needs of progressive segregation. Although specific segregation mechanisms of the prokaryotic chromosomes are still unknown, the segregation pattern itself is dramatically different from the eukaryotic one (Fig. 2A) and explains the lack of centromeres in prokaryotic chromosomes. I argue that the unique demands of progressive segregation keep prokaryotic chromosomes inadequately protected and hastily replicated, one obvious example being minimizing the duration of the critical period of sister-chromatid cohesion. However, the major and perhaps a related constrain is that progressive segregation strongly favors a single replication bubble (Fig. 1A). The obvious reason is that multiple replication bubbles, under conditions of progressive segregation, necessitate subnucleoid presorting to ensure that all the daughter subnucleoids with the parental "Watson strand" would group into one daughter nucleoid whereas all the daughter subnucleoids with the parental "Crick strand" would group into the other daughter nucleoid (Fig. 4B and C) (147). Such mechanisms of nonrandom segregation of parts of the chromosome are generally unknown, and there is no reason to suspect their existence in prokaryotes. Without subnucleoid presorting, the random assortment of individual subnucleoids forming around corresponding origins should hopelessly entangle sister nucleoids like two strings of beads (Fig. 4B to D and H). The only (theoretical) way to disentangle such fully replicated and intertwined sister nucleoids would be through "logistic negotiation" (Fig. 4D and C), another hypothetical transaction. Thus, progressive segregation should force prokaryotic chromosomes to assume the "single-duplicon" (replicon plus segregon) configuration (Fig. 1A), even discouraging insertion into the chromosome of plasmids with a copy number of 1.

How do eukaryotic cells solve this problem with their multi-origin chromosomes? Presorting may not even be necessary in eukaryotic chromosomes, because they begin condensation in preparation for segregation only after their replication is complete (Fig. 4E). Moreover, with some degree of coordination, some shorter sister chromosomes may be able to condense into continuous bodies (Fig. 4E to G), rather than into a string of several independently condensed domains (Fig. 4H), while unique centromeres on monocentric eukaryotic chromosomes should make it possible to untangle coordinately condensed chromosomes simply by spindle pulling (Fig. 4I). However, the suspected local presegregation in eukaryotic chromosomes (149, 150) and the likely lack of coordination between condensation events in different chromosome subdomains (Fig. 4H), especially in eukaryotes with holocentric chromosomes (Fig. 4J), make entangling of chromosome subdomains a potentially colossal problem for eukaryotic chromosomes. This problem in eukaryotes is likely addressed by the system of logistic negotiation hypothesized above that disentangles condensed sister-chromatid subdomains and groups all subdomains with the "Watson" strand on one side and all those with the "Crick" strand of the other (Fig. 4H to J). I speculate that the extended SCC period in eukaryotes that covers the good half of their cell cycle (Fig. 3C) is required to accomplish this logistic negotiation process. Remarkably, the crenarchaeote *Sulfolobus*, which has three replication origins in its chromosome (73), does not segregate sister chromatids concurrently with replication, like bacteria (or single-origin archaea [151, 152]), but instead keeps completely replicated daughter chromosomes together during a long G_2 phase employing some kind of DNA junctions, rapidly segregating them just before cell division (82, 153, 154). The long G_2 phase with no daughter chromosome separation may therefore mark the period of similar logistic negotiation in *Sulfolobus*.

PROKARYOTIC CHROMOSOME ORGANIZATION COMPENSATES FOR THE SINGLE DUPLICON

Forcing the entire chromosome to replicate and segregate by a single duplication bubble, while eliminating the need for subnucleoid presorting and logistic negotiation, makes the chromosome duplication round unacceptably long, demanding serious minimization of the limiting stage, which is the chromosome replication time. The minimal replication time in *E. coli* is an impressive 42 min (109) (translating into the overall DNA synthesis rate of ~100,000 kbp per min), and yet it is still much longer than *E. coli*'s shortest cell division time of 24 min (109). Minimization of the chromosome replication time in prokaryotes is achieved in multiple ways (Fig. 5A) as follows.

(i) The enzymatic bacterial replicase rate is at least 10 times higher than that of their eukaryotic counterparts. The *in vitro* rate of purified main bacterial replicase DNA polymerase (pol) III is ~500 nt/s (155–157). The directly measured rates of replication fork propagation *in vivo*, at 620 to 700 nt/s at 30°C (158, 159) and 1,300 nt/s at 42°C (159), are even higher, and there is evidence that the rate is limited by the rate of DNA pol III chain elongation

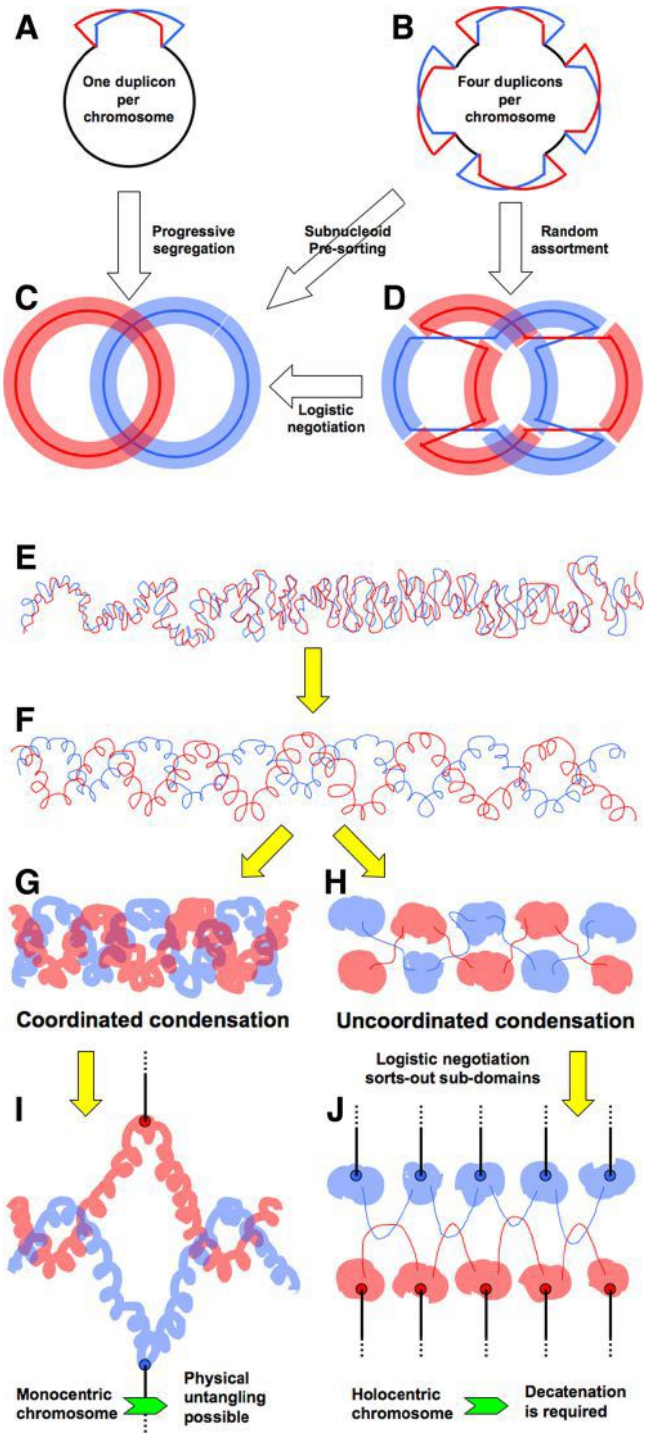


FIG 4 Subdomain presorting and logistic negotiation during chromosome segregation. Red and blue lines designate daughter duplexes containing, correspondingly, "Watson" or "Crick" strands of the parental duplex. (A to D) Prokaryotic chromosome. (A) The theta-replicating chromosome with a single-duplication bubble. (B) A similar replicating chromosome with four duplication bubbles. (C) Progressive segregation from a single-duplication bubble by default yields two completely separate daughter nucleoids. (D) Progressive segregation without subnucleoid presorting yields two daughter nucleoids intertwined due to misclustering of the individual subnucleoids. As a result, the daughter DNA duplex containing, for example, the "Watson" strand of the parental duplex finds itself in both daughter nucleoids. (E to J) Eukaryotic chromosome. (E) Still-to-be-condensed sister chromatids after replication. (F) Gradual condensation sorts sister chromatids out at the level of subdomains. (G) Coordinated condensation results in "single-body" chromosomes ready for segregation. (H) Uncoordinated independent condensation centers produce entangled subdomains. (I) Monocentric chromosomes condensed as "one body" should be able to disentangle during segregation. (J) Holocentric chromosomes likely need logistic negotiation to help sort out all the Watson subdomains (one sister) from all the Crick subdomains (the other sister) before segregation can even take place.

(160). In contrast, the maximal rate of the yeast leading-strand DNA polymerase epsilon *in vitro* is only 50 nt/s, although under the same conditions the lagging-strand DNA polymerase delta moves faster, at 200 nt/s (157). The directly measured rate of replication fork progression *in vivo* ranges between 10 and 100 bp/s in yeast (105, 161, 162) and between 5 and 100 bp/s in human cells (106).

(ii) Prokaryotes keep their DNA histone free to minimize impeding the progress of replication forks. We know this from *E. coli* mutants that replicate even faster than wild-type (WT) cells—they make chromosomal DNA even more "naked," by inactivating nucleoid-associated proteins H-NS and HU (163, 164) or by titrating another DNA-binding protein, DnaA (165).

(iii) Prokaryotes coorient reasonably expressed genes with replication, to minimize replication-transcription conflict. The replication fork trap in the terminus region apparently serves the same purpose, by not allowing replication forks to enter chromosomal regions with opposite transcription. The reduction in the number of promoters diminishes the number of idling RNA polymerases on DNA, further reducing the conflict.

(iv) The circularity of the prokaryotic chromosomes may have nothing to do with getting rid of telomeres, since linearization of the chromosome does not increase the chromosome replication or cell division time in *E. coli* (55). However, circularization automatically minimizes the replication time in a chromosome duplicated by a single bubble, by ensuring that the two replication forks always terminate at the same time, so there is no extra wait for one of them. In the circular chromosomes, termination of the two forks is simultaneous by definition, no matter how chromosomal rearrangements displace the origin relative to the terminus.

(v) In contrast to eukaryotic genomes, where deletions and insertions happen at equal rates, deletions outnumber insertions at least 10:1 in the prokaryotic genomes (30, 31), which systematically reduces the amount of DNA to replicate. The specific mechanisms favoring deletions over insertions are not known; the bias might be a mechanistic consequence of the way prokaryotes segregate their DNA (to be discussed elsewhere).

Remarkably, progressive segregation, having brought these serious demands for the fastest possible replication, also created a general solution for the problem of synthesizing enough DNA in case the replication fork rate becomes inadequate for the cell mass growth rate. The elegant solution is to permit multiple replication cycles on the same chromosome (166) (Fig. 5B). Eukaryotes can have only one replication round per chromosome, while prokaryotes have easily three consecutive rounds going on in the same chromosome (81), made possible by concurrent replication-segregation immediately generating daughter nucleoids that, though incomplete, are fully proficient to initiate their own replication round as soon as the eclipse period is over (167). For this system to work smoothly, synchronization of replication initiation would be a necessary feature, as again observed.

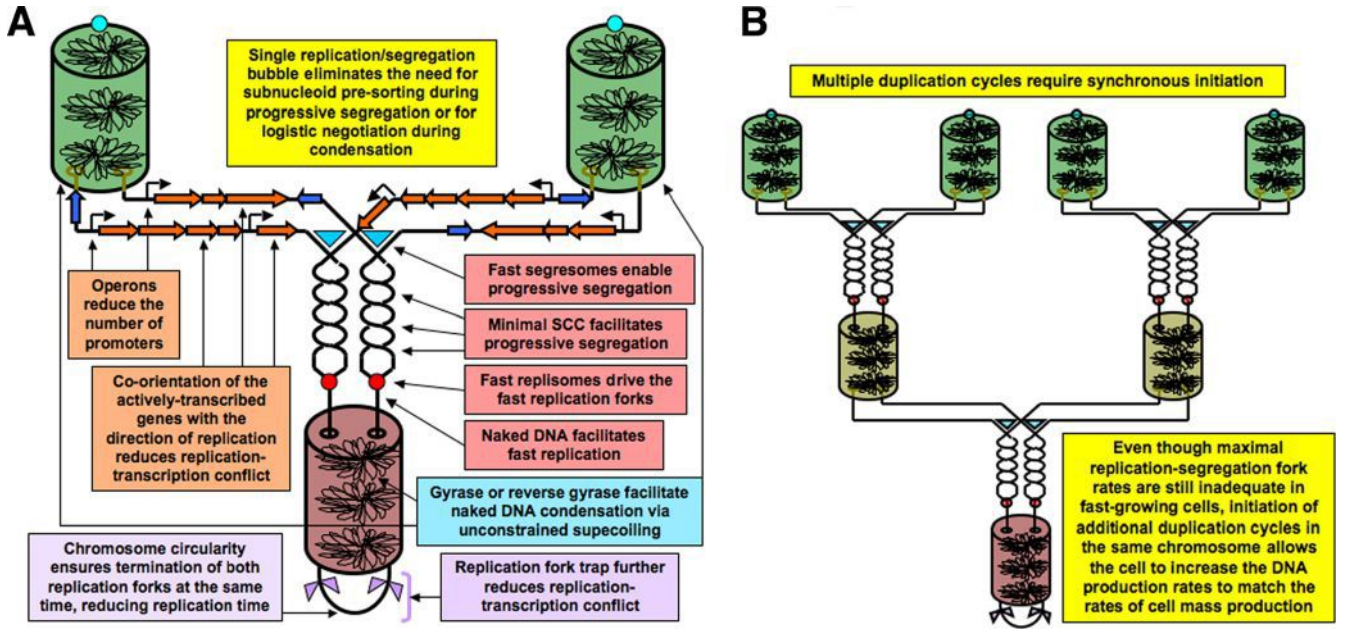


FIG 5 Prokaryotic chromosome organization compensates for the single duplicon but also creates a strategic opportunity. (A) Various factors minimizing the chromosome duplication time. (B) Multiple relocation cycles in the same chromosome strategically solve the duplication problem.

WHY DO PROKARYOTES PREFER A SINGLE MASTER CHROMOSOME?

Perhaps the only prevalent prokaryotic chromosomal feature that does not compensate for the burden of progressive segregation, and instead exacerbates it, is the single chromosome in most prokaryotes. In fact, prokaryotic cells could have easily solved their multiple chromosomal problems caused by unique duplicon if, instead of a single big chromosome, they had had multiple smaller (plasmid-like) chromosomes, even if each one of them was still driven by a unique replication/segregation origin. An example of such genome organization could be "multiple minicircles" of the chloroplast genome in dinoflagellates (168). With a genome comprising 10 to 30 small chromosomes, prokaryotic cells would have been able to bring down the rate of replication to match the rate of transcription. Moreover, they would not have had to perform progressive segregation, performing the plasmid "condense-and-polarize" segregation instead (96) and thus being able to extend SCC to guarantee the proper maturation of nascent DNA. The problem of logistic negotiation would have been permanently solved by pairwise and independent segregation of multiple-duplicon chromosomes. Yet, against all these apparent benefits, bacteria still prefer to "put all their eggs in one basket," evolving a single main chromosome and, sometimes, an additional plasmid-derived chromid, which is always smaller than the master chromosome.

One reason for consolidation of the whole genome into a single chromosome could be that prokaryotic cells have problems handling several independent plasmid segregation systems due to the various incompatibility issues the plasmid systems are known for (169); this explanation is corroborated by the paucity of prokaryotes harboring multiple plasmids. However, the key to the real reason may be the fact that, even with several replicons in the cell, there is always only one replicon driven by the *oriC*/DnaA pair. The unique *oriC*/DnaA pair per cell may be behind the preference for a single chromosome in bacteria, for example, due to the fact that it is initiation at *oriC* by DnaA that is believed to pace the bacterial cell cycle (148). According to this logic, since other replicons in the same cell are driven by their *oriP*/Rep pairs, they should not influence the cell cycle.

However, this simple idea is inconsistent with the fact that additional *oriC*/DnaA-driven plasmids are well tolerated in *E. coli*, at least under laboratory conditions (170). In fact, they initiate replication together with the chromosome (171), at the same time maintaining a higher copy number (170). It could be that, while the *oriC*/DnaA-specific initiation of the master chromosome starts the cell cycle, replication of the terminus (*terC*) in the same master chromosome signals its finish. According to this logic, the master duplicon drives the cell cycle by both its initiation and termination events, whereas other duplicons are tolerated as long as they duplicate within the duplication period of the master duplicon—this could be why the chromids are always smaller than the *oriC*-containing chromosome. If both the initiation and the termination of the master duplicon indeed pace the cell cycle in prokaryotes, this creates selection for the housekeeping genes to relocate from secondary duplicons to the master duplicon as the most stable one. At the same time, multiple *oriC-terC* duplicons would not be tolerated, because all *oirC* genes of the cell fire at once (replication synchrony), and if the variously sized *terC*-containing chromosomes were then to terminate at various times, this could disorient the cell cycle, which is anchored by both the initiation and termination events. Thus, a corollary of the arrangement when the cell cycle is driven by both initiation and termination of the master duplicon is migration of the housekeeping genes from other chromosomes to this particular chromosome, eventually making it a single chromosome in the cell.

CONCLUSION

We have presented an argument that it is the progressive chromosome segregation, possibly operating on naked DNA, that drives the evolution of prokaryotic chromosome organization to be so precarious and so different from the eukaryotic one. Progressive segregation is possible only when the duration of SCC is short and is practical only with a single replication-segregation bubble per chromosome, which, in turn, creates a real chromosome duplication rate crisis. To minimize the chromosome duplication time, prokaryotes employ the fastest known replisomes, keep their DNA naked, coorient most of their transcription with replication, reduce the number of sites where RNA polymerases idle (promoters), and keep the chromosome circular so that the two forks always terminate simultaneously, while the replication fork trap in the terminus prevents replication fork entry into the wrong replichore (Fig. 5A). However, all these features are not enough, and the minimal duplication time of the *E. coli* chromosome (~45 min) could be still almost two times longer than the minimal division time under the optimal growth conditions. The major relief comes from the possibility of having multiple duplication rounds in the same chromosome, synchronously initiated from the unique replication origins in the initiation-competent, though incomplete, daughter nucleoids (Fig. 5B).

It should be stressed that, even though the concurrent prokaryotic chromosome cycle was likely developed to minimize the chromosome duplication time and to disengage the chromosome cycle from the cell cycle (101), many bacteria always have a single chromosome cycle per cell, just as in eukaryotes. In fact, among the model bacteria illustrating the prokaryotic chromosome cycle, *Caulobacter* is incapable of multiple chromosome cycles in the same cell, and yet this does not make its chromosome cycle different (at least in the major aspects) from the one of *E. coli*. At the same time, the archaeote *Sulfolobus*, with three replication origins, does have a more eukaryote-like chromosome cycle, in that its segregation is a stage distinct from replication and is separated from it by an extended "postreplicative sister-chromatid synapsis" period. Thus, the chromosome cycle distinction is not between slow-growing versus fast-growing prokaryotes, but it might be between single-origin versus multiple-origin chromosomes.

Challenging the proposed argument with experimental tests should be facilitated by the various exceptions to the eukaryotic versus prokaryotic "chromosome rules." For example, does the mode of prokaryotic genome evolution apply to planktomycetes that house their nucleoid within the membranous compartment (2)? Dinoflagellates, the eukaryotic protists that, like prokaryotes, maintain condensed chromosomes throughout the interphase and lack histone-based nucleosome packaging of DNA (172), could be predicted to have prokaryote-like fast DNA replication and progressive chromosome segregation (whatever its mechanisms turn out to be). Spectacular pictures of mitosis in dinoflagellates (the so-called "dinomitosis") are indeed highly suggestive (173). It should be possible, as was demonstrated recently (174), to set up an experimental system to test the central prediction of the "duplicon" argument that the existence of several replication origins in the prokaryotic chromosomes would create a logistical problem with segregation of the resulting subnucleoids (Fig. 4). Even testing the idea that the prokaryotic chromosome evolution is driven by progressive segregation may become possible one day in a fantastic synthetic organism, in which the overall eukaryotic chromosome organization will be asked to evolve under the pressure of the prokaryotic progressive chromosome segregation as the only segregation mechanism available. Without such an experimental test, this otherwise compelling collective argument will retain its mostly philosophical nature.

ACKNOWLEDGMENTS

I thank Francois-Xavier Barre (Centre de Genetique Moleculaire, CNRS) for insightful discussions and encouragement and three anonymous reviewers for clarifying the presentation and strengthening the arguments.

Experimental work in this laboratory is supported by grant GM 073115 from the National Institutes of Health.

I declare that I have no conflicts of interest.

REFERENCES

1. **Stanier RY, van Niel CB.** 1962. The concept of a bacterium. Arch. Mikrobiol. **42:**17–35. http://dx.doi.org/10.1007/BF00425185.
2. **Fuerst JA, Sagulenko E.** 2011. Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. Nat. Rev. Microbiol. **9:**403–413. http://dx.doi.org/10.1038/nrmicro2578.
3. **Koonin EV.** 2011 The logic of chance: the nature and origin of biological evolution. FT Press Science, Upper Saddle River, NJ.
4. **Cavalier-Smith T.** 1985. Eukaryote gene numbers, non-coding DNA and genome size, p 69–103. *In* Cavalier-Smith T (ed), The evolution of genome size. John Wiley & Sons, Chichester, United Kingdom.
5. **Lechardeur D, Sohn KJ, Haardt M, Joshi PB, Monck M, Graham RW, Beatty B, Squire J, O'Brodovich H, Lukacs GL.** 1999. Metabolic instability of plasmid DNA in the cytosol: a potential barrier to gene transfer. Gene Ther. **6:**482–497. http://dx.doi.org/10.1038/sj.gt.3300867.
6. **Pollard H, Toumaniantz G, Amos JL, Avet-Loiseau H, Guihard G, Behr JP, Escande D.** 2001. Ca2+-sensitive cytosolic nucleases prevent efficient delivery to the nucleus of injected plasmids. J. Gene Med. **3:**153–164. http://dx.doi.org/10.1002/jgm.160.
7. **Shimizu N, Kamezaki F, Shigematsu S.** 2005. Tracking of microinjected DNA in live cells reveals the intracellular behavior and elimination of extrachromosomal genetic material. Nucleic Acid Res. **33:**6296–6307. http://dx.doi.org/10.1093/nar/gki946.
8. **Akita H, Ito R, Khalil IA, Futaki S, Harashima H.** 2004. Quantitative three-dimensional analysis of the intracellular trafficking of plasmid DNA transfected by a nonviral gene delivery system using confocal laser scanning microscopy. Mol. Ther. **9:**443–451. http://dx.doi.org/10.1016/j.ymthe.2004.01.005.
9. **Lechardeur D, Lukacs GL.** 2006. Nucleocytoplasmic transport of plasmid DNA: a perilous journey from the cytoplasm to the nucleus. Hum. Gene Ther. **17:**882–889. http://dx.doi.org/10.1089/hum.2006.17.882.
10. **Goff SP.** 1992. Genetics of retroviral integration. Annu. Rev. Genet. **26:**527–544. http://dx.doi.org/10.1146/annurev.ge.26.120192.002523.
11. **Whitcomb JM, Hughes SH.** 1992. Retroviral reverse transcription and integration: progress and problems. Annu. Rev. Cell Biol. **8:**275–306. http://dx.doi.org/10.1146/annurev.cb.08.110192.001423.
12. **Negroni M, Buc H.** 2001. Mechanisms of retroviral recombination. Annu. Rev. Genet. **35:**275–302. http://dx.doi.org/10.1146/annurev.genet.35.102401.090551.
13. **Boeke JD.** 2003. The unusual phylogenetic distribution of retrotransposons: a hypothesis. Genome Res. **13:**1975–1983. http://dx.doi.org/10.1101/gr.1392003.
14. **Feschotte C, Jiang N, Wessler SR.** 2002. Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. **3:**329–341. http://dx.doi.org/10.1038/nrg793.
15. **Löwer R, Löwer J, Kurth R.** 1996. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proc. Natl. Acad. Sci. U. S. A. **93:**5177–5184. http://dx.doi.org/10.1073/pnas.93.11.5177.
16. **Smit AF.** 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr. Opin. Genet. Dev. **9:**657–663. http://dx.doi.org/10.1016/S0959-437X(99)00031-3.
17. **Eichler EE, Sankoff D.** 2003. Structural dynamics of eukaryotic chromosome evolution. Science **301:**793–797. http://dx.doi.org/10.1126/science.1086132.

18. **Mieczkowski PA, Lemoine FJ, Petes TD.** 2006. Recombination between retrotransposons as a source of chromosome rearrangements in the yeast *Saccharomyces cerevisiae*. DNA Repair **5:**1010–1020. http://dx.doi.org/10.1016/j.dnarep.2006.05.027.
19. **Zhao H, Bourque G.** 2009. Recovering genome rearrangements in the mammalian phylogeny. Genome Res. **19:**934–942. http://dx.doi.org/10.1101/gr.086009.108.
20. **Andersson JO.** 2005. Lateral gene transfer in eukaryotes. Cell. Mol. Life Sci. **62:**1182–1197. http://dx.doi.org/10.1007/s00018-005-4539-z.
21. **Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ.** 2009. Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. Plant Cell **21:**1897–1911. http://dx.doi.org/10.1105/tpc.109.065805.
22. **Andersson JO.** 2009. Horizontal gene transfer between microbial eukaryotes. Methods Mol. Biol. **532:**473–487. http://dx.doi.org/10.1007/978-1-60327-853-9_27.
23. **Toro N.** 2003. Bacteria and Archaea group II introns: additional mobile genetic elements in the environment. Environ. Microbiol. **5:**143–151. http://dx.doi.org/10.1046/j.1462-2920.2003.00398.x.
24. **Koonin EV, Wolf YI.** 2008. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. Nucleic Acids Res. **36:**6688–6719. http://dx.doi.org/10.1093/nar/gkn668.
25. **Lorenz MG, Wackernagel W.** 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. **58:**563–602.
26. **Hartl DL, Sawyer SA.** 1988. Why do unrelated insertion sequences occur together in the genome of *Escherichia coli*? Genetics **118:**537–541.
27. **Bentley SD, Parkhill J.** 2004. Comparative genomic structure of prokaryotes. Annu. Rev. Genet. **38:**771–792. http://dx.doi.org/10.1146/annurev.genet.38.072902.094318.
28. **Dorman CJ.** 2013. Genome architecture and global gene regulation in bacteria: making progress towards a unified model? Nat. Rev. Microbiol. **11:**349–355. http://dx.doi.org/10.1038/nrmicro3007.
29. **Sobetzko P, Travers A, Muskhelishvili G.** 2012. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. Proc. Natl. Acad. Sci. U. S. A. **109:**E42–E50. http://dx.doi.org/10.1073/pnas.1108229109.
30. **Kuo CH, Ochman H.** 2009. Deletional bias across the three domains of life. Genome Biol. Evol. **1:**145–152. http://dx.doi.org/10.1093/gbe/evp016.
31. **Mira A, Ochman H, Moran NA.** 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet. **17:**589–596. http://dx.doi.org/10.1016/S0168-9525(01)02447-7.
32. **Guacci V, Hogan E, Koshland D.** 1994. Chromosome condensation and sister chromatid pairing in budding yeast. J. Cell Biol. **125:**517–530. http://dx.doi.org/10.1083/jcb.125.3.517.
33. **Holmes VF, Cozzarelli NR.** 2000. Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. Proc. Natl. Acad. Sci. U. S. A. **97:**1322–1324. http://dx.doi.org/10.1073/pnas.040576797.
34. **Trask B, Pinkel D, van den Engh G.** 1989. The proximity of DNA sequences in interphase cell nuclei is correlated to genomic distance and permits ordering of cosmids spanning 250 kilobase pairs. Genomics **5:**710–717. http://dx.doi.org/10.1016/0888-7543(89)90112-2.
35. **Trun NJ, Marko JF.** 1998. Architecture of a bacterial chromosome. ASM News **64:**276–283. http://tigger.uic.edu/~jmarko/Trun.ASMNews.98.pdf.
36. **Kavenoff R, Bowen BC.** 1976. Electron microscopy of membrane-free folded chromosomes from *Escherichia coli*. Chromosoma **59:**89–101. http://dx.doi.org/10.1007/BF00328479.
37. **Marsden MP, Laemmli UK.** 1979. Metaphase chromosome structure: evidence for a radial loop model. Cell **17:**849–858. http://dx.doi.org/10.1016/0092-8674(79)90325-8.
38. **Bendich AJ, Drlica K.** 2000. Prokaryotic and eukaryotic chromosomes: what's the difference? Bioessays **22:**481–486. http://dx.doi.org/10.1002/(SICI)1521-1878(200005)22:5<469::AID-BIES9>3.0.CO;2-4.
39. **Ris H, Kubai DF.** 1970. Chromosome structure. Annu. Rev. Genet. **4:**263–294. http://dx.doi.org/10.1146/annurev.ge.04.120170.001403.
40. **Wikipedia contributors.** 19 September 2013. List of organisms by chromosome count. *On* Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/wiki/List_of_organisms_by_chromosome_count. Accessed 17 March 2014.
41. **White MJD.** 1973. Animal cytology and evolution, 3rd ed. Cambridge University Press, London, United Kingdom.
42. **Crosland MW, Crozier RH.** 1986. *Myrmecia pilosula*, an ant with only one pair of chromosomes. Science **231:**1278. http://dx.doi.org/10.1126/science.231.4743.1278.
43. **Yamaichi Y, Iida T, Park KS, Yamamoto K, Honda T.** 1999. Physical and genetic map of the genome of *Vibrio parahaemolyticus*: presence of two chromosomes in *Vibrio* species. Mol. Microbiol. **31:**1513–1521. http://dx.doi.org/10.1046/j.1365-2958.1999.01296.x.
44. **Jumas-Bilak E, Michaux-Charachon S, Bourg G, O'Callaghan D, Ramuz M.** 1998. Differences in chromosome number and genome rearrangements in the genus *Brucella*. Mol. Microbiol. **27:**99–106. http://dx.doi.org/10.1046/j.1365-2958.1998.00661.x.
45. **Mardanov AV, Ravin NV.** 2012. The impact of genomics on research in diversity and evolution of archaea. Biochemistry (Mosc) **77:**799–812. http://dx.doi.org/10.1134/S0006297912080019.
46. **Baliga NS, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, Shannon P, Chiu Y, Weng RS, Gan RR, Hung P, Date SV, Marcotte E, Hood L, Ng WV.** 2004. Genome sequence of *Haloarcula marismortui*: a halophilic archaeon from the Dead Sea. Genome Res. **14:**2221–2234. http://dx.doi.org/10.1101/gr.2700304.
47. **Farrar NA, Williams KL.** 1988. Nuclear plasmids in the simple eukaryotes *Saccharomyces cerevisiae* and *Dictyostelium discoideum*. Trends Genet. **4:**343–348. http://dx.doi.org/10.1016/0168-9525(88)90054-6.
48. **Griffiths AJ.** 1995. Natural plasmids of filamentous fungi. Microbiol. Rev. **59:**673–685.
49. **Harrison PW, Lower RP, Kim NK, Young JP.** 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. Trends Microbiol. **18:**141–148. http://dx.doi.org/10.1016/j.tim.2009.12.010.
50. **Petersen J, Frank O, Göker M, Pradella S.** 2013. Extrachromosomal, extraordinary and essential—the plasmids of the *Roseobacter* clade. Appl. Microbiol. Biotechnol. **97:**2805–2815. http://dx.doi.org/10.1007/s00253-013-4746-8.
51. **Umeda M, Ohtsubo E.** 1989. Mapping of insertion elements IS1, IS2 and IS3 on the *Escherichia coli* K-12 chromosome. Role of the insertion elements in formation of Hfrs and F' factors and in rearrangement of bacterial chromosomes. J. Mol. Biol. **208:**601–614.
52. **Ishikawa F, Naito T.** 1999. Why do we have linear chromosomes? A matter of Adam and Eve. Mutat. Res. **434:**99–107.
53. **McClintock B.** 1932. A correlation of ring-shaped chromosomes with variegation in *Zea mays*. Proc. Natl. Acad. Sci. U. S. A. **18:**677–681. http://dx.doi.org/10.1073/pnas.18.12.677.
54. **Hinnebusch J, Tilly K.** 1993. Linear plasmids and chromosomes in bacteria. Mol. Microbiol. **10:**917–922. http://dx.doi.org/10.1111/j.1365-2958.1993.tb00963.x.
55. **Cui T, Moro-oka N, Ohsumi K, Kodama K, Ohshima T, Ogasawara N, Mori H, Wanner B, Niki H, Horiuchi T.** 2007. *Escherichia coli* with a linear genome. EMBO Rep. **8:**181–187. http://dx.doi.org/10.1038/sj.embor.7400880.
56. **Schulman I, Bloom KS.** 1991. Centromeres: an integrated protein/DNA complex required for chromosome movement. Annu. Rev. Cell Biol. **7:**311–336. http://dx.doi.org/10.1146/annurev.cb.07.110191.001523.
57. **Draper GC, Gober JW.** 2002. Bacterial chromosome segregation. Annu. Rev. Microbiol. **56:**567–597. http://dx.doi.org/10.1146/annurev.micro.56.012302.160729.
58. **Gordon GS, Wright A.** 2000. DNA segregation in bacteria. Annu. Rev. Microbiol. **54:**681–708. http://dx.doi.org/10.1146/annurev.micro.54.1.681.
59. **Possoz C, Junier I, Espeli O.** 2012. Bacterial chromosome segregation. Front. Biosci. (Landmark Ed) **17:**1020–1034. http://dx.doi.org/10.2741/3971.
60. **Toro E, Shapiro L.** 2010. Bacterial chromosome organization and segregation. Cold Spring Harb. Perspect. Biol. **2:**a000349. http://dx.doi.org/10.1101/cshperspect.a000349.
61. **Reyes-Lamothe R, Nicolas E, Sherratt DJ.** 2012. Chromosome replication and segregation in bacteria. Annu. Rev. Genet. **46:**121–143. http://dx.doi.org/10.1146/annurev-genet-110711-155421.
62. **Huberman JA.** 1973. Structure of chromosome fibers and chromosomes. Annu. Rev. Biochem. **42:**355–378. http://dx.doi.org/10.1146/annurev.bi.42.070173.002035.
63. **Sajan SA, Hawkins RD.** 2012. Methods for identifying higher-order chromatin structure. Annu. Rev. Genomics Hum. Genet. **13:**59–82. http://dx.doi.org/10.1146/annurev-genom-090711-163818.
64. **Widom J.** 1998. Structure, dynamics, and function of chromatin in vitro. Annu. Rev. Biophys. Biomol. Struct. **27:**285–327. http://dx.doi.org/10.1146/annurev.biophys.27.1.285.

65. **Rizzo PJ.** 1976. Basic chromosomal proteins in lower eukaryotes: relevance to the evolution and function of histones. J. Mol. Evol. **8:**79–94. http://dx.doi.org/10.1007/BF01738884.
66. **Postow L, Hardy CD, Arsuaga J, Cozzarelli NR.** 2004. Topological domain structure of the *Escherichia coli* chromosome. Genes Dev. **18:** 1766–1779. http://dx.doi.org/10.1101/gad.1207504.
67. **Wang X, Montero Llopis P, Rudner DZ.** 2013. Organization and segregation of bacterial chromosomes. Nat. Rev. Genet. **14:**191–203. http://dx.doi.org/10.1038/nrg3375.
68. **Forterre P, Gribaldo S, Gadelle D, Serre MC.** 2007. Origin and evolution of DNA topoisomerases. Biochimie **89:**427–446. http://dx.doi.org/10.1016/j.biochi.2006.12.009.
69. **Bates AD, Maxwell A.** 2007. Energy coupling in type II topoisomerases: why do they hydrolyze ATP? Biochemistry **46:**7929–7941. http://dx.doi.org/10.1021/bi700789g.
70. **Luijsterburg MS, White MF, van Driel R, Dame RT.** 2008. The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. Crit. Rev. Biochem. Mol. Biol. **43:**393–418. http://dx.doi.org/10.1080/10409230802528488.
71. **Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M.** 2010. Eukaryotic chromosome DNA replication: where, when, and how? Annu. Rev. Biochem. **79:**89–130. http://dx.doi.org/10.1146/annurev.biochem.052308.103205.
72. **Sernova NV, Gelfand MS.** 2008. Identification of replication origins in prokaryotic genomes. Brief. Bioinform. **9:**376–391. http://dx.doi.org/10.1093/bib/bbn031.
73. **Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R.** 2004. Three replication origins in *Sulfolobus species*: synchronous initiation of chromosome replication and asynchronous termination. Proc. Natl. Acad. Sci. U. S. A. **101:**7046–7051. http://dx.doi.org/10.1073/pnas.0400656101.
74. **Pelve EA, Lindås AC, Knöppel A, Mira A, Bernander R.** 2012. Four chromosome replication origins in the archaeon *Pyrobaculum calidifontis*. Mol. Microbiol. **85:**986–995. http://dx.doi.org/10.1111/j.1365-2958.2012.08155.x.
75. **Cha RS, Kleckner N.** 2002. ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. Science **297:**602–606. http://dx.doi.org/10.1126/science.1071398.
76. **Fasullo M, Tsaponina O, Sun M, Chabes A.** 2010. Elevated dNTP levels suppress hyper-recombination in *Saccharomyces cerevisiae* S-phase checkpoint mutants. Nucleic Acids Res. **38:**1195–1203. http://dx.doi.org/10.1093/nar/gkp1064.
77. **Duggin IG, Wake RG, Bell SD, Hill TM.** 2008. The replication fork trap and termination of chromosome replication. Mol. Microbiol. **70:**1323–1333. http://dx.doi.org/10.1111/j.1365-2958.2008.06500.x.
78. **Mirkin EV, Mirkin SM.** 2007. Replication fork stalling at natural impediments. Microbiol. Mol. Biol. Rev. **71:**13–35. http://dx.doi.org/10.1128/MMBR.00030-06.
79. **Diffley JF.** 2011. Quality control in the initiation of eukaryotic DNA replication. Philos. Trans. R. Soc. Lond. B Biol. Sci. **366:**3545–3553. http://dx.doi.org/10.1098/rstb.2011.0073.
80. **Bird RE, Louarn J, Martuscelli J, Caro L.** 1972. Origin and sequence of chromosome replication in *Escherichia coli*. J. Mol. Biol. **70:**549–566. http://dx.doi.org/10.1016/0022-2836(72)90559-1.
81. **Martín CM, Guzmán EC.** 2011. DNA replication initiation as a key element in thymineless death. DNA Repair **10:**94–101. http://dx.doi.org/10.1016/j.dnarep.2010.10.005.
82. **Lindås AC, Bernander R.** 2013. The cell cycle of archaea. Nat. Rev. Microbiol. **11:**627–638. http://dx.doi.org/10.1038/nrmicro3077.
83. **Skarstad K, Katayama T.** 2013. Regulating DNA replication in bacteria. Cold Spring Harb. Perspect. Biol. **5:**a012922. http://dx.doi.org/10.1101/cshperspect.a012922.
84. **Hirano T.** 2000. Chromosome cohesion, condensation, and separation. Annu. Rev. Biochem. **69:**115–144. http://dx.doi.org/10.1146/annurev.biochem.69.1.115.
85. **Nasmyth K, Peters JM, Uhlmann F.** 2000. Splitting the chromosome: cutting the ties that bind sister chromatids. Science **288:**1379–1385. http://dx.doi.org/10.1126/science.288.5470.1379.
86. **Balakrishnan L, Bambara RA.** 2013. Okazaki fragment metabolism. Cold Spring Harb. Perspect. Biol. **5:**a010173. http://dx.doi.org/10.1101/cshperspect.a010173.
87. **Kuzminov A.** 1999. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ. Microbiol. Mol. Biol. Rev. **63:**751–813.
88. **Peter BJ, Ullsperger C, Hiasa H, Marians KJ, Cozzarelli NR.** 1998. The structure of supercoiled intermediates in DNA replication. Cell **94:**819–827. http://dx.doi.org/10.1016/S0092-8674(00)81740-7.
89. **Sundin O, Varshavsky A.** 1980. Terminal stages of SV40 DNA replication proceed via multiply intertwined catenated dimers. Cell **21:**103–114. http://dx.doi.org/10.1016/0092-8674(80)90118-X.
90. **Díaz-Martínez LA, Giménez-Abián JF, Clarke DJ.** 2008. Chromosome cohesion - rings, knots, orcs and fellowship. J. Cell Sci. **121:**2107–2114. http://dx.doi.org/10.1242/jcs.029132.
91. **Farcas AM, Uluocak P, Helmhart W, Nasmyth K.** 2011. Cohesin's concatenation of sister DNAs maintains their intertwining. Mol. Cell **44:**97–107. http://dx.doi.org/10.1016/j.molcel.2011.07.034.
92. **Nielsen HJ, Youngren B, Hansen FG, Austin S.** 2007. Dynamics of Escherichia coli chromosome segregation during multifork replication. J. Bacteriol. **189:**8660–8666. http://dx.doi.org/10.1128/JB.01212-07.
93. **Lesterlin C, Gigant E, Boccard F, Espéli O.** 2012. Sister chromatid interactions in bacteria revealed by a site-specific recombination assay. EMBO J. **31:**3468–3479. http://dx.doi.org/10.1038/emboj.2012.194.
94. **Wang X, Reyes-Lamothe R, Sherratt DJ.** 2008. Modulation of *Escherichia coli* sister chromosome cohesion by topoisomerase IV. Genes Dev. **22:**2426–2433. http://dx.doi.org/10.1101/gad.487508.
95. **Joshi MC, Magnan D, Montminy TP, Lies M, Stepankiw N, Bates D.** 2013. Regulation of sister chromosome cohesion by the replication fork tracking protein SeqA. PLoS Genet. **9:**e1003673. http://dx.doi.org/10.1371/journal.pgen.1003673.
96. **Ghosh SK, Hajra S, Paek A, Jayaram M.** 2006. Mechanisms for chromosome and plasmid segregation. Annu. Rev. Biochem. **75:**211–241. http://dx.doi.org/10.1146/annurev.biochem.75.101304.124037.
97. **McIntosh JR, Grishchuk EL, West RR.** 2002. Chromosome-microtubule interactions during mitosis. Annu. Rev. Cell Dev. Biol. **18:** 193–219. http://dx.doi.org/10.1146/annurev.cellbio.18.032002.132412.
98. **Løbner-Olesen A, Kuempel PL.** 1992. Chromosome partitioning in *Escherichia coli*. J. Bacteriol. **174:**7883–7889.
99. **Lundgren M, Malandrin L, Eriksson S, Huber H, Bernander R.** 2008. Cell cycle characteristics of crenarchaeota: unity among diversity. J. Bacteriol. **190:**5362–5367. http://dx.doi.org/10.1128/JB.00330-08.
100. **Youngren B, Nielsen HJ, Jun S, Austin S.** 2014. The multifork *Escherichia coli* chromosome is a self-duplicating and self-segregating thermodynamic ring polymer. Genes Dev. **28:**71–84. http://dx.doi.org/10.1101/gad.231050.113.
101. **Kuzminov A.** 2013. The chromosome cycle of prokaryotes. Mol. Microbiol. **90:**214–227. http://dx.doi.org/10.1111/mmi.12372.
102. **Nielsen HJ, Li Y, Youngren B, Hansen FG, Austin S.** 2006. Progressive segregation of the *Escherichia coli* chromosome. Mol. Microbiol. **61:**383–393. http://dx.doi.org/10.1111/j.1365-2958.2006.05245.x.
103. **Vallet-Gely I, Boccard F.** 2013. Chromosomal organization and segregation in *Pseudomonas aeruginosa*. PLoS Genet. **9:**e1003492. http://dx.doi.org/10.1371/journal.pgen.1003492.
104. **Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L.** 2004. Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proc. Natl. Acad. Sci. U. S. A. **101:**9257–9262. http://dx.doi.org/10.1073/pnas.0402606101.
105. **Newlon CS.** 1988. Yeast chromosome replication and segregation. Microbiol. Rev. **52:**568–601.
106. **Berezney R, Dubey DD, Huberman JA.** 2000. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma **108:** 471–484. http://dx.doi.org/10.1007/s004120050399.
107. **Ardehali MB, Lis JT.** 2009. Tracking rates of transcription and splicing in vivo. Nat. Struct. Mol. Biol. **16:**1123–1124. http://dx.doi.org/10.1038/nsmb1109-1123.
108. **Marcello A.** 2012. RNA polymerase II transcription on the fast lane. Transcription **3:**29–34. http://dx.doi.org/10.4161/trns.3.1.19147.
109. **Bremer H, Dennis PP.** 1996. Modulation of chemical composition and other parameters of the cell by growth rate, p 1553–1569. *In* Neidhardt FC, Curtiss R, III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, DC.
110. **Brewer BJ.** 1988. When polymerases collide: replication and the transcriptional organization of the *E. coli* chromosome. Cell **53:**679–686. http://dx.doi.org/10.1016/0092-8674(88)90086-4.
111. **Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V,**

Bertero MG, Bessières P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignell SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi SK, Codani JJ, Connerton IF, Danchin A, et al. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature **390:**249–256. http://dx.doi.org/10.1038/36786.

112. **Bruggemann H, Baumer S, Fricke WF, Wiezer A, Liesegang H, Decker I, Herzberg C, Martinez-Arias R, Merkl R, Henne A, Gottschalk G.** 2003. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. Proc. Natl. Acad. Sci. U. S. A. **100:**1316–1321. http://dx.doi.org/10.1073/pnas.0335853100.
113. **De Septenville AL, Duigou S, Boubakri H, Michel B.** 2012. Replication fork reversal after replication-transcription collision. PLoS Genet. **8:**e1002622. http://dx.doi.org/10.1371/journal.pgen.1002622.
114. **Mirkin EV, Mirkin SM.** 2005. Mechanisms of transcription-replication collisions in bacteria. Mol. Cell. Biol. **25:**888–895. http://dx.doi.org/10.1128/MCB.25.3.888-895.2005.
115. **Wang JD, Berkmen MB, Grossman AD.** 2007. Genome-wide coorientation of replication and transcription reduces adverse effects on replication in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. **104:**5608–5613. http://dx.doi.org/10.1073/pnas.0608999104.
116. **Huvet M, Nicolay S, Touchon M, Audit B, d'Aubenton-Carafa Y, Arneodo A, Thermes C.** 2007. Human gene organization driven by the coordination of replication and transcription. Genome Res. **17:**1278–1285. http://dx.doi.org/10.1101/gr.6533407.
117. **Necsulea A, Guillet C, Cadoret JC, Prioleau MN, Duret L.** 2009. The relationship between DNA replication and human genome organization. Mol. Biol. Evol. **26:**729–741. http://dx.doi.org/10.1093/molbev/msn303.
118. **Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A.** 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature **457:**1038–1042. http://dx.doi.org/10.1038/nature07747.
119. **Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM.** 2009. Bidirectional promoters generate pervasive transcription in yeast. Nature **457:**1033–1037. http://dx.doi.org/10.1038/nature07728.
120. **Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engström PG, Frith MC, Forrest AR, Alkema WB, Tan SL, Plessy C, Kodzius R, Ravasi T, Kasukawa T, Fukuda S, Kanamori-Katayama M, Kitazume Y, Kawaji H, Kai C, Nakamura M, Konno H, Nakano K, Mottagui-Tabar S, Arner P, Chesi A, Gustincich S, Persichetti F, Suzuki H, Grimmond SM, Wells CA, Orlando V, Wahlestedt C, Liu ET, Harbers M, Kawai J, Bajic VB, Hume DA, Hayashizaki Y.** 2006. Genome-wide analysis of mammalian promoter architecture and evolution. Nat. Genet. **38:**626–635. http://dx.doi.org/10.1038/ng1789.
121. **Pedersen AG, Baldi P, Chauvin Y, Brunak S.** 1999. The biology of eukaryotic promoter prediction–a review. Comput. Chem. **23:**191–207. http://dx.doi.org/10.1016/S0097-8485(99)00015-7.
122. **Wei W, Pelechano V, Järvelin AI, Steinmetz LM.** 2011. Functional consequences of bidirectional promoters. Trends Genet. **27:**267–276. http://dx.doi.org/10.1016/j.tig.2011.04.002.
123. **Osbourn AE, Field B.** 2009. Operons. Cell. Mol. Life Sci. **66:**3755–3775. http://dx.doi.org/10.1007/s00018-009-0114-3.
124. **Hershberg R, Bejerano G, Santos-Zavaleta A, Margalit H.** 2001. PromEC: An updated database of Escherichia coli mRNA promoters with experimentally identified transcriptional start sites. Nucleic Acid Res. **29:**277. http://dx.doi.org/10.1093/nar/29.1.277.
125. **Sierro N, Makita Y, de Hoon M, Nakai K.** 2008. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. Nucleic Acids Res. **36**(Database issue): D93–D96. http://dx.doi.org/10.1093/nar/gkm910.
126. **Reppas NB, Wade JT, Church GM, Struhl K.** 2006. The transition between transcriptional initiation and elongation in E. coli is highly variable and often rate limiting. Mol. Cell **24:**747–757. http://dx.doi.org/10.1016/j.molcel.2006.10.030.
127. **Lawrence JG, Roth JR.** 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics **143:**1843–1860.
128. **Lawrence JG.** 2003. Gene organization: selection, selfishness, and serendipity. Annu. Rev. Microbiol. **57:**419–440. http://dx.doi.org/10.1146/annurev.micro.57.030502.090816.
129. **Jacob F, Monod J.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. **3:**318–356. http://dx.doi.org/10.1016/S0022-2836(61)80072-7.
130. **Dandekar T, Snel B, Huynen M, Bork P.** 1998. Conservation of gene order: a fingerprint of proteins that physically interact. Trends Biochem. Sci. **23:**324–328. http://dx.doi.org/10.1016/S0968-0004(98)01274-2.
131. **Itoh T, Takemoto K, Mori H, Gojobori T.** 1999. Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes. Mol. Biol. Evol. **16:**332–346. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026114.
132. **Price MN, Arkin AP, Alm EJ.** 2006. The life-cycle of operons. PLoS Genet. **2:**e96. http://dx.doi.org/10.1371/journal.pgen.0020096.
133. **de Daruvar A, Collado-Vides J, Valencia A.** 2002. Analysis of the cellular functions of *Escherichia coli* operons and their conservation in *Bacillus subtilis*. J. Mol. Evol. **55:**211–221. http://dx.doi.org/10.1007/s00239-002-2317-1.
134. **Price MN, Huang KH, Arkin AP, Alm EJ.** 2005. Operon formation is driven by co-regulation and not by horizontal gene transfer. Genome Res. **15:**809–819. http://dx.doi.org/10.1101/gr.3368805.
135. **Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV.** 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29:**22–28. http://dx.doi.org/10.1093/nar/29.1.22.
136. **Struhl K.** 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell **98:**1–4. http://dx.doi.org/10.1016/S0092-8674(00)80599-1.
137. **Semsey S, Tolstorukov MY, Virnik K, Zhurkin VB, Adhya S.** 2004. DNA trajectory in the Gal repressosome. Genes Dev. **18:**1898–1907. http://dx.doi.org/10.1101/gad.1209404.
138. **Semsey S, Virnik K, Adhya S.** 2005. A gamut of loops: meandering DNA. Trends Biochem. Sci. **30:**334–341. http://dx.doi.org/10.1016/j.tibs.2005.04.009.
139. **Losada A, Hirano T.** 2001. Shaping the metaphase chromosome: coordination of cohesion and condensation. Bioessays **23:**924–935. http://dx.doi.org/10.1002/bies.1133.
140. **Mazia D.** 1987. The chromosome cycle and the centrosome cycle in the mitotic cycle. Int. Rev. Cytol. **100:**49–92. http://dx.doi.org/10.1016/S0074-7696(08)61698-8.
141. **Baserga R, Wiebel F.** 1969. The cell cycle of mammalian cells. Int. Rev. Exp. Pathol. **7:**1–30.
142. **Bates D, Epstein J, Boye E, Fahrner K, Berg H, Kleckner N.** 2005. The *Escherichia coli* baby cell column: a novel cell synchronization method provides new insight into the bacterial cell cycle. Mol. Microbiol. **57:**380–391. http://dx.doi.org/10.1111/j.1365-2958.2005.04693.x.
143. **Bernander R, Poplawski A.** 1997. Cell cycle characteristics of thermophilic archaea. J. Bacteriol. **179:**4963–4969.
144. **Murray AW.** 2004. Recycling the cell cycle: cyclins revisited. Cell **116:**221–234. http://dx.doi.org/10.1016/S0092-8674(03)01080-8.
145. **Blow JJ, Tanaka TU.** 2005. The chromosome cycle: coordinating replication and segregation. EMBO Rep. **6:**1028–1034. http://dx.doi.org/10.1038/sj.embor.7400557.
146. **Stillman B.** 2005. Origin recognition and the chromosome cycle. FEBS Lett. **579:**877–884. http://dx.doi.org/10.1016/j.febslet.2004.12.011.
147. **Dingman CW.** 1974. Bidirectional chromosome replication: some topological considerations. J. Theor. Biol. **43:**187–195. http://dx.doi.org/10.1016/S0022-5193(74)80052-4.
148. **Abner K, Aaviksaar T, Adamberg K, Vilu R.** 2014. Single-cell model of prokaryotic cell cycle. J. Theor. Biol. **341:**78–87.
149. **Pflumm MF.** 2002. The role of DNA replication in chromosome condensation. Bioessays **24:**411–418. http://dx.doi.org/10.1002/bies.10092.
150. **Stack SM, Anderson LK.** 2001. A model for chromosome structure during the mitotic and meiotic cell cycles. Chromosome Res. **9:**175–198. http://dx.doi.org/10.1023/A:1016690802570.
151. **Lopez P, Philippe H, Myllykallio H, Forterre P.** 1999. Identification of putative chromosomal origins of replication in Archaea. Mol. Microbiol. **32:**883–886. http://dx.doi.org/10.1046/j.1365-2958.1999.01370.x.
152. **Majerník AI, Lundgren M, McDermott P, Bernander R, Chong JP.** 2005. DNA content and nucleoid distribution in *Methanothermobacter thermautotrophicus*. J. Bacteriol. **187:**1856–1858. http://dx.doi.org/10.1128/JB.187.5.1856-1858.2005.
153. **Popławski A, Bernander R.** 1997. Nucleoid structure and distribution in thermophilic Archaea. J. Bacteriol. **179:**7625–7630.
154. **Robinson NP, Blood KA, McCallum SA, Edwards PA, Bell SD.** 2007.

Sister chromatid junctions in the hyperthermophilic archaeon *Sulfolobus solfataricus*. EMBO J. **26:**816–824. http://dx.doi.org/10.1038/sj.emboj.7601529.

155. **O'Donnell ME, Kornberg A.** 1985. Dynamics of DNA polymerase III holoenzyme of *Escherichia coli* in replication of a multiprimed template. J. Biol. Chem. **260:**12875–12883.
156. **Tanner NA, Loparo JJ, Hamdan SM, Jergic S, Dixon NE, van Oijen AM.** 2009. Real-time single-molecule observation of rolling-circle DNA replication. Nucleic Acid Res. **37:**e27. http://dx.doi.org/10.1093/nar/gkp006.
157. **Yao NY, Schroeder JW, Yurieva O, Simmons LA, O'Donnell ME.** 2013. Cost of rNTP/dNTP pool imbalance at the replication fork. Proc. Natl. Acad. Sci. U. S. A. **110:**12942–12947. http://dx.doi.org/10.1073/pnas.1309506110.
158. **Breier AM, Weier HU, Cozzarelli NR.** 2005. Independence of replisomes in *Escherichia coli* chromosomal replication. Proc. Natl. Acad. Sci. U. S. A. **102:**3942–3947. http://dx.doi.org/10.1073/pnas.0500812102.
159. **Kouzminova EA, Kuzminov A.** 2012. Chromosome demise in the wake of ligase-deficient replication. Mol. Microbiol. **84:**1079–1096. http://dx.doi.org/10.1111/j.1365-2958.2012.08076.x.
160. **Pham TM, Tan KW, Sakumura Y, Okumura K, Maki H, Akiyama MT.** 18 September 2013. A single-molecule approach to DNA replication in *Escherichia coli* cells demonstrated that DNA polymerase III is a major determinant of fork speed. Mol. Microbiol. http://dx.doi.org/10.1111/mmi.12386.
161. **Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, Lockhart DJ, Davis RW, Brewer BJ, Fangman WL.** 2001. Replication dynamics of the yeast genome. Science **294:**115–121. http://dx.doi.org/10.1126/science.294.5540.115.
162. **Sekedat MD, Fenyö D, Rogers RS, Tackett AJ, Aitchison JD, Chait BT.** 2010. GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. Mol. Syst. Biol. **6:**353. http://dx.doi.org/10.1038/msb.2010.8.
163. **Atlung T, Hansen FG.** 2002. Effect of different concentrations of H-NS protein on chromosome replication and the cell cycle in *Escherichia coli*. J. Bacteriol. **184:**1843–1850. http://dx.doi.org/10.1128/JB.184.7.1843-1850.2002.
164. **Von Freiesleben U, Rasmussen KV, Atlung T, Hansen FG.** 2000. Rifampicin-resistant initiation of chromosome replication from oriC in *ihf* mutants. Mol. Microbiol. **37:**1087–1093. http://dx.doi.org/10.1046/j.1365-2958.2000.02060.x.
165. **Morigen Løbner-Olesen A, Skarstad K.** 2003. Titration of the *Escherichia coli* DnaA protein to excess *datA* sites causes destabilization of replication forks, delayed replication initiation and delayed cell division. Mol. Microbiol. **50:**349–362. http://dx.doi.org/10.1046/j.1365-2958.2003.03695.x.
166. **Cooper S, Helmstetter CE.** 1968. Chromosome replication and the division cycle of *Escherichia coli* B/r. J. Mol. Biol. **31:**519–540. http://dx.doi.org/10.1016/0022-2836(68)90425-7.
167. **von Freiesleben U, Krekling MA, Hansen FG, Løbner-Olesen A.** 2000. The eclipse period of *Escherichia coli*. EMBO J. **19:**6240–6248. http://dx.doi.org/10.1093/emboj/19.22.6240.
168. **Howe CJ, Nisbet RE, Barbrook AC.** 2008. The remarkable chloroplast genome of dinoflagellates. J. Exp. Bot. **59:**1035–1045. http://dx.doi.org/10.1093/jxb/erm292.
169. **Bouet JY, Nordström K, Lane D.** 2007. Plasmid partition and incompatibility–the focus shifts. Mol. Microbiol. **65:**1405–1414. http://dx.doi.org/10.1111/j.1365-2958.2007.05882.x.
170. **Jensen MR, Løbner-Olesen A, Rasmussen KV.** 1990. *Escherichia coli* minichromosomes: random segregation and absence of copy number control. J. Mol. Biol. **215:**257–265. http://dx.doi.org/10.1016/S0022-2836(05)80344-4.
171. **Helmstetter CE, Leonard AC.** 1987. Coordinate initiation of chromosome and minichromosome replication in *Escherichia coli*. J. Bacteriol. **169:**3489–3494.
172. **Moreno Díaz de la Espina S, Alverca E, Cuadrado A, Franca S.** 2005. Organization of the genome and gene expression in a nuclear environment lacking histones and nucleosomes: the amazing dinoflagellates. Eur. J. Cell Biol. **84:**137–149. http://dx.doi.org/10.1016/j.ejcb.2005.01.002.
173. **Costas E, Goyanes V.** 2005. Architecture and evolution of dinoflagellate chromosomes: an enigmatic origin. Cytogenet. Genome Res. **109:**268–275. http://dx.doi.org/10.1159/000082409.
174. **Wang X, Lesterlin C, Reyes-Lamothe R, Ball G, Sherratt DJ.** 2011. Replication and segregation of an *Escherichia coli* chromosome with two replication origins. Proc. Natl. Acad. Sci. U. S. A. **108:**E243–E250. http://dx.doi.org/10.1073/pnas.1100874108.

JB Journals.ASM.org

Fifty Years after the Replicon Hypothesis: Cell-Specific Master Regulators as New Players in Chromosome Replication Control

Marcin Wolański,[a] Dagmara Jakimowicz,[a,b] Jolanta Zakrzewska-Czerwińska[a,b]

Faculty of Biotechnology, University of Wrocław, Wrocław, Poland[a]; Department of Microbiology, Ludwik Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland[b]

Numerous free-living bacteria undergo complex differentiation in response to unfavorable environmental conditions or as part of their natural cell cycle. Developmental programs require the *de novo* expression of several sets of genes responsible for morphological, physiological, and metabolic changes, such as spore/endospore formation, the generation of flagella, and the synthesis of antibiotics. Notably, the frequency of chromosomal replication initiation events must also be adjusted with respect to the developmental stage in order to ensure that each nascent cell receives a single copy of the chromosomal DNA. In this review, we focus on the master transcriptional factors, Spo0A, CtrA, and AdpA, which coordinate developmental program and which were recently demonstrated to control chromosome replication. We summarize the current state of knowledge on the role of these developmental regulators in synchronizing the replication with cell differentiation in *Bacillus subtilis*, *Caulobacter crescentus*, and *Streptomyces coelicolor*, respectively.

Faithful transmission of genetic material to daughter cells requires the precise regulation of chromosomal replication and its coordination with the cell cycle. In all three domains of life, chromosomal replication is mainly regulated at the initiation step, an important cell cycle checkpoint designed to guarantee that chromosomal replication occurs only once per cell division cycle. Over 50 years ago, Sydney Brenner, Francois Cuzin, and Francois Jacob proposed the replicon theory (1) explaining how chromosome replication in bacteria is coordinated with the cell cycle and cell division. The theory assumed the existence of only two key elements required for initiation of bacterial chromosome replication: the *trans*-acting element (a structural gene encoding an initiator) and the *cis*-acting element (replicator). According to their hypothesis, the initiator positively regulates replication initiation by interaction with the replicator. Twenty to 30 years later, the theory was experimentally proven not only for bacteria, phages, and plasmids but also for archaea and, to some extent, for eukaryotes. Over the last 20 years, researchers have made considerable progress in understanding the mechanisms of replication initiation, particularly the structures and functions of the key elements across the three domains of life.

In contrast to *Eukaryota*, bacterial chromosomes possess a single, unique replication origin, in which the DNA synthesis starts generating a single replication eye per chromosome (2, 3). Replication of the bacterial chromosome is initiated by the binding of the initiator protein called DnaA to specific 9-mer sequences (called DnaA boxes) within the *oriC* region (origin of chromosomal replication—the replicator) (originally described by the Kornberg group; see, e.g., references 4 and 5). The DnaA protein is composed of four functional domains that are responsible for self-oligomerization and interactions with other proteins (e.g., DnaB and HU), cofactors (ATP/ADP), and DNA (i.e., DnaA boxes) (see Table 1) (6–11). Bacterial *oriC* regions differ in size (from ~200 to 1,000 bp), and their sequences are conserved only among closely related organisms. In addition to the repertoire of DnaA boxes, they usually include an AT-rich region with a DNA-unwinding element (DUE) and the binding sites for regulatory proteins. Through the sequential binding of high-, medium-, and low-affinity DnaA boxes, the DnaA protein forms a highly ordered nucleoprotein complex (orisome), which promotes separation of DNA strands at the DUE. Notably, only ATP-bound DnaA is able to unwind DNA. Furthermore, ATP-bound DnaA is also required for the binding of low-affinity DnaA boxes (e.g., those in *Escherichia coli* and *Streptomyces*) (12–15). The unwound region provides an entry site for the key replication enzymes helicase, primase, and DNA polymerase (Pol) III, the latter of which forms the replication fork (for a review, see reference 16).

Bacterial replication begins at a single *oriC* region, and the activity and availability of this region and the initiator protein have to be tightly regulated to ensure that chromosomal DNA is entirely replicated only once per cell cycle. Among several mechanisms involved in regulating replication initiation, the inactivation of DnaA-ATP by ATP hydrolysis is likely to be the most common in bacteria (10), whereas other mechanisms appear to be specific for particular bacteria. These include the sequestration of the hemimethylated *oriC* region seen in *E. coli*; the titration of DnaA proteins by clusters of DnaA boxes localized outside the *oriC* regions of *E. coli*, *Bacillus subtilis*, and *Streptomyces coelicolor*; the modulation of DnaA activity by different proteins, such as Hda homologues in *E. coli* and *Caulobacter crescentus* and SirA and YabA in *B. subtilis*; and the proteolytic degradation of DnaA by the Lon and ClpP proteases in *C. crescentus* (Table 2) (for reviews or details, see references 7, 10, 17, 18, 19, 20, 21, 22, and 23). Recently, master transcriptional regulators were demonstrated to be involved in controlling chromosome replication as well as other cellular processes. Examples of these are CtrA, Spo0A, and AdpA proteins, which, in *C. crescentus*, *B. subtilis*, and *S. coelicolor*, re-

Published ahead of print 9 June 2014

Address correspondence to Jolanta Zakrzewska-Czerwińska, jolanta.zakrzewska@uni.wroc.pl.

This article is dedicated to the memory of Walter Messer.

doi:10.1128/JB.01706-14

TABLE 1 DnaA—replication initiator protein[a]

Microorganism	DnaA protein features
Bacillus subtilis	50.8 kDa; 5′-TTATCCACA-3′—consensus binding sequence; regulon contains a minimum of 52 genes (including the *dnaA* gene)
Caulobacter crescentus	53.9 kDa; 5′-TGATCCACA-3′—G box consensus binding sequence; 5′-NNNTCCCCA-3′—W box consensus sequence; regulon contains a minimum of 14 genes
Streptomyces coelicolor	73.2 kDa; 5′-TTGTCCACA-3′—consensus binding sequence; regulon not yet determined

[a] DnaA is a highly conserved eubacterial protein that specifically recognizes several 9-mer sequences (DnaA boxes) within the origin of chromosome replication (*oriC*) and oligomerizes along them to facilitate local DNA unwinding. DnaA belongs to the AAA$^+$ (ATPases associated with diverse cellular activities) protein family and requires ATP to form the initiation-active nucleoprotein oligomer. It consists of four domains (I to IV) that contain functional regions responsible for interactions with protein partners (I and III), self-oligomerization (I and III), ATP binding and hydrolysis (III), and the DNA-binding HTH motif (IV).

spectively, play crucial roles in coordination of replication with differentiation of these bacteria (Table 3).

In bacteria that undergo differentiation, the regulatory networks that control replication initiation are likely to be intricate and require specific mechanisms that can synchronize chromosomal replication initiation with developmental processes, which involve functional specialization of specific cell types in response to environmental signals (e.g., nutrient limitations). In this review, we summarize the current state of knowledge regarding the checkpoints that couple replication to cell cycle progression. We focus on the master transcription factors, CtrA, Spo0A, and AdpA (Table 3), which control the morphological developments of *C. crescentus*, *B. subtilis*, and *S. coelicolor*, respectively, and discuss how these regulators temporally and spatially coordinate replication initiation with cell differentiation.

MASTER TRANSCRIPTION FACTORS CONTROL CELL CYCLE PROGRESSION IN BACTERIA THAT UNDERGO COMPLEX DEVELOPMENTAL CYCLES

Free-living bacteria develop specific strategies to survive unfavorable environmental conditions. Such adaptations may be achieved in certain bacteria through the acquisition of a complex life cycle that includes the formation of spores or dormant cells. A complex bacterial life cycle requires the specific coordination of cell division, chromosomal replication, segregation, and morphological differentiation (Fig. 1). Differentiation includes morphological changes that depend on the activation of numerous genes, such as those responsible for spore coat synthesis or flagellar assembly. Thus, complex developmental pathways require the presence of specific regulatory proteins capable of coordinating multiple processes, including chromosomal replication. This often arises via the specific spatial, temporal, and asymmetrical (depending on the cell type/cell compartment) regulation of gene expression and protein accumulation controlled by master regulators. Among the master regulators identified in *Bacillus*, *Streptomyces*, and *Caulobacter*, Spo0A, AdpA, and CtrA, respectively (see Table 3), have been shown to coordinate differentiation with the regulation of chromosomal replication.

TABLE 2 Negative-regulatory mechanisms for DnaA-dependent replication initiation

Bacterial species and element	Mechanism	Reference
Bacillus subtilis		
Spo0A	Directly binds to *oriC* and inhibits DnaA binding	74
YabA	Inhibits DnaA helix formation	19
SirA	Inhibits the binding of DnaA to *oriC*	20
Soj monomer	Inhibits DnaA helix formation	109
Caulobacter crescentus		
CtrA	Directly binds to *oriC* and inhibits DnaA binding	77
HdaA	Inactivates DnaA in the replisome	21
ClpP	Degrades DnaA	18
Lon	Degrades DnaA	17
GcrA	Inhibits the expression of *dnaA*	72
Streptomyces coelicolor		
AdpA	Directly binds to *oriC* and inhibits DnaA binding	15
D78 cluster	Titrates DnaA from *oriC*	23

Sporulation, which is a survival strategy in such diverse genera as *Bacillus*, *Clostridium*, *Myxococcus*, and *Streptomyces*, is triggered by nutrient limitation and/or high cell density (24–28). Spore formation may follow very diverse routes depending on the genus. During the sporulation of *B. subtilis* (Fig. 1), for example, formation of a single endospore starts with an asymmetric cell division that delimits a forespore, followed by engulfment of the forespore in a phagocytic-like process, the formation of a protective spore coat, and, subsequently, lysis of the mother cell to release the spore (reviewed in references 29, 30, and 31). In contrast, mycelial soil-inhabiting *Streptomyces* spp. form the chains of exospores (Fig. 1) (32). Here, the transition from vegetative growth to sporulation involves the differentiation of the branched network of vegetative hyphae into morphologically distinct aerial hyphae, which are further converted into chains of prespores via synchronized multiple-cell division (33). Thereafter, the spores mature via synthesis of a thick cell wall and eventually undergo dispersal. Unlike bacteria in which sporulation is an optional developmental pathway, the water-inhabiting *Caulobacter* sp. undergoes morphological differentiation as an intrinsic part of its natural cell cycle (34). Asymmetric division of a *C. crescentus* cell produces two daughter cells with different morphologies: a nonmotile stalked cell able to adhere to surfaces and a motile swarmer cell with a polar flagellum (Fig. 1). Only the stalked cell is able to divide; when the swarmer cell is provided with sufficient nutrients, it differentiates into a stalked cell (via ejection of the flagellum and formation of a stalk at the same pole) and undergoes division.

Normally, before initiation of sporulation of *B. subtilis*, ongoing chromosomal replication is terminated (Sda protein is a factor

TABLE 3 Global transcription regulators

Transcription regulator—bacterial species	Features
Spo0A—*B. subtilis*	29.7 kDa; N-terminal phosphoacceptor-dimerization domain; C-terminal HTH DNA-binding motif and transcription activation domain; Spo0A consensus-binding sequence (0A-box), 5′-TGTCGAA-3′; essential for the activation of sporulation genes; responsible for regulation of over 500 genes and for direct regulation of 120 genes; response regulator of the two-component signal transduction pathway type, activated by Spo0B-mediated phosphorylation
CtrA—*C. crescentus*	25.8 kDa; N-terminal phosphoacceptor-dimerization domain; C-terminal HTH DNA-binding and effector domain; CtrA-binding sequence, 5′-TTAA-N7-TTAA-3′; responsible for the direct regulation of nearly 100 genes, including those responsible for the biogenesis of flagella and pili and for cell division; response regulator of the two-component signal transduction pathway type, activated by ChpT-mediated phosphorylation
AdpA—*S. coelicolor*	42.8 kDa; N-terminal dimerization domain; C-terminal DNA-binding domain containing a double HTH motif; member of the AraC/XylS family of transcriptional regulators; AdpA consensus-binding sequence, 5′-TGGCSNGW WY-3′; responsible for activating the production of over 40 proteins, including those essential for morphological development (e.g., proteases, protease inhibitors, and sigma factors); does not appear to be subject to phosphorylation

that prevents entry into sporulation when cells are initiating DNA replication [35]) and the newly replicated chromosomes are remodeled into an axial filament extending between the poles (36, 37). Asymmetrical positioning of the septum, which is typical of sporulation, traps the forespore chromosome within the closing septum and generates a transient asymmetry wherein the mother cell is diploid for two-thirds of the chromosomal genes. Completion of chromosome segregation requires active translocation of the chromosome through the closing septum (38).

Similarly to the situation in *B. subtilis*, the sporulation-related cell division of *Streptomyces* employs sporulation-specific control of the typical vegetative cell division machinery. During vegetative growth, occasional septa delimit adjacent elongated multinucleoid hyphal compartments; during sporulation, in contrast, the closely spaced sporulation septa form a ladder along the sporogenic multinucleoid compartment (39, 40) (Fig. 1). Formation of the long chains of spores is preceded by rapid extension of sporogenic hyphae accompanied by intensive chromosome replication, after which the ongoing rounds of replication are terminated and the chromosomes are condensed and precisely positioned between developing septa to form unigenomic spores (41, 42). In sporulating *Streptomyces*, therefore, cell division is tightly coordinated with the shutdown of chromosomal replication and the sporulation-specific segregation and organization of the chromosomes.

Precise synchronization of chromosomal replication with the cell cycle is also seen in *C. crescentus* (Fig. 1) (for recent reviews, see references 22 and 43). The newly replicated chromosome is moved to the nascent half of the cell immediately after initiation of replication in stalked cells (44), and after cell division, each daughter cell contains a single chromosome that is attached to the cell pole by a polarly localized replication origin region (45, 46). As noted above, chromosomal replication is blocked in swarmer cells until they transform into stalked cells. Septation is coordinated with replication initiation and the bipolar positioning of the two sister origin regions (47).

In *B. subtilis*, the sporulation "decision" depends on the master transcription factor, Spo0A (Table 3). A combination of signals associated with high population density and cell cycle progression accelerates a multicomponent phosphorelay cascade, resulting in the phosphorylation of Spo0A (48). Upon phosphorylation, Spo0A initiates the sporulation pathway by affecting the transcription of ~500 genes (~120 of which are directly controlled). The binding of Spo0A~P to a specific DNA sequence called the Spo0A-box in promoter regions activates the genes involved in sporulation and represses many genes that had been expressed during vegetative growth (35, 49, 50). Spo0A~P is responsible for activating mother cell/forespore-specific sigma factors (and other regulators) and genes whose products control asymmetric division (51–53) and sporulation-specific chromosomal segregation (36, 37). Interestingly, sporulation is triggered only when a high level of phosphorylated Spo0A is attained, whereas the intermediate levels of Spo0A~P in the so-called "transition state" are associated with increased protease production, motility, competence for transformation, biofilm formation, and cannibalism, in a hierarchy of developmental decision-making (54).

Similarly to the case of *B. subtilis*, the sporulation of *S. coelicolor* (a model organism for developmental studies in *Streptomyces*) is governed by numerous transcriptional regulators, including sigma factors (55). The set of regulators responsible for controlling the erection of aerial hyphae is called Bld due to the "bald" colony phenotype caused by mutations in their encoding genes ("bald" colonies do not produce the outer fluffy layer—the aerial hyphae) (56, 57). The transition from fluffy white aerial hyphae to chains of gray spores requires the activation of regulatory proteins encoded by the *whi* genes (so named for the white appearance of mutant colonies) (58, 59). As the development of *Streptomyces* colonies is correlated with secondary metabolism and antibiotic production, mutations in *bld* genes can affect both morphological differentiation and the production of antibiotics (reviewed in reference 60). The product of the *adpA* gene (also known as *bldH*) (Table 3) plays a key coordinating role in regulating the morphological differentiation of *Streptomyces*. AdpA is a master transcription factor responsible for controlling several dozen genes (61, 62) whose encoded products are required for morphological development (e.g., SapB, a regulator of the aerial mycelium-promoting peptide, and SsgA, a protein involved in septum formation during sporulation [63]; reviewed by Higo et al. [64]). *In silico* analysis has identified AdpA-binding sequences in more than 150 intergenic regions of *S. coelicolor* (63).

As in sporulating bacteria, control of development in *C. cres-*

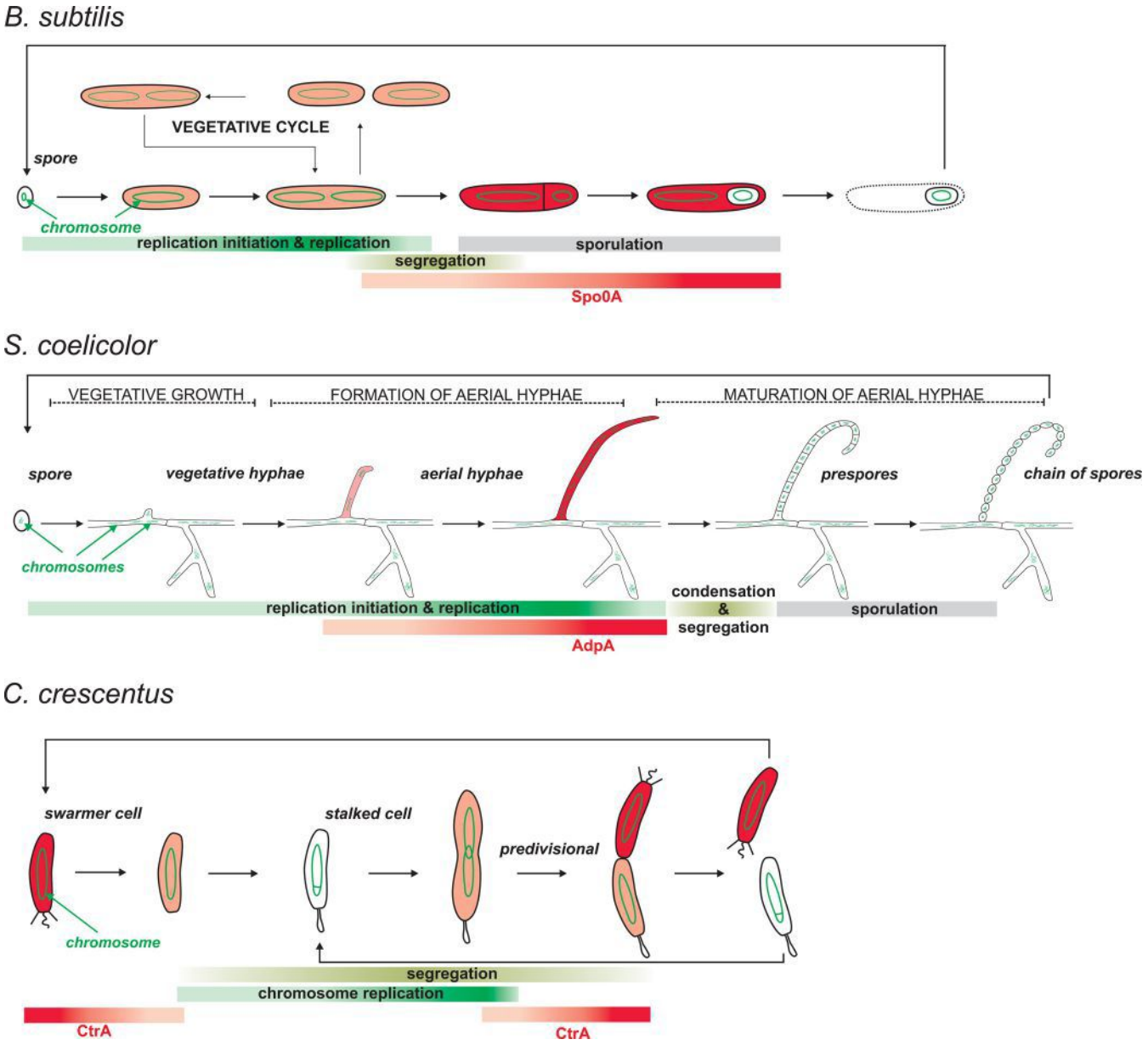


FIG 1 Chromosome distributions during the developmental life cycles of *B. subtilis*, *S. coelicolor*, and *C. crescentus*. Darker colors indicate intensive chromosomal replication or higher protein concentrations of the master regulators, Spo0A, AdpA, and CtrA.

centus is associated with the spatiotemporal control of numerous genes (about 20% of all genes) (65). Many of these are controlled by CtrA (Table 3) and three other master regulators, GcrA (a DNA binding protein), DnaA, and CcrM (a DNA methylase), that are regulated at one or several levels (synthesis, activation, localization, degradation). Together, these four proteins precisely coordinate chromosomal replication and segregation with cell growth, cell division, and pole morphogenesis (22, 66, 67). CtrA recognizes and binds a specific DNA sequence (the CtrA box) to directly control the transcription of ~100 genes within 55 operons affecting the expression of genes involved in flagellum assembly, holdfast synthesis, DNA methylation, and cell division (68). Activation of CtrA is controlled by its phosphorylation, in which a regulatory circuit of several polarly localized histidine kinases and phosphatases is involved (69, 70). Due to the polarity of *C. crescentus* predivisional cells, the swarmer cell inherits an active, phosphorylated form of CtrA whereas CtrA undergoes specific proteolysis in the stalked cell (71). When a swarmer cell transitions to a stalked cell, CtrA is inactivated and the *gcrA* gene, which encodes the stalked-cell-specific master regulator, is induced. GcrA affects the transcription of over 125 genes, including those encoding components of the DNA replication and chromosomal segregation machineries (72, 73). Upon initiation of cell division, clearance of GcrA is coincident with the gradual accumulation of CtrA (72).

In sum, the master regulators play crucial roles in the spatial and temporal regulation of sequential events during bacterial differentiation. Their abundance and activity are precisely controlled at different levels, including gene expression (at transcriptional

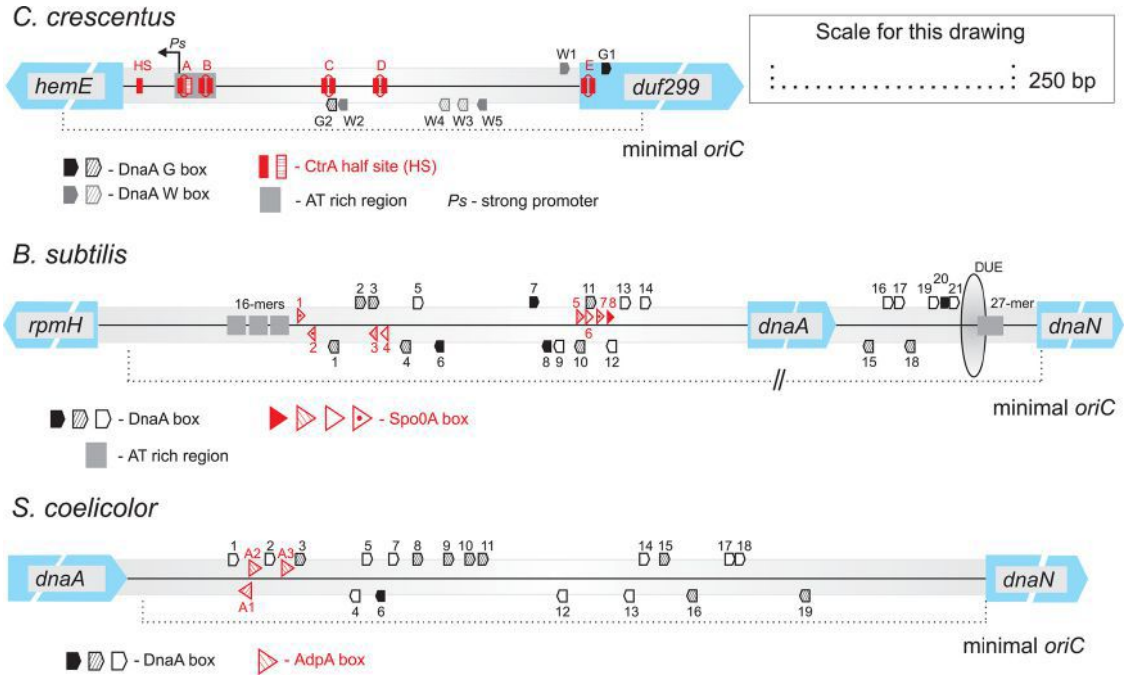


FIG 2 Organization of the origin regions (*oriC*) from *C. crescentus*, *B. subtilis*, and *S. coelicolor*. Positions of the binding sites of master regulators CtrA, Spo0A, AdpA (red), and DnaA (black) are shown. Filled, stripped, empty, and dotted boxes (in that order) represent increasing deviations from the consensus sequence. Gray rectangles indicate AT-rich sequences; the oval DUE represents the DNA unwinding element; Ps represents the transcription start site from the strong *hemE* promoter. The DUE for *C. crescentus* and *S. coelicolor* origins has not been defined.

and translational levels), proteolysis, and posttranslational modification (phosphorylation).

MASTER REGULATORS INHIBIT REPLICATION INITIATION DURING THE TRANSITION TO NONREPLICATING CELLS

Despite extensive studies on the mechanisms responsible for controlling the initiation of chromosomal replication (for reviews, see references 7, 9, and 10), we have begun only recently to understand how the master regulators coordinate differentiation with chromosomal replication. The initiation of chromosomal replication requires the ATP-bound form of DnaA (Table 1) and many other factors that organize and/or regulate initiation complex (orisome) formation, ensuring that DNA unwinding occurs once per cell cycle. AdpA, Spo0A, and CtrA (Table 3) have been identified as factors that coordinate replication with the cell cycle; they inhibit the initiation of chromosomal replication by binding to the cognate *oriC* region in *S. coelicolor* and *B. subtilis* (AdpA and Spo0A, respectively) and to the region called *Cori* in *C. crescentus* (CtrA) (15, 74–76). Compared with *E. coli* (14), the origin regions of *B. subtilis*, *C. crescentus*, and *S. coelicolor* are structurally more complex and relatively long (1,000 bp or longer) and contain multiple high-, moderate-, and low-affinity DnaA boxes (Fig. 2) (74, 77, 78). In the *oriC* regions of *S. coelicolor* and *B. subtilis*, the DnaA boxes are organized in two clusters (boxes 1 to 11 and boxes 13 to 19 in *S. coelicolor* and boxes separated by the *dnaA* gene in *B. subtilis*), both of which are required for autonomous replication (79, 80). *In vitro* observations indicate that upon binding of the DnaA protein, a loop is formed between the clusters; this mechanism may promote the *in vivo* unwinding of DNA at the DUE (81, 82). Interestingly, the *oriC* region of each three model organisms possesses multiple binding sites for their respective master regulators, AdpA, CtrA, and Spo0A, located near or partially overlapping the DnaA-binding sites that are indispensable for replication initiation (Fig. 2) (15, 74, 77, 79, 83). Moreover, the specific interactions between these master regulators and their *oriC* regions were confirmed by *in vitro* (electrophoretic mobility shift assay [EMSA], DNase I footprinting, or surface plasmon resonance) and *in vivo* (chromatin immunoprecipitation [ChIP]) assays in all three model organisms (15, 49, 68, 84). The interaction of master regulators with their targets presumably interferes with the formation of functional orisomes, thereby preventing replication initiation. Indeed, *in vivo* experiments demonstrated that *B. subtilis*, *C. crescentus*, and *S. coelicolor* with *oriC* regions lacking the Spo0A-, CtrA-, and AdpA-binding sequences, respectively, exhibited chromosome overreplication (15, 76, 85). The master regulators can probably displace DnaA from *oriC* (as shown for CtrA [77]) or prevent DnaA from oligomerizing at *oriC* and thus block the formation of a functional orisome by limiting DnaA access to the replication origin. Thus, the master regulators and initiator proteins presumably compete for binding to a particular part(s) of the *oriC* region to inhibit and promote, respectively, the initiation of replication. The balance of this "battle" depends on the cell cycle stage, since both elements are temporally and spatially regulated.

In *C. crescentus*, the levels of CtrA and DnaA oscillate during the cell cycle, peaking in swarmer and stalked cells, respectively (18). Replication initiation occurs in stalked cells or in the stalked compartment of late predivisional cells, whereas it is inhibited by the activated master regulator, CtrA~P, in swarmer cells (84; for a review, see also reference 22). Interestingly, besides being involved in DnaA displacement from the origin of replication region (77), CtrA also negatively regulates the strong *hemE* promoter (located at the 5′ end of the origin; Fig. 2), the activity of which is essential

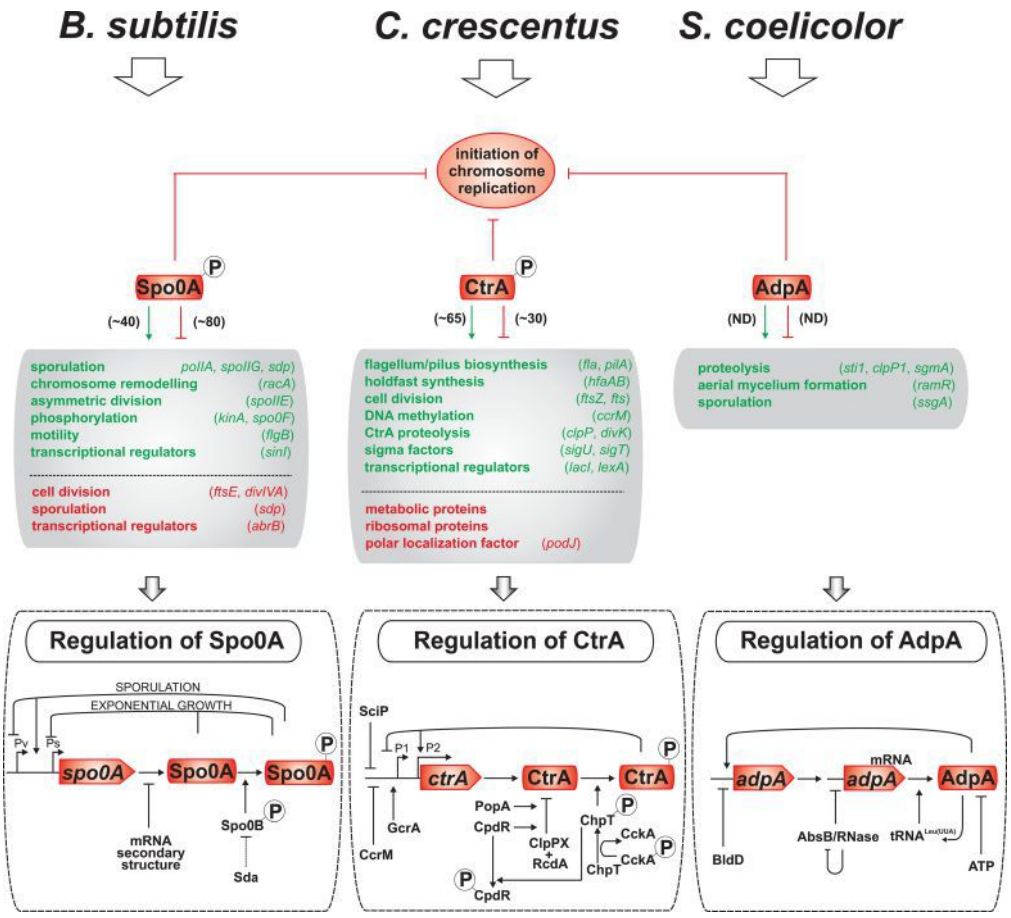


FIG 3 Master regulators coordinate the cell cycle with DNA replication. Examples of genes involved in cell cycle developmental processes and replication, which are regulated by Spo0A, CtrA, and AdpA proteins, respectively, are listed in boxes (for full lists of genes regulated by these transcription factors, see references 49, 50, 63, 68, and 118). The dotted line indicates an indirect multilevel interaction.

for the initiation of DNA replication (86). It is assumed that transcription from this promoter allows formation of a RNA-DNA hybrid inside the origin of replication and thus promotes the melting of DNA strands. CtrA is spatiotemporally regulated by phosphorylation (by the CckA kinase [69]—where ChpT phosphotransferase is a direct donor of a phosphate group for CtrA [87]) and by proteolysis (by the ClpXP protease [88]—where the RcdA protein assists ClpXP-dependent CtrA degradation at the cell pole while another PopA recruits both RcdA and CtrA to the cell pole [89, 90]). The other protein, CpdR, in its unphosphorylated form binds to ClpXP and recruits this proteolytic complex to the cell pole, leading to CtrA degradation in stalked cells. CpdR is subject to phosphorylation by the CckA/ChpT proteins (91).

Moreover, the *ctrA* gene expression level is also spatiotemporally regulated, with maximum and minimum transcriptional activity seen in predivisional and swarmer cells, respectively (92). The transcription of *ctrA* starts from two promoters, P1 and P2, one of which (P1) is regulated by the orchestrated actions of three master proteins, CcrM, GcrA, and its own product, CtrA (Fig. 3). Interestingly, both promoter P1 and promoter P2 are subject, respectively, to negative and positive feedback control by CtrA~P (92). CcrM inhibits transcription from P1, whereas GcrA activates transcription from P1 (72, 93). Recent studies showed that P1 activity could also be modulated (negatively) by another transcription factor, SciP, which is a CtrA antagonist responsible for repression of genes controlled by CtrA in the swarmer cell (94, 95). Thus, the presence of two promoters, whose activities are regulated in a sophisticated manner, allows *C. crescentus* to regulate precisely in time and space the level of CtrA during the cell cycle. CcrM and GcrA also temporally regulate the expression of *dnaA* during the cell cycle but with effects that are opposite their effects on the *ctrA* P1 promoter (72, 96). The DnaA protein of *C. crescentus* is subject to degradation in a cell cycle-dependent manner (first reported by Gorbatyuk and Marczynski [18]). Interestingly, Laub's group recently showed that the stability of DnaA under certain stress conditions is mainly controlled by the Lon protease, although they did not exclude a minor role (or a role under alternative conditions) for the ClpP protease (17). Additionally, DnaA activity in *C. crescentus* is negatively controlled by HdaA (Hda homologue from *E. coli*), as cells depleted of this protein often overinitiate DNA replication (21). The HdaA protein, in complex with the β-clamp, probably contributes to the RIDA (regulatory inactivation of DnaA) mechanism that stimulates the ATPase activity of DnaA protein and leads to initiator inactivation (97). Interestingly, DnaA directly stimulates gene expression of its negative regulator, HdaA (21). It is worth mentioning that, during the swarmer-cell-to-stalked-cell transition, DnaA accumulation (accompanied by CtrA degradation) leads to activation of *gcrA* tran-

scription and subsequent CtrA expression induced by the GcrA cell cycle regulator (98). Interestingly, though there are also CcrM methylation sites within the *C. crescentus* origin, it seems that methylation of the origin does not play a role in regulating the initiation of chromosomal replication as it does with Dam methylase (99, 100).

During sporulation of *B. subtilis*, binding of the activated master regulator, Spo0A~P, to *oriC* prevents replication initiation, possibly by directly impeding the interaction of DnaA with *oriC* or by altering the ability of DnaA to form the orisome (76). *In vitro* experiments demonstrated that Spo0A prevents DnaA-dependent unwinding within the *oriC* region (74), while an *in vivo* analysis showed that growth and chromosomal replication were inhibited in a Spo0A-dependent manner in mother cells (101). Spo0A is regulated by multiple kinases via the multicomponent phosphorelay (48) where the Spo0B transferase is a direct donor of phosphate group for Spo0A (Fig. 3). The Spo0A activity is also temporally controlled on the transcriptional level by two promoters: a vegetative-state-related σ^A-recognized promoter (Pv) and a sporulation-related σ^H-recognized promoter (Ps) (102). The σ^A housekeeping sigma factor controls a weak Pv promoter, which provides a basal level of Spo0A that is required for efficient firing of a strong Ps promoter (Spo0A~P and σ^H activated) during the transition to the stationary phase (103–105). *spo0A* transcription is autoregulated by Spo0A itself in a manner similar to (but more complex than) that seen with *ctrA* in *C. crescentus*. The *spo0A* regulatory region contains multiple binding sites for Spo0A~P; one is responsible for repressing Pv during the transition to the stationary phase, while the others are responsible for repressing Ps during growth or activating Ps upon entry into sporulation. Moreover, the translation of Pv-originating mRNA is impeded by the presence of a secondary RNA structure, thereby decreasing the synthesis of Spo0A (103). This intricate mechanism for regulating *spo0A* ensures that high levels of Spo0A are appropriately available for the activation of key sporulation genes (50, 106). In contrast to the case in *C. crescentus*, the transcription of *spo0A* and *dnaA* in *B. subtilis* does not depend on methylation, as there are no *dam*, *ccrM*, or *seqA* homologues. Transcription of *dnaA* is also autoregulated, as binding of DnaA to the DnaA boxes within the promoter region represses the transcription of *dnaA* (107). The activity of *B. subtilis* DnaA is negatively regulated by SirA, YabA, DnaD, and monomeric Soj, which directly interact with DnaA to prevent its assembly on the *oriC* region (19, 20, 108–110). Additionally, DNA-bound Soj promotes replication initiation, presumably by stimulating DnaA protein helix assembly. As a ParA ortholog involved in chromosomal segregation, Soj enables the coordination of chromosomal replication and segregation in *B. subtilis* (109). Two other proteins, YabA and DnaD, have been also shown recently to inhibit DnaA helix formation during the initiation of replication, suggesting that oligomerization of the initiator protein is a strictly regulated step in *B. subtilis* (19). Additionally, the SirA sporulation protein negatively influences the activity of the DnaA protein by inhibiting the initiator binding to (or the removal of the initiator from) *oriC* (20). Expression of the *sirA* gene is positively regulated by the Spo0A regulator (49, 50). Conversely, the sporulation is blocked by DnaA-induced Sda protein when the replication is impaired. Sda acts by blocking one of the histidine kinases involved in activation of Spo0A (35, 111).

In contrast to the results from *B. subtilis* and *C. crescentus*, phosphorylation of the *S. coelicolor* master regulator, AdpA, has not been observed to date. Interestingly, AdpA competes with DnaA for binding of the *oriC* region, with ATP acting as a key regulator of the DNA-binding activities of both proteins (15). ATP is well known to strengthen the binding of DnaA proteins to DnaA boxes (particularly medium- and low-affinity boxes) and is also required for DNA unwinding. In contrast, the interaction between AdpA and *oriC* is profoundly weakened by ATP (15). The mechanism responsible for this phenomenon has not yet been elucidated. When AdpA reaches its highest level (which presages the emergence of aerial branches) and the ATP level is low, DnaA is not able to efficiently bind the *oriC* region and initiate replication (15, 112). Similarly to the cases of *spo0A* and *ctrA*, the expression of *adpA* is subject to regulation at different levels (Fig. 3). The abundance of the *adpA* transcript is modulated by AbsB/RNase III cleavage (113), and AdpA itself positively stimulates the transcription of *adpA* (63). Additionally, another transcriptional regulator, the BldD protein, encoded by one of the "bald" genes is involved in negative regulation of *adpA* transcription (114). Moreover, translation of *adpA* mRNA depends on the presence of a leucyl-tRNA (encoded by *bldA*) for a rarely used TTA codon (115, 116). Interestingly, the accumulation of this tRNA species coincides with aerial mycelium development (117). In a mutual feed-forward mechanism, AdpA directly activates *bldA* transcription (118). Thus, these mechanisms collectively ensure that AdpA protein levels peak at the proper time during colony development (Fig. 1 and 3). As in *B. subtilis*, the expression of *S. coelicolor dnaA* is negatively autoregulated, with two DnaA-binding sites located within the *dnaA* promoter region (119). Despite extensive studies, we have not yet identified any protein responsible for modulating *S. coelicolor* DnaA activity. Thus, DnaA activity is thought to be strictly dependent on its nucleotide-bound state, as the ATP-bound form of DnaA is required for DNA unwinding (our unpublished results) and the binding of low- and moderate-affinity DnaA boxes (15, 120).

The protein levels and activities of DnaA and the master regulators are subject to stringent and complex regulation during the cell cycle to ensure their effective and coordinated functioning (Fig. 3). Although the mechanistic principles of CtrA, Spo0A, and AdpA action may differ, since they originate from very diverse organisms, these proteins similarly regulate replication by binding *oriC* and negatively regulating initiation (Fig. 3). CtrA, Spo0A, and AdpA regulators are conserved in the alphaproteobacteria, endospore-forming bacilli, and *Streptomyces*, respectively, justifying adoption of *C. crescentus*, *B. subtilis*, and *S. coelicolor* as model organisms to study these regulators. In addition, although replication initiation has been extensively studied, we do not yet fully understand its regulation in microorganisms that undergo complex life cycles, and the participation of as-yet-unknown factors in this process cannot be excluded.

SUMMARY AND PERSPECTIVES

In all three domains of life, the synthesis of genetic material must be tightly regulated because under- or overreplication frequently leads to serious aberrations and/or genomic instability. In organisms that undergo complex life cycles, diverse checkpoint mechanisms have evolved to coordinate DNA replication with their sophisticated differentiation programs. Recent advances in cell biology have increased our understanding of how replication is regulated at the different steps of the cell cycle. Initiation of chromosomal replication is an essential checkpoint that seems to be

common to all domains of life. In bacteria undergoing the transition to a dormant state, master transcription factors (e.g., Spo0A, CtrA, and AdpA) regulate the expression levels of dozens of genes involved in morphological differentiation and inhibit replication initiation by binding to the origin of replication. Analyses of *oriC* regions and their affinity toward the master regulators, DnaA, and other regulatory proteins suggest that the origin regions have evolved to coordinate chromosomal replication with the complex developmental cell cycles of certain bacteria. In the future, an improved understanding of replication regulation could facilitate the experimental control of bacterial replication, potentially allowing us to inhibit DNA replication in pathogens, optimize the production of valuable secondary metabolites (e.g., antibiotics), and/or generate synchronized cultures for various physiological and genetic studies.

ACKNOWLEDGMENTS

We are grateful to Keith Chater and Patrick Viollier for providing valuable comments on the manuscript.

This work was supported by the National Science Centre, Poland (Maestro, grant no. 2012/04/A/NZ1/00057).

REFERENCES

1. **Jacob F, Brenner S, Cuzin F.** 1963. On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp. Quant. Biol. **28:**329–348. http://dx.doi.org/10.1101/SQB.1963.028.01.048.
2. **Kuzminov A.** 2014. The precarious prokaryotic chromosome. J. Bacteriol. **196:**1793–1806. http://dx.doi.org/10.1128/JB.00022-14.
3. **Leonard AC, Mechali M.** 2013. DNA replication origins. Cold Spring Harb. Perspect. Biol. **5:**a010116. http://dx.doi.org/10.1101/cshperspect.a010116.
4. **Fuller RS, Funnell BE, Kornberg A.** 1984. The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites. Cell **38:**889–900. http://dx.doi.org/10.1016/0092-8674(84)90284-8.
5. **Hwang DS, Kornberg A.** 1992. Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. J. Biol. Chem. **267:**23083–23086.
6. **Kaguni JM.** 2006. DnaA: controlling the initiation of bacterial DNA replication and more. Annu. Rev. Microbiol. **60:**351–375. http://dx.doi.org/10.1146/annurev.micro.60.080805.142111.
7. **Katayama T, Ozaki S, Keyamura K, Fujimitsu K.** 2010. Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. Nat. Rev. Microbiol. **8:**163–170. http://dx.doi.org/10.1038/nrmicro2314.
8. **Leonard AC, Grimwade JE.** 2010. Regulating DnaA complex assembly: it is time to fill the gaps. Curr. Opin. Microbiol. **13:**766–772. http://dx.doi.org/10.1016/j.mib.2010.10.001.
9. **Messer W.** 2002. The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. FEMS Microbiol. Rev. **26:**355–374. http://dx.doi.org/10.1111/j.1574-6976.2002.tb00620.x.
10. **Zakrzewska-Czerwińska J, Jakimowicz D, Zawilak-Pawlik A, Messer W.** 2007. Regulation of the initiation of chromosomal replication in bacteria. FEMS Microbiol. Rev. **31:**378–387. http://dx.doi.org/10.1111/j.1574-6976.2007.00070.x.
11. **Messer W, Blaesing F, Majka J, Nardmann J, Schaper S, Schmidt A, Seitz H, Speck C, Tungler D, Wegrzyn G, Weigel C, Welzeck M, Zakrzewska-Czerwinska J.** 1999. Functional domains of DnaA proteins. Biochimie **81:**819–825. http://dx.doi.org/10.1016/S0300-9084(99)00215-1.
12. **McGarry KC, Ryan VT, Grimwade JE, Leonard AC.** 2004. Two discriminatory binding sites in the Escherichia coli replication origin are required for DNA strand opening by initiator DnaA-ATP. Proc. Natl. Acad. Sci. U. S. A. **101:**2811–2816. http://dx.doi.org/10.1073/pnas.0400340101.
13. **Ozaki S, Noguchi Y, Hayashi Y, Miyazaki E, Katayama T.** 2012. Differentiation of the DnaA-oriC subcomplex for DNA unwinding in a replication initiation complex. J. Biol. Chem. **287:**37458–37471. http://dx.doi.org/10.1074/jbc.M112.372052.
14. **Rozgaja TA, Grimwade JE, Iqbal M, Czerwonka C, Vora M, Leonard AC.** 2011. Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during Escherichia coli pre-RC assembly. Mol. Microbiol. **82:**475–488. http://dx.doi.org/10.1111/j.1365-2958.2011.07827.x.
15. **Wolański M, Jakimowicz D, Zakrzewska-Czerwińska J.** 2012. AdpA, key regulator for morphological differentiation regulates bacterial chromosome replication. Open Biol. **2:**120097. http://dx.doi.org/10.1098/rsob.120097.
16. **Kaguni JM.** 2011. Replication initiation at the Escherichia coli chromosomal origin. Curr. Opin. Chem. Biol. **15:**606–613. http://dx.doi.org/10.1016/j.cbpa.2011.07.016.
17. **Jonas K, Liu J, Chien P, Laub MT.** 2013. Proteotoxic stress induces a cell-cycle arrest by stimulating Lon to degrade the replication initiator DnaA. Cell **154:**623–636. http://dx.doi.org/10.1016/j.cell.2013.06.034.
18. **Gorbatyuk B, Marczynski GT.** 2005. Regulated degradation of chromosome replication proteins DnaA and CtrA in Caulobacter crescentus. Mol. Microbiol. **55:**1233–1245. http://dx.doi.org/10.1111/j.1365-2958.2004.04459.x.
19. **Scholefield G, Murray H.** 2013. YabA and DnaD inhibit helix assembly of the DNA replication initiation protein DnaA. Mol. Microbiol. **90:**147–159. http://dx.doi.org/10.1111/mmi.12353.
20. **Rahn-Lee L, Merrikh H, Grossman AD, Losick R.** 2011. The sporulation protein SirA inhibits the binding of DnaA to the origin of replication by contacting a patch of clustered amino acids. J. Bacteriol. **193:**1302–1307. http://dx.doi.org/10.1128/JB.01390-10.
21. **Collier J, Shapiro L.** 2009. Feedback control of DnaA-mediated replication initiation by replisome-associated HdaA protein in Caulobacter. J. Bacteriol. **191:**5706–5716. http://dx.doi.org/10.1128/JB.00525-09.
22. **Collier J.** 2012. Regulation of chromosomal replication in Caulobacter crescentus. Plasmid **67:**76–87. http://dx.doi.org/10.1016/j.plasmid.2011.12.007.
23. **Smulczyk-Krawczyszyn A, Jakimowicz D, Ruban-Osmialowska B, Zawilak-Pawlik A, Majka J, Chater K, Zakrzewska-Czerwinska J.** 2006. Cluster of DnaA boxes involved in regulation of Streptomyces chromosome replication: from in silico to in vivo studies. J. Bacteriol. **188:**6184–6194. http://dx.doi.org/10.1128/JB.00528-06.
24. **Grossman AD, Losick R.** 1988. Extracellular control of spore formation in Bacillus subtilis. Proc. Natl. Acad. Sci. U. S. A. **85:**4369–4373. http://dx.doi.org/10.1073/pnas.85.12.4369.
25. **Solomon JM, Lazazzera BA, Grossman AD.** 1996. Purification and characterization of an extracellular peptide factor that affects two different developmental pathways in Bacillus subtilis. Genes Dev. **10:**2014–2024. http://dx.doi.org/10.1101/gad.10.16.2014.
26. **Zusman DR, Scott AE, Yang Z, Kirby JR.** 2007. Chemosensory pathways, motility and development in Myxococcus xanthus. Nat. Rev. Microbiol. **5:**862–872. http://dx.doi.org/10.1038/nrmicro1770.
27. **Chater KF.** 2000. Developmental decisions during sporulation in the aerial mycelium in *Streptomyces*, p 33–48. *In* Brun YV, Shimkets LJ (ed), Prokaryotic development. American Society for Microbiology, Washington, DC.
28. **Durre P, Hollergschwandner C.** 2004. Initiation of endospore formation in Clostridium acetobutylicum. Anaerobe **10:**69–74. http://dx.doi.org/10.1016/j.anaerobe.2003.11.001.
29. **Errington J.** 2003. Regulation of endospore formation in Bacillus subtilis. Nat. Rev. Microbiol. **1:**117–126. http://dx.doi.org/10.1038/nrmicro750.
30. **Piggot PJ, Hilbert DW.** 2004. Sporulation of Bacillus subtilis. Curr. Opin. Microbiol. **7:**579–586. http://dx.doi.org/10.1016/j.mib.2004.10.001.
31. **Stragier P, Losick R.** 1996. Molecular genetics of sporulation in Bacillus subtilis. Annu. Rev. Genet. **30:**297–241. http://dx.doi.org/10.1146/annurev.genet.30.1.297.
32. **Chater KF.** 1998. Taking a genetic scalpel to the *Streptomyces* colony. Microbiology (Reading, England) **144:**1465–1478. http://dx.doi.org/10.1099/00221287-144-6-1465.
33. **Chater KF.** 2011. Differentiation in Streptomyces: the properties and programming of diverse cell-types, 43–86. *In* Dyson P (ed), *Streptomyces*: molecular biology and biotechnology. Caister Academic Press, Norfolk, England.
34. **Brun YV, Janakiraman R.** 2000. The dimorphic life cycle of *Caulobacter* and stalked bacteria, p 297–317. In Brun YV, Shimkets LJ (ed), Prokaryotic development. American Society for Microbiology, Washington, DC.
35. **Veening JW, Murray H, Errington J.** 2009. A mechanism for cell cycle

regulation of sporulation initiation in Bacillus subtilis. Genes Dev. **23:** 1959–1970. http://dx.doi.org/10.1101/gad.528209.
36. **Ben-Yehuda S, Rudner DZ, Losick R.** 2003. RacA, a bacterial protein that anchors chromosomes to the cell poles. Science **299:**532–536. http://dx.doi.org/10.1126/science.1079914.
37. **Ryter A, Schaeffer P, Ionesco H.** 1966. Cytologic classification, by their blockage stage, of sporulation mutants of Bacillus subtilis Marburg. Ann. de l'Institut Pasteur (Paris) **110:**305–315. (In French.)
38. **Bath J, Wu LJ, Errington J, Wang JC.** 2000. Role of Bacillus subtilis SpoIIIE in DNA transport across the mother cell-prespore division septum. Science **290:**995–997. http://dx.doi.org/10.1126/science.290.5493.995.
39. **Grantcharova N, Lustig U, Flardh K.** 2005. Dynamics of FtsZ assembly during sporulation in Streptomyces coelicolor A3(2). J. Bacteriol. **187:** 3227–3237. http://dx.doi.org/10.1128/JB.187.9.3227-3237.2005.
40. **Schwedock J, McCormick JR, Angert JR, Nodwell J, Losick R.** 1997. Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of *Streptomyces coelicolor*. Mol. Microbiol. **25:**847–858. http://dx.doi.org/10.1111/j.1365-2958.1997.mmi507.x.
41. **Jakimowicz D, Zydek P, Kois A, Zakrzewska-Czerwinska J, Chater KF.** 2007. Alignment of multiple chromosomes along helical ParA scaffolding in sporulating *Streptomyces* hyphae. Mol. Microbiol. **65:**625–641. http://dx.doi.org/10.1111/j.1365-2958.2007.05815.x.
42. **Ruban-Ośmiałowska B, Jakimowicz D, Smulczyk-Krawczyszyn A, Chater KF, Zakrzewska-Czerwińska J.** 2006. Replisome localization in vegetative and aerial hyphae of *Streptomyces coelicolor*. J. Bacteriol. **188:** 7311–7316. http://dx.doi.org/10.1128/JB.00940-06.
43. **Curtis PD, Brun YV.** 2010. Getting in the loop: regulation of development in Caulobacter crescentus. Microbiol. Mol. Biol. Rev. **74:**13–41. http://dx.doi.org/10.1128/MMBR.00040-09.
44. **Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L.** 2004. Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proc. Natl. Acad. Sci. U. S. A. **101:**9257–9262. http://dx.doi.org/10.1073/pnas.0402606101.
45. **Bowman GR, Comolli LR, Zhu J, Eckart M, Koenig M, Downing KH, Moerner WE, Earnest T, Shapiro L.** 2008. A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. Cell **134:** 945–955. http://dx.doi.org/10.1016/j.cell.2008.07.015.
46. **Ebersbach G, Briegel A, Jensen GJ, Jacobs-Wagner C.** 2008. A self-associating protein critical for chromosome attachment, division, and polar organization in caulobacter. Cell **134:**956–968. http://dx.doi.org/10.1016/j.cell.2008.07.016.
47. **Thanbichler M, Shapiro L.** 2006. Mipz, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. Cell **126:**147–162. http://dx.doi.org/10.1016/j.cell.2006.05.038.
48. **Burbulys D, Trach KA, Hoch JA.** 1991. Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. Cell **64:**545–552. http://dx.doi.org/10.1016/0092-8674(91)90238-T.
49. **Molle V, Fujita M, Jensen ST, Eichenberger P, González-Pastor JE, Liu JS, Losick R.** 2003. The Spo0A regulon of Bacillus subtilis. Mol. Microbiol. **50:**1683–1701. http://dx.doi.org/10.1046/j.1365-2958.2003.03818.x.
50. **Fujita M, Gonzalez-Pastor JE, Losick R.** 2005. High- and low-threshold genes in the Spo0A regulon of Bacillus subtilis. J. Bacteriol. **187:**1357–1368. http://dx.doi.org/10.1128/JB.187.4.1357-1368.2005.
51. **Ben-Yehuda S, Losick R.** 2002. Asymmetric cell division in B. subtilis involves a spiral-like intermediate of the cytokinetic protein FtsZ. Cell **109:**257–266. http://dx.doi.org/10.1016/S0092-8674(02)00698-0.
52. **Khvorova A, Zhang L, Higgins ML, Piggot PJ.** 1998. The spoIIE locus is involved in the Spo0A-dependent switch in the location of FtsZ rings in Bacillus subtilis. J. Bacteriol. **180:**1256–1260.
53. **Levin PA, Losick R.** 1996. Transcription factor Spo0A switches the localization of the cell division protein FtsZ from a medial to a bipolar pattern in Bacillus subtilis. Genes Dev. **10:**478–488. http://dx.doi.org/10.1101/gad.10.4.478.
54. **Lopez D, Vlamakis H, Kolter R.** 2009. Generation of multiple cell types in Bacillus subtilis. FEMS Microbiol. Rev. **33:**152–163. http://dx.doi.org/10.1111/j.1574-6976.2008.00148.x.
55. **Flärdh K, Buttner MJ.** 2009. *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. Nat. Rev. Microbiol. **7:**36–49. http://dx.doi.org/10.1038/nrmicro1968.
56. **Hopwood DA.** 1967. Genetic analysis and genome structure in *Streptomyces coelicolor*. Bacteriol. Rev. **31:**373–403.
57. **Merrick MJ.** 1976. A morphological and genetic mapping study of bald colony mutants of Streptomyces coelicolor. J. Gen. Microbiol. **96:**299–315. http://dx.doi.org/10.1099/00221287-96-2-299.
58. **Chater KF.** 1972. A morphological and genetic mapping study of white colony mutants of Streptomyces coelicolor. J. Gen. Microbiol. **72:**9–28. http://dx.doi.org/10.1099/00221287-72-1-9.
59. **Hopwood DA, Wildermuth H, Palmer HM.** 1970. Mutants of Streptomyces coelicolor defective in sporulation. J. Gen. Microbiol. **61:**397–408. http://dx.doi.org/10.1099/00221287-61-3-397.
60. **Liu G, Chater KF, Chandra G, Niu G, Tan H.** 2013. Molecular regulation of antibiotic biosynthesis in streptomyces. Microbiol. Mol. Biol. Rev. **77:**112–143. http://dx.doi.org/10.1128/MMBR.00054-12.
61. **Horinouchi S.** 2007. Mining and polishing of the treasure trove in the bacterial genus Streptomyces. Biosci. Biotechnol. Biochem. **71:**283–299. http://dx.doi.org/10.1271/bbb.60627.
62. **Ohnishi Y, Yamazaki H, Kato JY, Tomono A, Horinouchi S.** 2005. AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci. Biotechnol. Biochem. **69:**431–439. http://dx.doi.org/10.1271/bbb.69.431.
63. **Wolanski M, Donczew R, Kois-Ostrowska A, Masiewicz P, Jakimowicz D, Zakrzewska-Czerwinska J.** 2011. The level of AdpA directly affects expression of developmental genes in *Streptomyces coelicolor*. J. Bacteriol. **193:**6358–6365. http://dx.doi.org/10.1128/JB.05734-11.
64. **Higo A, Hara H, Horinouchi S, Ohnishi Y.** 2012. Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in Streptomyces, revealed the extent and complexity of the AdpA regulatory network. DNA Res. **19:**259–274. http://dx.doi.org/10.1093/dnares/dss010.
65. **Laub MT.** 2000. Global analysis of the genetic network controlling a bacterial cell cycle. Science **290:**2144–2148. http://dx.doi.org/10.1126/science.290.5499.2144.
66. **Jenal U.** 2009. The role of proteolysis in the Caulobacter crescentus cell cycle and development. Res. Microbiol. **160:**687–695. http://dx.doi.org/10.1016/j.resmic.2009.09.006.
67. **McAdams HH, Shapiro L.** 2011. The architecture and conservation pattern of whole-cell control circuitry. J. Mol. Biol. **409:**28–35. http://dx.doi.org/10.1016/j.jmb.2011.02.041.
68. **Laub MT, Chen SL, Shapiro L, McAdams HH.** 2002. Genes directly controlled by CtrA, a master regulator of the Caulobacter cell cycle. Proc. Natl. Acad. Sci. U. S. A. **99:**4632–4637. http://dx.doi.org/10.1073/pnas.062065699.
69. **Jacobs C, Domian IJ, Maddock JR, Shapiro L.** 1999. Cell cycle-dependent polar localization of an essential bacterial histidine kinase that controls DNA replication and cell division. Cell **97:**111–120. http://dx.doi.org/10.1016/S0092-8674(00)80719-9.
70. **Tsokos CG, Laub MT.** 2012. Polarity and cell fate asymmetry in Caulobacter crescentus. Curr. Opin. Microbiol. **15:**744–750. http://dx.doi.org/10.1016/j.mib.2012.10.011.
71. **Domian IJ, Quon KC, Shapiro L.** 1997. Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. Cell **90:**415–424. http://dx.doi.org/10.1016/S0092-8674(00)80502-4.
72. **Holtzendorff J, Hung D, Brende P, Reisenauer A, Viollier PH, McAdams HH, Shapiro L.** 2004. Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. Science **304:**983–987. http://dx.doi.org/10.1126/science.1095191.
73. **Kirkpatrick CL, Viollier PH.** 2012. Decoding Caulobacter development. FEMS Microbiol. Rev. **36:**193–205. http://dx.doi.org/10.1111/j.1574-6976.2011.00309.x.
74. **Castilla-Llorente V, Munoz-Espin D, Villar L, Salas M, Meijer WJ.** 2006. Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication. EMBO J. **25:**3890–3899. http://dx.doi.org/10.1038/sj.emboj.7601266.
75. **Siam R, Marczynski GT.** 2000. Cell cycle regulator phosphorylation stimulates two distinct modes of binding at a chromosome replication origin. EMBO J. **19:**1138–1147. http://dx.doi.org/10.1093/emboj/19.5.1138.
76. **Boonstra M, de Jong IG, Scholefield G, Murray H, Kuipers OP, Veening JW.** 2013. Spo0A regulates chromosome copy number during sporulation by directly binding to the origin of replication in Bacillus

subtilis. Mol. Microbiol. **87:**925–938. http://dx.doi.org/10.1111/mmi.12141.
77. **Taylor JA, Ouimet MC, Wargachuk R, Marczynski GT.** 2011. The Caulobacter crescentus chromosome replication origin evolved two classes of weak DnaA binding sites. Mol. Microbiol. **82:**312–326. http://dx.doi.org/10.1111/j.1365-2958.2011.07785.x.
78. **Jakimowicz D, Majka J, Messer W, Speck C, Fernandez M, Martin MC, Sanchez J, Schauwecker F, Keller U, Schrempf H, Zakrzewska-Czerwinska J.** 1998. Structural elements of the Streptomyces oriC region and their interactions with the DnaA protein. Microbiology **144**(Pt 5): 1281–1290. http://dx.doi.org/10.1099/00221287-144-5-1281.
79. **Moriya S, Atlung T, Hansen FG, Yoshikawa H, Ogasawara N.** 1992. Cloning of an autonomously replicating sequence (ars) from the Bacillus subtilis chromosome. Mol. Microbiol. **6:**309–315. http://dx.doi.org/10.1111/j.1365-2958.1992.tb01473.x.
80. **Zakrzewska-Czerwińska J, Majka J, Schrempf H.** 1995. Minimal requirements of the *Streptomyces lividans* 66 oriC region and its transcriptional and translational activities. J. Bacteriol. **177:**4765–4771.
81. **Jakimowicz D, Majka J, Konopa G, Wegrzyn G, Messer W, Schrempf H, Zakrzewska-Czerwinska J.** 2000. Architecture of the *Streptomyces lividans* DnaA protein-replication origin complexes. J. Mol. Biol. **298:** 351–364. http://dx.doi.org/10.1006/jmbi.2000.3686.
82. **Krause M, Ruckert B, Lurz R, Messer W.** 1997. Complexes at the replication origin of Bacillus subtilis with homologous and heterologous DnaA protein. J. Mol. Biol. **274:**365–380. http://dx.doi.org/10.1006/jmbi.1997.1404.
83. **Marczynski GT, Shapiro L.** 1992. Cell-cycle control of a cloned chromosomal origin of replication from Caulobacter crescentus. J. Mol. Biol. **226:**959–977. http://dx.doi.org/10.1016/0022-2836(92)91045-Q.
84. **Quon KC, Yang B, Domian IJ, Shapiro L, Marczynski GT.** 1998. Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. Proc. Natl. Acad. Sci. U. S. A. **95:**120–125. http://dx.doi.org/10.1073/pnas.95.1.120.
85. **Bastedo DP, Marczynski GT.** 2009. CtrA response regulator binding to the Caulobacter chromosome replication origin is required during nutrient and antibiotic stress as well as during cell cycle progression. Mol. Microbiol. **72:**139–154. http://dx.doi.org/10.1111/j.1365-2958.2009.06630.x.
86. **Marczynski GT, Lentine K, Shapiro L.** 1995. A developmentally regulated chromosomal origin of replication uses essential transcription elements. Genes Dev. **9:**1543–1557. http://dx.doi.org/10.1101/gad.9.12.1543.
87. **Biondi EG, Reisinger SJ, Skerker JM, Arif M, Perchuk BS, Ryan KR, Laub MT.** 2006. Regulation of the bacterial cell cycle by an integrated genetic circuit. Nature **444:**899–904. http://dx.doi.org/10.1038/nature05321.
88. **Jenal U, Fuchs T.** 1998. An essential protease involved in bacterial cell-cycle control. EMBO J. **17:**5658–5669. http://dx.doi.org/10.1093/emboj/17.19.5658.
89. **McGrath PT, Iniesta AA, Ryan KR, Shapiro L, McAdams HH.** 2006. A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. Cell **124:**535–547. http://dx.doi.org/10.1016/j.cell.2005.12.033.
90. **Duerig A, Abel S, Folcher M, Nicollier M, Schwede T, Amiot N, Giese B, Jenal U.** 2009. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. Genes Dev. **23:**93–104. http://dx.doi.org/10.1101/gad.502409.
91. **Iniesta AA, McGrath PT, Reisenauer A, McAdams HH, Shapiro L.** 2006. A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. Proc. Natl. Acad. Sci. U. S. A. **103:**10935–10940. http://dx.doi.org/10.1073/pnas.0604554103.
92. **Domian IJ, Reisenauer A, Shapiro L.** 1999. Feedback control of a master bacterial cell-cycle regulator. Proc. Natl. Acad. Sci. U. S. A. **96:**6648–6653. http://dx.doi.org/10.1073/pnas.96.12.6648.
93. **Reisenauer A, Shapiro L.** 2002. DNA methylation affects the cell cycle transcription of the CtrA global regulator in Caulobacter. EMBO J. **21:** 4969–4977. http://dx.doi.org/10.1093/emboj/cdf490.
94. **Gora KG, Tsokos CG, Chen YE, Srinivasan BS, Perchuk BS, Laub MT.** 2010. A cell-type-specific protein-protein interaction modulates transcriptional activity of a master regulator in Caulobacter crescentus. Mol. Cell **39:**455–467. http://dx.doi.org/10.1016/j.molcel.2010.06.024.
95. **Tan MH, Kozdon JB, Shen X, Shapiro L, McAdams HH.** 2010. An essential transcription factor, SciP, enhances robustness of Caulobacter cell cycle regulation. Proc. Natl. Acad. Sci. U. S. A. **107:**18985–18990. http://dx.doi.org/10.1073/pnas.1014395107.
96. **Collier J, McAdams HH, Shapiro L.** 2007. A DNA methylation ratchet governs progression through a bacterial cell cycle. Proc. Natl. Acad. Sci. U. S. A. **104:**17111–17116. http://dx.doi.org/10.1073/pnas.0708112104.
97. **Fernandez-Fernandez C, Grosse K, Sourjik V, Collier J.** 2013. The beta-sliding clamp directs the localization of HdaA to the replisome in Caulobacter crescentus. Microbiology **159:**2237–2248. http://dx.doi.org/10.1099/mic.0.068577-0.
98. **Collier J, Murray SR, Shapiro L.** 2006. DnaA couples DNA replication and the expression of two cell cycle master regulators. EMBO J. **25:**346–356. http://dx.doi.org/10.1038/sj.emboj.7600927.
99. **Shaheen SM, Ouimet MC, Marczynski GT.** 2009. Comparative analysis of Caulobacter chromosome replication origins. Microbiology **155:** 1215–1225. http://dx.doi.org/10.1099/mic.0.025528-0.
100. **Gonzalez D, Kozdon JB, McAdams HH, Shapiro L, Collier J.** 2014. The functions of DNA methylation by CcrM in Caulobacter crescentus: a global approach. Nucleic Acids Res. **42:**3720–3735. http://dx.doi.org/10.1093/nar/gkt1352.
101. **Xenopoulos P, Piggot PJ.** 2011. Regulation of growth of the mother cell and chromosome replication during sporulation of Bacillus subtilis. J. Bacteriol. **193:**3117–3126. http://dx.doi.org/10.1128/JB.00204-11.
102. **Chibazakura T, Kawamura F, Takahashi H.** 1991. Differential regulation of spo0A transcription in Bacillus subtilis: glucose represses promoter switching at the initiation of sporulation. J. Bacteriol. **173:**2625–2632.
103. **Chastanet A, Losick R.** 2011. Just-in-time control of Spo0A synthesis in Bacillus subtilis by multiple regulatory mechanisms. J. Bacteriol. **193:** 6366–6374. http://dx.doi.org/10.1128/JB.06057-11.
104. **Chibazakura T, Kawamura F, Asai K, Takahashi H.** 1995. Effects of spo0 mutations on spo0A promoter switching at the initiation of sporulation in Bacillus subtilis. J. Bacteriol. **177:**4520–4523.
105. **Strauch MA, Trach KA, Day J, Hoch JA.** 1992. Spo0A activates and represses its own synthesis by binding at its dual promoters. Biochimie **74:**619–626. http://dx.doi.org/10.1016/0300-9084(92)90133-Y.
106. **Grossman AD.** 1995. Genetic networks controlling the initiation of sporulation and the development of genetic competence in Bacillus subtilis. Annu. Rev. Genet. **29:**477–508. http://dx.doi.org/10.1146/annurev.ge.29.120195.002401.
107. **Ogura Y, Imai Y, Ogasawara N, Moriya S.** 2001. Autoregulation of the dnaA-dnaN operon and effects of DnaA protein levels on replication initiation in Bacillus subtilis. J. Bacteriol. **183:**3833–3841. http://dx.doi.org/10.1128/JB.183.13.3833-3841.2001.
108. **Rahn-Lee L, Gorbatyuk B, Skovgaard O, Losick R.** 2009. The conserved sporulation protein YneE inhibits DNA replication in Bacillus subtilis. J. Bacteriol. **191:**3736–3739. http://dx.doi.org/10.1128/JB.00216-09.
109. **Scholefield G, Errington J, Murray H.** 2012. Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA. EMBO J. **31:**1542–1555. http://dx.doi.org/10.1038/emboj.2012.6.
110. **Briggs GS, Smits WK, Soultanas P.** 2012. Chromosomal replication initiation machinery of low-G+C-content Firmicutes. J. Bacteriol. **194:** 5162–5170. http://dx.doi.org/10.1128/JB.00865-12.
111. **Burkholder WF, Kurtser I, Grossman AD.** 2001. Replication initiation proteins regulate a developmental checkpoint in Bacillus subtilis. Cell **104:**269–279. http://dx.doi.org/10.1016/S0092-8674(01)00211-2.
112. **Li R, Liu G, Xie Z, He X, Chen W, Deng Z, Tan H.** 2010. PolY, a transcriptional regulator with ATPase activity, directly activates transcription of *polR* in polyoxin biosynthesis in *Streptomyces cacaoi*. Mol. Microbiol. **75:**349–364. http://dx.doi.org/10.1111/j.1365-2958.2009.06968.x.
113. **Xu W, Huang J, Lin R, Shi J, Cohen SN.** 2010. Regulation of morphological differentiation in S. coelicolor by RNase III (AbsB) cleavage of mRNA encoding the AdpA transcription factor. Mol. Microbiol. **75:**781–791. http://dx.doi.org/10.1111/j.1365-2958.2009.07023.x.
114. **den Hengst CD, Tran NT, Bibb MJ, Chandra G, Leskiw BK, Buttner MJ.** 2010. Genes essential for morphological development and antibiotic production in Streptomyces coelicolor are targets of BldD during vegetative growth. Mol. Microbiol. **78:**361–379. http://dx.doi.org/10.1111/j.1365-2958.2010.07338.x.
115. **Nguyen KT, Tenor J, Stettler H, Nguyen LT, Nguyen LD, Thompson CJ.** 2003. Colonial differentiation in *Streptomyces coelicolor* depends on

translation of a specific codon within the *adpA* gene. J. Bacteriol. **185:** 7291–7296. http://dx.doi.org/10.1128/JB.185.24.7291-7296.2003.
116. **Takano E, Tao M, Long F, Bibb MJ, Wang L, Li W, Buttner MJ, Deng ZX, Chater KF.** 2003. A rare leucine codon in *adpA* is implicated in the morphological defect of bldA mutants of *Streptomyces coelicolor*. Mol. Microbiol. **50:**475–486. http://dx.doi.org/10.1046/j.1365-2958.2003.03728.x.
117. **Pettersson BM, Kirsebom LA.** 2011. tRNA accumulation and suppression of the bldA phenotype during development in Streptomyces coelicolor. Mol. Microbiol. **79:**1602–1614. http://dx.doi.org/10.1111/j.1365-2958.2011.07543.x.
118. **Higo A, Horinouchi S, Ohnishi Y.** 2011. Strict regulation of morphological differentiation and secondary metabolism by a positive feedback loop between two global regulators AdpA and BldA in Streptomyces griseus. Mol. Microbiol. **81:**1607–1622. http://dx.doi.org/10.1111/j.1365-2958.2011.07795.x.
119. **Majka J, Jakimowicz D, Messer W, Schrempf H, Lisowski M, Zakrzewska-Czerwinska J.** 1999. Interactions of the *Streptomyces lividans* initiator protein DnaA with its target. Eur. J. Biochem. **260:**325–335.
120. **Madiraju MV, Moomey M, Neuenschwander PF, Muniruzzaman S, Yamamoto K, Grimwade JE, Rajagopalan M.** 2006. The intrinsic ATPase activity of *Mycobacterium tuberculosis* DnaA promotes rapid oligomerization of DnaA on *oriC*. Mol. Microbiol. **59:**1876–1890. http://dx.doi.org/10.1111/j.1365-2958.2006.05068.x.

Marcin Wolański obtained his M.Sc. (Biotechnology) from the Faculty of Chemistry at the Technical University of Wrocław, Poland, in 2005 and his Ph.D. (Molecular Biology) from the Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland, in 2011. Shortly after obtaining his Ph.D., he has started working as a postdoctoral fellow at the Department of Molecular Microbiology (Faculty of Biotechnology) of the University of Wrocław. During the course of his Ph.D. studies, he investigated protein factors involved in the regulation of chromosome replication and differentiation in *Streptomyces coelicolor*. Since 2011, his research has been also focused on the transcription regulation of antibiotic gene clusters in *Actinomycetales*.

Dagmara Jakimowicz obtained her M.Sc. in biotechnology from the Faculty of Natural Sciences, University of Wrocław, Poland, in 1996, and her Ph.D. in biological sciences from the Department of Microbiology at the Institute of Immunology and Experimental Therapy, Polish Academy of Science, Wroclaw, Poland, in 2000. She worked as a postdoctoral research assistant at the John Innes Centre, Norwich, United Kingdom (2001 to 2005), where she carried out research on chromosome segregation in *Streptomyces*. Since her return to Poland (2005), she has been continuing research on chromosome segregation and its coordination with replication in *Streptomyces* and also in *Mycobacterium*. Since 2007, she has been Associate Professor and the group leader in the Molecular Microbiology Department at the Faculty of Biotechnology, University of Wrocław.

Jolanta Zakrzewska-Czerwińska is a Professor of molecular microbiology. She obtained her M.Sc. (chemistry/bioengineering) from the Faculty of Chemistry at the Wrocław University of Technology, Poland, in 1984 and her Ph.D. from the Institute of Immunology and Experimental Therapy, Wrocław, Poland, in 1989. She did postdoctoral research as the Alexander von Humboldt fellow at the Osnabruck University, Germany (1991 to 1992), where she carried out research on initiation of chromosome replication in *Streptomyces*. In 1999 to 2013, she was a head of the Department of Microbiology at the Institute of Immunology and Experimental Therapy and her research focused on the molecular taxonomy of *Actinomycetales* and initiation of chromosome replication in bacteria. Since 2007, she has worked as a head of the Department of Molecular Microbiology at the Faculty of Biotechnology, University of Wrocław, Poland. Currently, her main area of research is chromosome replication and regulation of this process in various bacteria, particularly in *Streptomyces* and *Mycobacterium*.

JB Journals.ASM.org

DNA Looping in Prokaryotes: Experimental and Theoretical Approaches

Axel Cournac,[a] Jacqueline Plumbridge[b]

Laboratoire de Physique Théorique de la Matière Condensée, UMR7600 (associated with the Université Pierre et Marie Curie), Paris, France[a]; CNRS UPR9073 (associated with the Université Paris Diderot, Sorbonne Paris Cité), Institut de Biologie Physico-Chimique, Paris, France[b]

Transcriptional regulation is at the heart of biological functions such as adaptation to a changing environment or to new carbon sources. One of the mechanisms which has been found to modulate transcription, either positively (activation) or negatively (repression), involves the formation of DNA loops. A DNA loop occurs when a protein or a complex of proteins simultaneously binds to two different sites on DNA with looping out of the intervening DNA. This simple mechanism is central to the regulation of several operons in the genome of the bacterium *Escherichia coli*, like the *lac* operon, one of the paradigms of genetic regulation. The aim of this review is to gather and discuss concepts and ideas from experimental biology and theoretical physics concerning DNA looping in genetic regulation. We first describe experimental techniques designed to show the formation of a DNA loop. We then present the benefits that can or could be derived from a mechanism involving DNA looping. Some of these are already experimentally proven, but others are theoretical predictions and merit experimental investigation. Then, we try to identify other genetic systems that could be regulated by a DNA looping mechanism in the genome of *Escherichia coli*. We found many operons that, according to our set of criteria, have a good chance to be regulated with a DNA loop. Finally, we discuss the proposition recently made by both biologists and physicists that this mechanism could also act at the genomic scale and play a crucial role in the spatial organization of genomes.

Different levels of DNA organization exist within bacterial chromosomes. In the case of *Escherichia coli*, the genome has been shown to be organized, on the largest scale, in four individual macrodomains (Ter, Ori, Right, and Left) and two less-structured regions (1) that have a precise localization within the cell throughout the cell cycle and are associated with specific binding proteins (2). Large-scale DNA loops have been visualized by nucleoid-spreading techniques and are thought to be stabilized by membrane and/or RNA components (3, 4). Then, at the scale of 10 kb, there are topological domains formed by supercoiled structures (5, 6) whose barriers are not placed stably at fixed sites but instead are randomly distributed (7). These intermediate loops can be stabilized with nucleoid-associated proteins like H-NS (8). Finally, there are smaller loops of a few hundred base pairs made by specific transcription factors that have a direct impact on transcription. Although loops of different sizes can have functional consequences for genomic organization and genetic regulation, it is the last category that we focus on in this review.

A first hint that a transcription factor can bind simultaneously to two sites derived from the work of Kania and Müller-Hill in 1977 (9). However, the first experimental demonstration and clear proposal for the existence of a DNA loop affecting gene regulation was in 1984, by the team of Robert Schleif (10), working on the regulatory region of the *ara* operon. Since then, this phenomenon has been found to play a role in the regulation of several other operons in *Escherichia coli*. Several historical reviews published in the beginning of the 1990s describe and compare these different cases of DNA loop formation and start to consider the energetics of loop formation (11, 12, 13). Regulatory mechanisms involving DNA looping have also been found in eukaryotes. These often involve more-sophisticated models requiring long-range interactions, like "enhancers" or "insulators" (14) forming higher-order chromatin structures.

A DNA loop is formed when a protein or a complex of proteins simultaneously binds to two different sites on DNA. Depending on the location of the proteins relative to the transcriptional start site, the formation of the DNA loop can be responsible for transcriptional repression or for transcriptional activation. Examples of both have been documented experimentally, but the phenomenon of transcriptional repression has been better studied at the theoretical level (15, 16, 17). Figure 1 presents the two simple schemes of how a DNA loop can be at the origin of transcriptional repression or activation. In Fig. 1A, a bivalent transcription factor binds simultaneously to two binding sites, creating a loop, generally of the order of a hundred base pairs. The activity of RNA polymerase is blocked, and there is no expression of the operon. Inhibition of transcription can be either at the level of polymerase binding, e.g., due to competition between repressor and polymerase for overlapping binding sites, or at later stages, because RNA polymerase can be trapped in the loop or not be able to recruit an activator (e.g., cyclic AMP receptor protein [CRP]) (11). Repression requiring DNA looping has been experimentally shown in several operons in *Escherichia coli*; the best-studied examples are *ara* (10), *lac* (18, 19), *gal* (20, 21), *deo* (22, 23), and *nag* (24, 25).

DNA looping can also be at the origin of transcriptional activation. The activation of the expression of the *glnALG* operon of *Escherichia coli* by NtrC, which is sometimes called a bacterial "enhancer," is one of the best-studied cases of this type of system (26). Several binding sites for the transcription factor are placed upstream of the promoter site. RNA polymerase, which in all cases studied is of the σ^{54} type, binds only weakly to the promoter and is

Published ahead of print 4 January 2013

Address correspondence to Jacqueline Plumbridge, Jackie.Plumbridge@ibpc.fr.

doi:10.1128/JB.02038-12

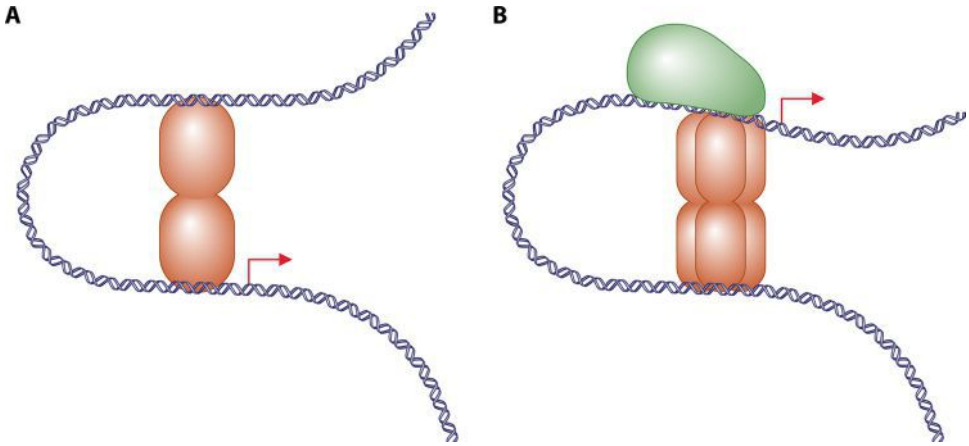


FIG 1 The two main mechanisms of DNA looping in transcriptional regulation are depicted. (A) A DNA loop can be responsible for transcriptional repression. A bivalent transcription factor binds simultaneously to two binding sites and blocks access to the RNA polymerase (e.g., the regulation of the *lac* operon, the best studied in *E. coli*). (B) A DNA loop can be responsible for transcriptional activation. Transcription factors bind away from the site of fixation of RNA polymerase (normally of σ^{54} type) and help the recruitment of RNA polymerase and formation of an open complex (in *E. coli*, the *glnALG* operon is regulated by this mechanism).

unable to form an open complex. Supplementary proteins can facilitate the binding of RNA polymerase and are absolutely required to melt the DNA at the transcription start site. These proteins are called EBPs, for "enhancer binding proteins." Interestingly, these proteins act at a distance from the promoter site, but generally on the order of about a hundred base pairs, and can function independently of the orientation of the binding site. It is the latter characteristic that recalls the enhancers of eukaryote genomes (27). In their inactive state, the prokaryotic enhancer proteins are usually dimers that bind to pairs of tandem sites in the enhancer elements. Upon encountering inducing conditions, they oligomerize into hexameric or heptameric ATPase-active rings that use the energy of ATP hydrolysis to physically remodel the otherwise stable "closed" complex formed by RNA polymerase with σ^{54} at its promoters. ATP hydrolysis changes the conformation of the EBP from ATP bound to ADP bound and alters its contact with σ^{54}. This overcomes the blockage to open complex formation, resulting in separation of the double-stranded DNA and binding of the template strand in the active site of the polymerase, producing a strong transcriptional activation of the operon (28, 29). In the genome of *Escherichia coli*, there exist several operons that have been shown to be regulated by a similar mechanism, including *fdhF*, *glnH*, *hypA*, *prpB*, etc. (28, 30). Examples are known in other species (31).

Since the early years of the 21st century, transcription factor-DNA interactions and the phenomenon of DNA looping have become a subject of analysis and research by theoretical physicists (15, 17, 32, 33, 34). DNA looping is seen as a thermodynamic system, and the formalism of statistical physics allows, for example, calculation of the probabilities of a molecule being in one of the different accessible states, i.e., looped or not-looped states (16, 35, 36, 37). This approach brought new ideas and made predictions concerning DNA loops that are interesting to compare with the experimental results.

In this review, we will start by describing various experimental techniques that can be relevant for the detection and justification of the presence of a DNA loop inside a regulatory region. In the second section, we will present the advantages that a mechanism involving DNA looping could bring to gene regulation systems. Several of these properties are predictions derived from recent theoretical work. In the third section, we attempt to identify other operons that could be regulated by such a mechanism. We will focus on the model organism *Escherichia coli* and make use of some bioinformatics tools to look for new loops that we compare to those in the existing literature.

EXPERIMENTAL TECHNIQUES TO SHOW DNA LOOPING

Several different techniques have been used to demonstrate the existence of a DNA loop. Some of them are classical molecular biology, whereas others involve electron microscopy. The various techniques give information on different aspects of the DNA looping phenomenon, i.e., its impact on gene expression and, also, its thermodynamics. Interestingly, we will see that gene expression and the thermodynamics of the system are connected.

Requirement of two operator sites acting synergistically. One of the simplest experiments routinely employed to show DNA looping is to mutate one or other of the binding sites and see an effect on the transcriptional regulation, for example, by quantifying the activity of a reporter gene. If the loss of one binding site is equivalent to the loss of both binding sites, it means that the two sites act with "cooperativity." Regulation does not have to be 100% dependent on the two sites for them to be cooperative; e.g., at *lac*, binding to O1 and either O2 or O3 is cooperative but there is still some repression by the single O1 site (18, 19). The two sites must act synergistically and not additively. For example, a mutation that inactivates one of the two binding sites for GalR suppresses the repression at the *gal*P2 promoter of the *gal* operon (38). It should be noted that cooperativity is a necessary but not sufficient condition for the existence of a DNA loop. It is not possible to completely exclude the possibility that other mechanisms responsible for the observed cooperativity could exist (for example, propagation of a physical constraint along the DNA). Therefore, other experiments are necessary to prove that the observed cooperativity is effectively due to a mechanism of DNA looping.

DNase I footprinting. Another classic technique is DNase I

footprinting, because it shows precisely where a regulatory protein can bind on a DNA sequence. Complementing the previous technique, it can confirm the existence of two sites and, under certain circumstances, can show the cooperative binding to the two sites, e.g., if the loss of one operator reduces the affinity of the repressor for the other site (25). Moreover, in the case of certain small DNA loops, it has been noticed by several workers that a pattern of hypersensitive DNase I cleavages every 10 to 11 bp, (with some protection of the intervening DNA) appears between the two binding sites implicated in the DNA loop (24, 39), as shown originally for phage λ repressor (40). The bending that is required to form a small DNA loop compresses the grooves on the face of the DNA on the inside of the loop and makes the grooves on the outside of the loop wider. Because DNase I cuts in the minor groove and its activity is highly sensitive to the width of the minor groove (41), looped regions are readily identified by the formation of hypersensitive cleavages separated by about one turn of a B-form DNA helix. Thus, this technique can bring convincing evidence in favor of DNA looping in a regulatory region. (It should be noted that this technique is carried out on a population of DNA fragments, so there is still the possibility that partial protection could come from different fragments.)

Requirement for the two binding sites to be in phase. For the formation of DNA loops, and particularly for small loops of about 100 bp, it is necessary that the two binding sites are in phase on the double-helix to allow the interaction. DNA has natural torsional rigidity and develops a resistive torque when it is twisted. Torsional stiffness thus affects the cyclization of DNA, and it can be measured by single-molecule experiments (42). If the two sites are on opposite sides of the double helix, torsional energy is required to twist the DNA so that the operators are available to be simultaneously bound by the oligomeric protein. The torsional energy amounts to 4 kcal/mol to twist by one half turn a linear DNA molecule of 200 bp *in vitro* (43). It is comparable to the binding affinities between a transcription factor (TF) and its binding site that are between 5 and 15 kcal/mol (44, 45). So when two binding sites are dephased by adding 5 to 6 bp (i.e., half a turn of a double helix), a loss of regulation can be expected. This is the method used by the team of Robert Schleif to demonstrate DNA looping in the regulatory region of the *ara* operon (10). Repression was impaired in cases in which half-integral turns of the DNA helix were introduced, but repression was nearly normal for the insertions of +11 and +31 bp. It can also be noted that the sequence of the intervening DNA is not completely neutral. In the case of NagC repression of the divergent *nagE-nagB* genes, the CRP binding site introduces an intrinsic bend in the interoperator DNA. Displacing the CRP binding site by half a turn derepresses the expression of both genes, even though the interoperator distance is not altered (46).

Band shift experiments. Electrophoretic mobility shift analysis (band shift) can help to indicate the formation of a DNA loop inside a region. The migration of DNA-protein complexes in a polyacrylamide gel depends upon molecular mass, charge, and shape. It is generally used to determine the affinity of a protein for a specific DNA sequence and can be used to calculate a value of the dissociation constant of a protein for a binding sequence. Complexes forming a DNA loop migrate more slowly than the same components in a complex on a linear DNA molecule, i.e., without a DNA loop (19). Therefore, this technique could detect the formation of DNA looping, but as the migration of a protein-DNA complex in a gel depends upon many variables, it is important to be able to compare the same components in a linear and looped conformation. Oehler et al. (19) observed, in a single lane, two shifted bands, one corresponding to two dimers of LacI binding independently and one to a tetramer of LacI binding and forming a loop. The latter band was only observed with LacI, which was capable of forming tetramers, which is a strong argument in favor of its identification as a looped complex. It should be noticed that the position of migration of a looped complex varies considerably with the sequence of the DNA, as observed in reference 47.

Requirement for oligomeric regulators. To be capable of forming a DNA loop, the regulatory protein must form oligomers with two independent DNA-binding domains. For the majority of standard helix-turn-helix-containing prokaryotic regulators binding to palindromic sequences, this implies the formation of a tetramer. Two methods have been employed to demonstrate the necessity of an oligomeric repressor for regulation.

In the case of the *gal* operon, the two GalR operators, which were suspected to be at the origin of a DNA loop, were replaced by LacI binding sites. Then, repression of the system in the presence of LacI was compared to that with GalR. Comparable repression in the two cases implied loop formation. This interpretation was confirmed by the use of a mutated form of LacI that is incapable of forming tetramers but that exhibits the same DNA binding properties. In this case, the *gal* operon was derepressed, showing that a DNA loop was very likely formed in the case of GalR (21). This method works well for related proteins with similar affinities for DNA and whose sites do not interfere with RNA polymerase binding but could be complicated if the operator covers the promoter and its affinity for DNA is not similar to that of LacI for its sites. A more general method is to create mutations within the oligomerization domain of the regulatory protein (48). This method was first applied (by chance) to the *lac* operon, where a frameshift mutation in LacI produced a protein missing the C-terminal oligomerization domain. This protein repressed much less well than the wild-type protein that is capable of forming tetramers (19).

It should also be noted that, in some cases, a repressor can require an auxiliary protein to stabilize a DNA loop. This is the case for GalR. The laboratory of Sankar Adhya demonstrated that the Gal repressor can only form a loop at the *gal* operon when the nucleoid-associated protein HU is bound to a specific site near the apex of the loop. The resulting nucleoprotein complex, called a "repressosome," was shown to carry an antiparallel loop (49, 50).

Plasmid concatenation. An original method was devised by the team of Kustu to support the model of DNA loop formation responsible for the transcriptional activation in the *glnALG* operon (51). They showed that the consequence of DNA looping was an increase in the local concentration of the transcriptional activator near the promoter. The transcription of *glnA* requires NtrC (nitrogen regulatory protein C) to bind upstream of the *glnA* promoter. Their method consisted of forming concatemers of two plasmids, one carrying the enhancer binding sites that bind NtrC regulatory protein and the other one carrying the *glnA* promoter. The 3-dimensional (3-D) interaction of NtrC with the RNA promoter-polymerase complex by DNA looping was shown by the comparison of the quantity of transcripts synthesized in the case where the plasmids were concatenated and in the case where they were not. For certain concentrations of the regulatory proteins, the system with concatenated plasmids showed a greater production of *glnA* transcripts. This experiment demonstrated that one

function of DNA loop formation is to enhance the local concentration of the activator in the vicinity of the promoter. To our knowledge, this method has not been applied to loops responsible for transcriptional repression, but it might still work for this type of regulation.

Electron microscopy. A direct proof of DNA loop formation is the observation of DNA loops between bound proteins by electron microscopy. In 1986, Griffith et al. (52) took pictures of the DNA loop formed by λ repressor binding to operator sites separated by an integral number of turns of the DNA helix. DNA loops of different sizes have been seen with LacI on *lac* DNA (39). In 1990, striking double loops of octomeric DeoR on native *deo* DNA were observed (22). To be convincing, several control experiments are needed, as well as rigorous statistics. Indeed, DNA loops can form with a certain probability but without any biological significance (13).

Tethered particle motion. A more recent technique is tethered particle motion (TPM) (53), which uses concepts from statistical physics and gives information on a single molecule. A DNA molecule is attached by one of its extremities to a glass plate. The second extremity, to which a ball is attached (for example, a microball of polystyrene with a diameter of 0.2 μm), is free to move. The Brownian motion of the ball is then followed by microscopy over a certain time period. If the DNA molecule is in a looped state, the motion of the ball is modified: its motions are more restricted and its statistical properties (as measured by the amplitude of the fluctuations) have changed. An advantage of this technique is that it can distinguish between different configurations of loops (54). It can also give information concerning the duration of the looped and nonlooped states (55). In concept and in the type of information it can bring, this technique is much more oriented to the thermodynamics of the loop than to effects on gene regulation. This technique has been used to study the antiparallel loops formed by GalR in the repressosome (56).

Atomic force microscopy. Atomic force microscopy (AFM) allows the visualization of the surface topography of a sample. It is based on the physical interactions (as measured by the repulsion of electronic clouds) between the sample surface and a mechanical probe. Lyubchenko et al. (57) analyzed samples of supercoiled minicircles containing the regulatory region of *gal*. They saw asymmetric structures that correspond to the formation of a loop between the two operator sites with GalR in the presence of HU. This technique also requires careful controls to distinguish protein-bound loops from those formed on plectonemic superhelical DNA. AFM was also used to study the regulatory region of *melR* (58), using a linear DNA fragment, and seemed to exclude loop formation in this case.

In this section, we have described several experimental methods to show the presence of a DNA loop inside a regulatory region. It is important to keep in mind that it is necessary to combine several methods to validate the physical interaction and demonstrate a role in gene regulation *in vivo*.

ADVANTAGES OBTAINED VIA DNA LOOPING

Several of the conceivable benefits to the cell from employing the formation of a DNA loop have been described and tested experimentally. Theoretical physics suggests additional advantages that need to be tested, although there are hints to be found in the literature that they do apply *in vivo*.

High local concentration. The major advantage that is gained by the formation of a DNA loop is that it produces a high local concentration at the right place (13, 51, 59, 60, 61). The same level of repression could be achieved by higher cellular levels of the transcription factors, but this runs the risk of allowing nonspecific binding to similar but unrelated sites. The DNA loop permits specific binding at a lower concentration than achieved by a single site. If we take the example of the *lac* operon, the two auxiliary operators will help to increase the concentration of Lac repressor around the principal binding site. Intuitively, the loop keeps the Lac repressor trapped in the vicinity of the principal binding site. In general, transcription factors are only synthesized in small amounts (for example, about 10 LacI tetramers per *E. coli* cell [59]), so it is necessary to find strategies to keep regulatory proteins at the correct place. The advantage offered by a DNA loop is already experimentally proven, as well as theoretically understood. Müller-Hill and colleagues showed that inactivation of one auxiliary binding site produced a significant loss in repression (18, 19). This effect of local concentration is no longer visible if the concentration of LacI repressors is increased, which shows that the formation of a DNA loop is equivalent to an increase in the concentration of regulatory proteins. The effect has now been quantified using the formalism of statistical physics. Vilar and Leibler (17) calculated that adding an auxiliary binding site is equivalent to increasing the effective concentration of repressors per cell, in agreement with the experimental data. This important concept is related to the chelate effect, which explains local concentration effects on enzymatic and intramolecular reactions (61, 62).

Attenuation of fluctuations. It is now well accepted that gene regulation has a stochastic component that can have a crucial impact on many biological processes (63, 64). For example, transcription can be observed to occur in "bursts" within individual cells (65). One of the sources of noise is the discrete and small number of molecules that are involved in the regulation of individual genes. Physicists and biologists have proposed that during the course of evolution, mechanisms developed to attenuate these fluctuations (66, 67). DNA looping has been proposed to have such a property. Vilar and colleagues (17, 68), using stochastic simulations (Gillespie algorithm [69]), showed that DNA looping attenuates the temporal fluctuations during gene expression compared to the fluctuations in a system with a single operator. This interesting result is at the moment a numerical prediction and has not been investigated *in vivo*. Looping allows a fast switching between active and inactive states because the repressor can be quickly recaptured by the main operator and thus maintain a greater repression level for a given level of repressor. If the switching rate is very slow, as in the case of binding to and release from a single operator, there are long periods of time in which mRNA is produced constantly or not at all, which generates larger fluctuations in the expression level.

Another argument in favor of a role of DNA looping in stabilizing a system is that the formation of the loop could prevent the access and the binding of other transcription factors that could interfere with the regulation process. As stated by Michèle Amouyal (70), "loops might insulate a gene and its expression from the genomic environment."

Lower variability inside a clonal population. Another theoretical prediction from the work of Vilar and Leibler (17) is that DNA looping can generate a lower variability inside a clonal population. Indeed, they compared the equations for the repression

level of a system with a single operator and one with two operators forming a DNA loop. They showed that the system with the DNA loop produced a more homogeneous population of gene expression. This difference is due to the fact that in the looping case, the repression level is a nonlinear function of the number of repressors, which decreases the sensitivity of the repression level to variations in the number of repressors. To our knowledge, no work has been undertaken to test this prediction experimentally. However, we think it would be technically possible by means of time-lapse microscopy techniques, as described in reference 71, or with flow cytometry techniques, as described in reference 72. These experiments could measure the fluorescence distribution of a reporter gene (like *gfp*) regulated by a single operator or by two operators that can interact by DNA looping and see if the fluorescence distribution of the looping system is narrower than in the case of DNA looping. (The operators would have to be chosen so that the average expression levels from the two systems are similar.)

Faster search of a target location. It has recently been proposed that the mechanism of DNA looping could accelerate the search by a transcription factor for its target location (73). Indeed, in its search for its target, the transcription factor is proposed to first diffuse in 3 dimensions and then undergo a 1-dimensional (1-D) diffusion along the DNA molecule (sliding). The actual distance that can be covered during the 1-D diffusion is still under debate. Elegant experiments with the restriction enzyme BbvCI show that this distance is 50 bp *in vitro* (74), and a similar distance has been measured for the Lac repressor *in vivo* (75). So if two binding sites are separated by more than this distance, they can be considered independent and DNA looping will facilitate the transfer of the protein from the auxiliary binding site to the principal binding site, which is responsible for the regulation. So auxiliary operators can be seen as "waymarks" for the searching protein, allowing transfer between sites separated by distances greater than the 1-D diffusion limit.

Bistability. The phenomenon of "bistability" is another concept first described by theoretical physics that has now found its place in experimental biology (76). A bistable system can be in two different stable states at a given level of a stimulus, depending upon its history. As illustrated in Figure 2, in the bistability area (in pink in the figure), the system can correspond to either the ON or the OFF state depending on its initial condition (e.g., the presence or absence of an inducer in a preculture). Such a system will not respond gradually according to the intensity of a stimulus but will have an all or none response (77). Since the expression level depends upon its previous state, this kind of system is often linked to the notion of memory (77, 78). It has been proposed that bistable systems are at the origin of the decision for cell differentiation (79).

We would like to suggest that DNA looping could contribute to the formation of a bistable state. Two conditions are necessary to generate bistability: the presence of strong nonlinearities and a positive feedback in the regulatory network (80). The cooperativity produced by DNA looping generates a nonlinearity in the response of the system, so it can facilitate the appearance of bistability. Indeed, for the *lac* operon, the cooperativity introduced by the DNA loop (17, 81) is responsible for the sigmoidicity of the induction curve of the response (82). Bistability was experimentally shown in early work on the *lac* operon (83) (but the term bistability was not used) and, more recently, by Ozbudak et al.

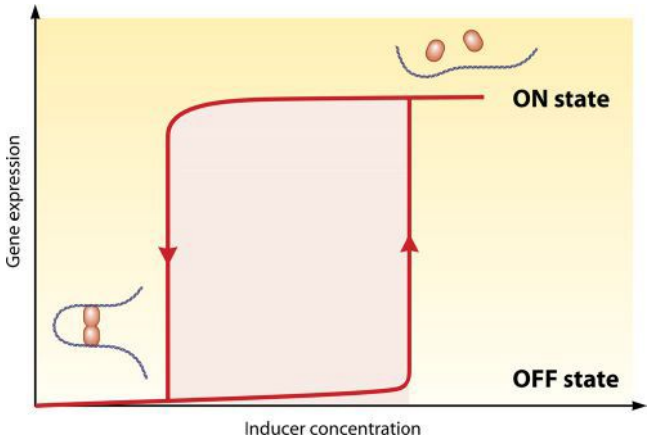


FIG 2 The mechanism of DNA looping could generate the phenomenon of "bistability." The nonlinear response associated with the cooperativity of DNA looping, as well as a positive feedback in the regulation of the system, can generate bistability. Bistable systems can be in one state or another depending on their history, which can have consequences for the biology of the system. The pink area indicates the concentrations of inducer where the gene can be either repressed or expressed depending upon its initial condition (with low inducer or high inducer concentration).

(82), looking at the expression of a *lac-gfp* reporter in individual cells. However, the link between the appearance of bistability and the presence of multiple operators with the formation of a DNA loop has not been directly addressed, to our knowledge. Using the same experimental techniques, it would be interesting to verify whether bistability is observed when one or both of the two auxiliary sites is inactivated or if the dimeric LacI protein is used. Although bistability could arise from other phenomena, (e.g., active transport [82]) and has been observed in bacterial systems where nonlinearities arise from positive autoregulation not invoking DNA loop formation (84), it might be worth looking for multiple operators capable of forming a loop in other systems exhibiting bistability. It is also worth noting that DNA looping systems generally present much larger response factors than non-DNA looping systems. Such systems allow a cell to rapidly pass from a genetic state with low gene expression to a genetic state with very high gene expression. For example, wild-type *lac* operon expression increases by a factor of 1,300 after treatment with the inducer, whereas the construct with just a single O1 operator increases just 18-fold (19). DNA looping appears to be a very practical method of producing a rapid and large-scale biological switch, e.g., when a bacterium has to respond to an environmental signal such as a new carbon source. This is the reason why this type of regulation is widely used in synthetic biology constructs (85).

Robustness to binding site mutations. Morelli et al. (86) used forward flux sampling simulations to analyze the contribution that DNA looping makes to the stability of the bacteriophage lambda (λ) lysogenic state and its resistance to the effects of mutations in the λ repressor operator sites. In their stochastic simulations, they changed the affinities of the λ repressor binding sites to represent a mutation and found that the global functionality of the system (i.e., its ability to change from the lysogenic to the lytic phase) was unaltered when the loop was present. Their results suggest that DNA looping is crucial for stability and provides a mechanism to minimize the effects of operator site mutations. This property of robustness has been experimentally observed in

the λ phage system, by examining the effect of mutations interchanging the OR1 and OR3 operators on the lysogenic/lytic switch (87). Despite large effects on the relative binding of Cro and CI, they found that the phage was still capable of lysogeny and regulated lytic development. A plausible explanation of this robustness is that the cooperativity present in the DNA looping mechanism allows it to compensate when binding sites become less efficient. The analysis by Morelli et al. (86) also showed that DNA looping is necessary for the bistability observed in the bacteriophage λ switch.

ARE THERE OTHER SYSTEMS USING DNA LOOPING TO REGULATE GENE EXPRESSION IN *ESCHERICHIA COLI*?

It seems surprising that few examples of gene regulation involving DNA looping have been reported in the literature. In *E. coli*, only six operons have been experimentally investigated and conclusively shown to be regulated by a mechanism involving a DNA loop (*ara*, *lac*, *gal*, *deo*, *nag*, and *ptsG*), which represents only a small proportion of the regulatory regions. As a first approach to look for more systems potentially regulated by DNA looping, we examined the database RegulonDB (88) for gene regulatory regions that could have the requisite characteristics to form a DNA loop. (We realize that not all the transcription factor [TF] binding sites listed in RegulonDB are experimentally demonstrated, but some have been inferred on the basis of sequence analysis. The data set is, however, an unbiased starting point to look for the possible existence of more looping systems.) Using bioinformatics tools, we selected, from among the 600 regulatory regions present in the database (RegulonDB "TF binding sites" file 2011), the ones that have two binding sites separated by a distance of about 90 bp for the same regulator. This distance is the minimum observed in naturally occurring DNA loops, although smaller loops have been shown to function in artificial *lac* constructs (89). Ninety base pairs is distinctly shorter than the persistence length for DNA, which is about 150 bp for naked DNA (90), as usually measured *in vitro* by cyclization assays (91, 92). To make smaller loops, the curvature energy of the DNA becomes too great and the formation of the loop is not favorable. However, intrinsic curvature in the intervening DNA and/or the presence of additional DNA-bending proteins like CRP or integration host factor (IHF) can facilitate DNA loop formation (46). Garcia et al. (91) have argued that the cellular environment, which means a supercoiled genome and a high concentration of other specific, as well as nucleoid-associated, DNA binding proteins, favors the formation of smaller DNA loops *in vivo*. Moreover, Wiggins et al. (93) showed, by AFM, that spontaneous large-angle bends in short DNA fragments were many times more prevalent than expected from classical models of polymers. In addition, the configuration and flexibility of the protein affects its ability to form a loop (94). More precisely, a protein that can exist in a V-shaped structure (as proposed for LacI [95]) will facilitate the formation of a DNA loop with either a parallel or antiparallel configuration (94).

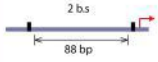
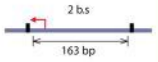
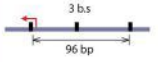
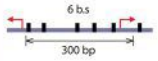
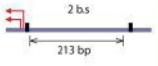
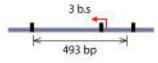
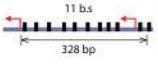
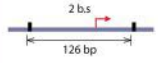
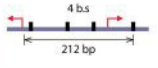
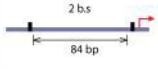
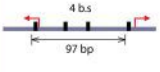
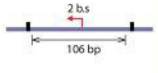
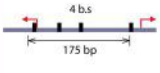
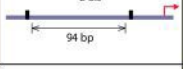
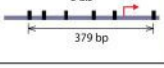
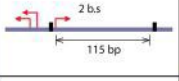
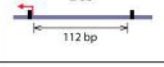
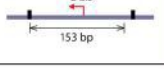
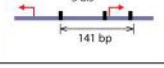
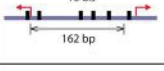
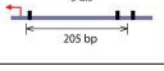
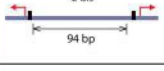
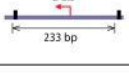
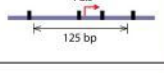
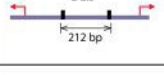
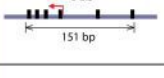
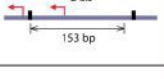
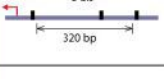
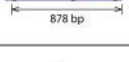
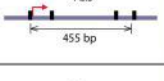
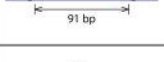
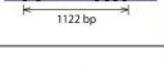
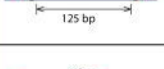
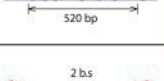
In our bioinformatics survey, we have considered only the simplest case and have looked for those regions with two or more binding sites for the same repressor, obtaining many candidates that are listed in Table 1. We have not looked for examples of loops involving heterologous proteins, (e.g., between σ^{54} RNA polymerase and an EBP [28, 30]). Finally, we have only considered binding sites for a gene-specific regulator; that is to say, we eliminated regions with multiple binding sites for the global transcriptional regulators, i.e., CRP, integration host factor (IHF), FNR, Fis, ArcA, Lrp, and H-NS (118). Indeed, we looked for configurations similar to that of the *lac* operon, involving a specific regulator. We found 48 regions that, according to these criteria, are good candidates for the detection of a DNA loop. We then compared them to the existing literature (Table 1). Only 6 have been convincingly shown to be regulated by loops (Table 1, indicated in boldface). For 16, there are indications in the literature that a loop is involved. Most of the others have not been studied at the molecular level. It is also striking to note that 14 of these systems are part of divergent promoter systems, so the total number of operons potentially regulated by loops is 61; however, in some cases, only one of the genes has been studied, e.g., *exuT-exuC*, where only *exuT* is regulated by ExuR, or *fadL-yfcZ*, where the OmpR sites regulate *fadL*. The possibility of coordinate regulation of both directions in this and other cases should be considered. Divergent promoters can provide an economical target for gene regulation via a DNA loop involving two operator sites.

CONCLUSION AND PERSPECTIVES

DNA looping mechanisms are now well characterized in the regulatory regions of several bacterial operons, either for transcriptional activation or repression. Although this mechanism seems to offer advantages for genetic regulation, relatively few examples have been experimentally investigated. We forecast that loops will be found in other operons of *Escherichia coli* and hope this review will encourage researchers to investigate potential DNA loops in their systems. We think an exhaustive experimental search for such loops (as predicted in Table 1) would be interesting in a well-studied genome like that of *Escherichia coli*. Indeed, cooperative binding of transcription factors has been shown for several operons; e.g., see reference 99. If DNA looping is not the cause of this cooperativity, then it means that some other mechanism involving long-range interactions might be discovered.

Looping at the genome scale? Several biologists and physicists proposed that DNA looping could occur on a genomic scale (16, 119, 120, 121). Bivalent proteins could bind to sites belonging to different genes of a regulon (which could be distant on a 1-D representation of the genome but very close in 3 dimensions). This hypothesis is present in the theoretical work of Buchler et al. (119). The geometrical confinement of the *E. coli* genome implies that regulatory proteins (the same or different) bound to two different regulatory regions can interact and regulate expression. They point out, however, that excessive "cross talk" between regulatory regions could have negative consequences on gene regulation (119). Coregulation of distal genes is the basis of the solenoidal model of the genome imagined by Képès (121). This author proposed that genes regulated by the same transcription factor are placed periodically on the DNA molecule, thus allowing transcription factors to act on several operons located at one locus inside the cell. Other evidence from analysis of transcriptome data also suggested a correlation in the pattern of expression from distant loci (122). The advantages offered by DNA looping can apply at the genome scale and could help optimize transcriptional regulation (121). At this scale, entropy is the major contributor to the energy required for the formation of macro-DNA loops, and specific mechanisms must come into play to overcome the entropy cost (16). On the other hand, it has been proposed that entropy can be the motor of macroloop formation (123, 124). Indeed, taking into account the size of macromolecules like RNA poly-

TABLE 1 Genes or operons in *Escherichia coli* known or predicted to be regulated with a DNA loop responsible for transcriptional repression[a]

Operon name	TF	Configuration	Other TF	ref	Operon name	TF	Configuration	Other TF	ref	Operon name	TF	Configuration	Other TF	ref
aceBAK	IclR		CRP, IHF	[96]	*flhDC*	OmpR		CRP, IHF	[100, 101]	*hybOABC DEFG*	NarL			
agaS	AgaR			[97]	*fliAZY*	NsrR				***lacZYA***	**LacI**		**CRP, H-NS**	**[18, 19]**
agaZ-agaR	AgaR			[97]	*focA*	NarL		CRP, IHF		*lldPRD*	LldR			[110]
araC-araBAD	**AraC**		**CRP**	**[10]**	*ftsQAZ*	SdiA				*manXYZ*	NagC		CRP	[111]
asnA-asnC	AsnC				***galETKM***	**GalR**		**CRP, H-NS, HU**	**[20]**	*melAB-melR*	MelR		CRP	[58, 112]
carAB	PepA		IHF, Fis		*galP*	GalR/S		CRP	[102]	*metR-metE*	MetJ			
chbBCARFG	NagC		CRP	[98]	*galS*	GalS		CRP		*metY-argG*	ArgR		CRP, Fis	
csgD-csgB	CpxR		CRP, IHF	[99]	*gcvTHP*	GcvA			[103]	***nagE-nagBACD***	**NagC**		**CRP**	**[24, 25]**
cspD	Mqs A/R				*gdhA*	Nac		CRP		*nanCM-fimB*	NagC		CRP, IHF	[113]
dcuB, fumB	NarL		CRP		*glmUS*	NagC			[104]	*ompF*	OmpR		CRP, IHF	[114, 115, 116]
deoCABD	**DeoR**		**CRP**	**[22, 23]**	*glpD*	GlpR		CRP	[105]	*phr*	LexA			
edd, eda	GntR				*glpFKX*	GlpR		CRP	[106]	***ptsG***	**Mlc**		**CRP, Fis**	**[117]**
exuT-exuC	ExuR		CRP		*glpT-glpA*	GlpR		CRP, Fis	[107]	*purA*	PurR		CRP, IHF	
fadL-yfcZ	OmpR		CRP		*gntP-uxuA*	UxuR		CRP		*uspA*	FadR		IHF	
fepB-entCEBAH	Fur		CRP		*gntT*	GntR		CRP	[108, 109]	*uxuAB*	UxuR		CRP	
fhuF	OxyR				*hipBA*	HipB		IHF		*ygbA*	NsrR			

[a] Data under the column heads are as follows. TF, transcription factor forming or predicted to form the DNA loop. Configuration, configuration of the binding sites responsible for DNA loop formation: the black boxes show the positions of the repressor binding sites relative to the transcription start site, indicated by a bent arrow, and the total number of binding sites (bs) and the distance between the first and last sites are indicated. Other TF, includes other transcription factors (CRP, H-NS, IHF, Fis, or HU) binding in the regulatory region that could have an impact on DNA loop formation. ref, reference(s) to the literature that validate or suggest regulation involving a DNA loop mechanism for this operon or give evidence against looping. Boldface indicates those operons where the loop has been clearly demonstrated.

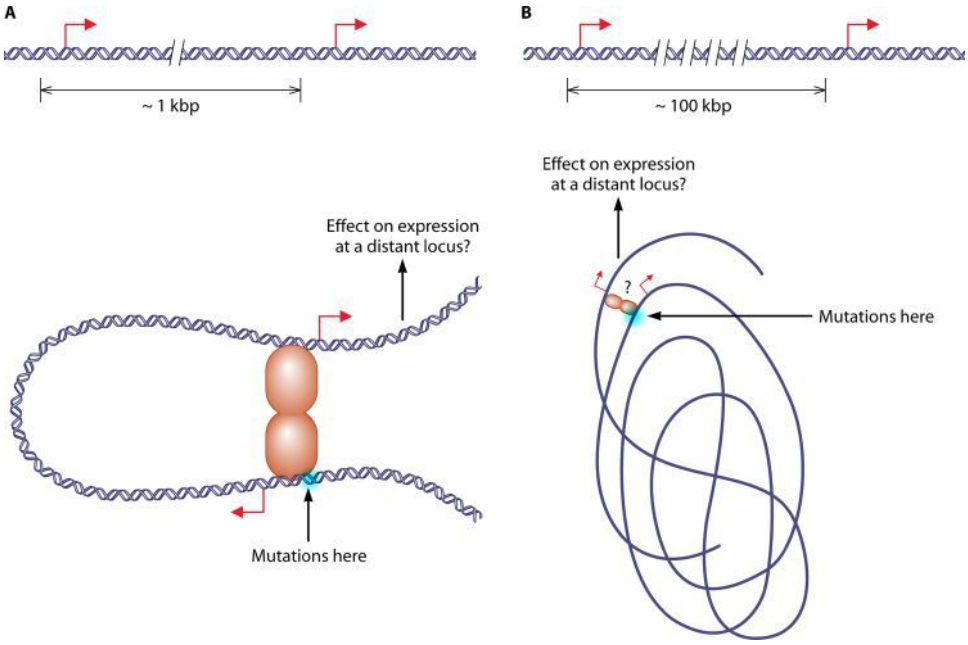


FIG 3 The mechanism of DNA looping could apply to the regulatory regions of two different genes so that binding to distant sites becomes cooperative (16, 121). We can imagine that cooperativity can occur between the regulatory regions of two genes that are relatively close on the genome (separation of the order of 1 kb [120]) (A) or between two genes that are far away in a 1-D representation of the genome (separation of the order of 100 kb) but near in a 3-D one (B). If these kinds of interactions exist, inactivation of binding sites adjacent to one gene would also affect the expression of other genes via the mechanism of DNA looping.

merases and of the crowded nature of the intracellular environment, the global entropy of the system can be minimized by bringing together these large molecules. This forces the grouping of transcription factors, RNA polymerases, and regulated genes into discrete loci, so-called "transcription factories" (125, 126). Computational modeling even suggests that the physical structure of the chromosome is a direct result of regulatory interactions by transcription factors (127). Moreover, the fluorescent probe localization experiments of Wiggins et al. (128) suggest that classical models for DNA being a "wormlike" random coil might not apply to the bacterial genome but that intranucleoid interactions organize the *E. coli* chromosome into a nucleoid filament precisely positioned within the cell.

As illustrated in Figure 3, if loops form between regulatory regions of distally located genes, mutations or deletions in binding sites placed elsewhere in the genome could affect the expression of a distally located gene. Classical molecular biology techniques could be used to test this interesting hypothesis. The development of new biochemical techniques to visualize the 3-D conformation of the eukaryotic genome, like capturing chromosome conformation (3C) (129, 130, 131) and circularized chromosome conformation capture (4C), carbon-copy chromosome conformation capture (5C), and high-throughput 3C (Hi-C), etc. (132, 133, 134), as well as the genetic approach based on site-specific recombination developed in reference 1, offers potential ways to validate the existence of such "long-range" interactions inside a bacterial genome. A very recent 3C study in *E. coli* has demonstrated the existence of long-range interactions between specific DNA binding sites for the GalR repressor in stationary-phase cells (135). This is the first demonstration of a transcription factor organizing the chromosome structure in space at the bacterial genome scale. We believe other examples will be found in the next few years in other bacterial genomes.

The mechanism of DNA looping is one example where we can see how the thermodynamic properties of macromolecules can have a direct impact on gene regulation. Thus, viewing biological systems as thermodynamic systems, coupled with experimental investigation, produces original ideas and information. Finally, we can say that DNA looping is an interesting phenomenon in transcriptional regulation since it represents one of the first examples of "gene regulation in the third dimension" (136). Understanding the phenomenon of DNA looping and its biological and physical implications will bring new insight to our understanding of other biological phenomena, like the 3-D organization of genomes, in prokaryotes and eukaryotes.

ACKNOWLEDGMENT

We are extremely grateful to Annie Kolb for useful suggestions and advice, critical reading of the manuscript, and encouragement.

REFERENCES

1. **Valens M, Penaud S, Rossignol M, Cornet F, Boccard F.** 2004. Macrodomain organization of the *Escherichia coli* chromosome. EMBO J. **23**:4330–4341.
2. **Dame RT, Kalmykowa OJ, Grainger DC.** 2011. Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in Gram negative bacteria. PLoS Genet. **7**:e1002123. doi: 10.1371/journal.pgen.1002123.
3. **Pettijohn D.** 1996. The nucleoid, p 158–166. *In* Neidhardt F. C., Curtiss R., III, Ingraham J. L., Lin E. C. C., Low K. B., Magasanik B., Reznikoff W. S., Riley M., Schaechter M, Umbarger H. E. (ed), *Escherichia coli* and *Salmonella:* cellular and molecular biology, 2nd ed. ASM Press, Washington, DC.
4. **Kavenoff R, Bowen BC.** 1976. Electron microscopy of membrane-free folded chromosomes from *Escherichia coli*. Chromosoma **59**:89–101.
5. **Deng S, Stein RA, Higgins NP.** 2005. Organization of supercoil domains and their reorganization by transcription. Mol. Microbiol. **57**: 1511–1521.
6. **Sinden RR, Carlson JO, Pettijohn DE.** 1980. Torsional tension in the

DNA double helix measured with trimethylpsoralen in living *Escherichia coli* cells: analogous measurements in insect and human cells. Cell **21**: 773–783.
7. **Postow L, Hardy CD, Arsuaga J, Cozzarelli NR.** 2004. Topological domain structure of the *Escherichia coli* chromosome. Genes Dev. **18**: 1766–1779.
8. **Dillon SC, Dorman CJ.** 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat. Rev. Microbiol. **8**:185–195.
9. **Kania J, Müller-Hill B.** 1977. Construction, isolation and implications of repressor-galactosidase-beta-galactosidase hybrid molecules. Eur. J. Biochem. **79**:381–386.
10. **Dunn TM, Hahn S, Ogden S, Schleif RF.** 1984. An operator at −280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. U. S. A. **81**:5017–5020.
11. **Adhya S.** 1989. Multipartite genetic control elements: communication by DNA loop. Annu. Rev. Genet. **23**:227–250.
12. **Matthews KS.** 1992. DNA looping. Microbiol. Rev. **56**:123–136.
13. **Schleif R.** 1992. DNA looping. Annu. Rev. Biochem. **61**:199–223.
14. **Bondarenko VA, Liu YV, Jiang YI, Studitsky VM.** 2003. Communication over a large distance: enhancers and insulators. Biochem. Cell Biol. **81**:241–251.
15. **Saiz L, Vilar JMG.** 2008. *Ab initio* thermodynamic modeling of distal multisite transcription regulation. Nucleic Acids Res. **36**:726–731.
16. **Sumedha, Weigt M.** 2008. A thermodynamic model for the agglomeration of DNA-looping proteins. J. Stat. Mech. **11**:005. doi:10.1088/1742-5468/2008/11/P11005.
17. **Vilar JMG, Leibler S.** 2003. DNA looping and physical constraints on transcription regulation. J. Mol. Biol. **331**:981–989.
18. **Oehler S, Amouyal M, Kolkhof P, von Wilcken-Bergmann B, Müller-Hill B.** 1994. Quality and position of the three *lac* operators of *Escherichia coli* define efficiency of repression. EMBO J. **13**:3348–3355.
19. **Oehler S, Eismann ER, Krämer H, Müller-Hill B.** 1990. The three operators of the *lac* operon cooperate in repression. EMBO J. **9**:973–979.
20. **Haber R, Adhya S.** 1988. Interaction of spatially separated protein-DNA complexes for control of gene expression: operator conversions. Proc. Natl. Acad. Sci. U. S. A. **85**:9683–9687.
21. **Mandal N, Su W, Haber R, Adhya S, Echols H.** 1990. DNA looping in cellular repression of transcription of the galactose operon. Genes Dev. **4**:410–418.
22. **Amouyal M, Mortensen L, Buc H, Hammer K.** 1989. Single and double loop formation when DeoR repressor binds to its natural operator sites. Cell **58**:545–551.
23. **Valentin-Hansen P, Albrechtsen B, Larsen JEL.** 1986. DNA-protein recognition: demonstration of three genetically separated operator elements that are required for repression of the *Escherichia coli deoCABD* promoters by the DeoR repressor. EMBO J. **5**:2015–2021.
24. **Plumbridge J, Kolb A.** 1991. CAP and Nag repressor binding to the regulatory regions of the *nagE-B* and *manX* genes of *Escherichia coli*. J. Mol. Biol. **217**:661–679.
25. **Plumbridge J, Kolb A.** 1993. DNA loop formation between Nag repressor molecules bound to its two operator sites is necessary for repression of the *nag* regulon of *Escherichia coli in vivo*. Mol. Microbiol. **10**:973–981.
26. **Kustu S, North AK, Weiss DS.** 1991. Prokaryotic transcriptional enhancers and enhancer-binding proteins. Trends Biochem. Sci. **16**:397–402.
27. **Reitzer LJ, Magasanik B.** 1986. Transcription of *glnA* in *Escherichia coli* is stimulated by activator bound to sites far from the promoter. Cell **45**:785–792.
28. **Ghosh T, Bose D, Zhang X.** 2010. Mechanisms for activating bacterial RNA polymerase. FEMS Microbiol. Rev. **34**:611–627.
29. **Bush M, Dixon R.** 2012. The role of bacterial enhancer binding proteins as specialized activators of σ^{54}-dependent transcription. Microbiol. Mol. Biol. Rev. **76**:497–529.
30. **Morett E, Segovia L.** 1993. The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J. Bacteriol. **175**:6067–6074.
31. **Studholme DJ, Dixon R.** 2003. Domain architectures of σ^{54}-dependent transcriptional activators. J. Bacteriol. **185**:1757–1767.
32. **Hirsh AD, Lillian TD, Lionberger TA, Perkins NC.** 2011. DNA modeling reveals an extended *lac* repressor conformation in classic in vitro binding assays. Biophys. J. **101**:718–726.
33. **Rutkauskas D, Zhan H, Matthews KS, Pavone FS, Vanzi F.** 2009. Tetramer opening in LacI-mediated DNA looping. Proc. Natl. Acad. Sci. U. S. A. **106**:16627–16632.
34. **Saiz L, Vilar JMG.** 2006. DNA looping: the consequences and its control. Curr. Opin. Struct. Biol. **16**:344–350.
35. **Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Kuhlman T, Phillips R.** 2005. Transcriptional regulation by the numbers: applications. Curr. Opin. Genet. Dev. **15**:125–135.
36. **Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R.** 2005. Transcriptional regulation by the numbers: models. Curr. Opin. Genet. Dev. **15**:116–124.
37. **Gerland U, Moroz JD, Hwa T.** 2002. Physical constraints and functional characteristics of transcription factor-DNA interaction. Proc. Natl. Acad. Sci. U. S. A. **99**:12015–12020.
38. **Majumdar A, Adhya S.** 1984. Demonstration of two operator elements in *gal*: *in vitro* repressor binding studies. Proc. Natl. Acad. Sci. U. S. A. **81**:6100–6104.
39. **Krämer H, Niemöller M, Amouyal M, Revet B, von Wilcken-Bergmann B, Müller-Hill B.** 1987. Lac repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. EMBO J. **6**:1481–1491.
40. **Hochschild A, Ptashne M.** 1986. Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. Cell **44**:681–687.
41. **Suck D, Oefner C.** 1986. Structure of DNase I at 2.0 Angstrom resolution suggests a mechanism for binding to and cutting DNA. Nature **321**: 620–625.
42. **Lipfert J, Kerssemakers JWJ, Jager T, Dekker NH.** 2010. Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments. Nat. Methods **7**:977–980.
43. **Horowitz DS, Wang JC.** 1984. Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. J. Mol. Biol. **173**:75–91.
44. **Frank DE, Saecker RM, Bond JP, Capp MW, Tsodikov OV, Melcher SE, Levandoski MM, Record MT.** 1997. Thermodynamics of the interactions of Lac repressor with variants of the symmetric *lac* operator: effects of converting a consensus site to a non-specific site. J. Mol. Biol. **267**:1186–1206.
45. **Morozov AV, Havranek JJ, Baker D, Siggia ED.** 2005. Protein-DNA binding specificity predictions with structural models. Nucleic Acids Res. **33**:5781–5798.
46. **Plumbridge J, Kolb A.** 1998. DNA bending and expression of the divergent *nagE-B* operons. Nucleic Acids Res. **26**:1254–1260.
47. **Mehta RA, Kahn JD.** 1999. Designed hyperstable Lac repressor.DNA loop topologies suggest alternative loop geometries. J. Mol. Biol. **294**: 67–77.
48. **Reed WL, Schleif RF.** 1999. Hemiplegic mutations in AraC protein. J. Mol. Biol. **294**:417–425.
49. **Aki T, Choy HE, Adhya S.** 1996. Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of *gal* transcription by Gal repressor. Genes Cells **1**:179–188.
50. **Semsey S, Tolstorukov MY, Virnik K, Zhurkin VB, Adhya S.** 2004. DNA trajectory in the *gal* repressosome. Genes Dev. **18**:1898–1907.
51. **Wedel A, Weiss DS, Popham D, Dröge P, Kustu S.** 1990. A bacterial enhancer functions to tether a transcriptional activator near a promoter. Science **248**:486–490.
52. **Griffith J, Hochschild A, Ptashne M.** 1986. DNA loops induced by cooperative binding of lambda repressor. Nature **322**:750–752.
53. **Schafer DA, Gelles J, Sheetz MP, Landick R.** 1991. Transcription by single molecules of RNA polymerase observed by light microscopy. Nature **352**:444–448.
54. **Towles KB, Beausang JF, Garcia HG, Phillips R, Nelson PC.** 2009. First-principles calculation of DNA looping in tethered particle experiments. Phys. Biol. **6**:25001. doi:10.1088/1478-3975/6/2/025001.
55. **Pouget N, Turlan C, Destainville N, Salomé L, Chandler M.** 2006. IS911 transpososome assembly as analysed by tethered particle motion. Nucleic Acids Res. **34**:4313–4323.
56. **Lia G, Semsey S, Lewis DEA, Adhya S, Bensimon D, Dunlap D, Finzi L.** 2008. The antiparallel loops in *gal* DNA. Nucleic Acids Res. **36**:4204–4210.
57. **Lyubchenko YL, Shlyakhtenko LS, Aki T, Adhya S.** 1997. Atomic force microscopic demonstration of DNA looping by GalR and HU. Nucleic Acids Res. **25**:873–876.
58. **Samarasinghe S, El-Robh MS, Grainger DC, Zhang W, Soultanas P, Busby SJW.** 2008. Autoregulation of the *Escherichia coli melR* promoter:

repression involves four molecules of MelR. Nucleic Acids Res. **36**:2667–2676.
59. **Müller-Hill B.** 1998. The function of auxiliary operators. Mol. Microbiol. **29**:13–18.
60. **Müller-Hill B.** 2006. What is life? The paradigm of DNA and protein cooperation at high local concentrations. Mol. Microbiol. **60**:253–255.
61. **Oehler S, Müller-Hill B.** 2010. High local concentration: a fundamental strategy of life. J. Mol. Biol. **395**:242–253.
62. **Page MI, Jencks WP.** 1971. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. Proc. Natl. Acad. Sci. U. S. A. **68**:1678–1683.
63. **Rao CV, Wolf DM, Arkin AP.** 2002. Control, exploitation and tolerance of intracellular noise. Nature **420**:231–237.
64. **Cox CD, McCollum JM, Austin DW, Allen MS, Dar RD, Simpson ML.** 2006. Frequency domain analysis of noise in simple gene circuits. Chaos **16**:026102. http://dx.doi.org/10.1063/1.2204354.
65. **Golding I, Paulsson J, Zawilski SM, Cox EC.** 2005. Real-time kinetics of gene activity in individual bacteria. Cell **123**:1025–1036.
66. **Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS.** 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. Nature **441**:840–846.
67. **Wang Z, Zhang J.** 2011. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proc. Natl. Acad. Sci. U. S. A. **108**:E67–E76.
68. **Vilar JMG, Saiz L.** 2005. DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. Curr. Opin. Genet. Dev. **15**:136–144.
69. **Gillespie D.** 1977. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. **81**:2340–2361.
70. **Amouyal M.** 2005. Gene regulation at-a-distance in *Escherichia coli*: new insights. C. R. Biol. **328**:1–9.
71. **Becskei A, Serrano L.** 2000. Engineering stability in gene networks by autoregulation. Nature **405**:590–593.
72. **Ansel J, Bottin H, Rodriguez-Beltran C, Damon C, Nagarajan M, Fehrmann S, François J, Yvert G.** 2008. Cell-to-cell stochastic variation in gene expression is a complex genetic trait. PLoS Genet. **4**:e1000049. doi:10.1371/journal.pgen.1000049.
73. **Li GW, Berg OG, Elf J.** 2009. Effects of macromolecular crowding and DNA looping on gene regulation kinetics. Nat. Phys. **5**:294–297.
74. **Gowers DM, Wilson GG, Halford SE.** 2005. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. Proc. Natl. Acad. Sci. U. S. A. **102**:15883–15888.
75. **Hammar P, Leroy P, Mahmutovic A, Marklund EG, Berg OG, Elf J.** 2012. The Lac repressor displays facilitated diffusion in living cells. Science **336**:1595–1598.
76. **Kramer BP, Fussenegger M.** 2005. Hysteresis in a synthetic mammalian gene network. Proc. Natl. Acad. Sci. U. S. A. **102**:9517–9522.
77. **Gardner TS, Cantor CR, Collins JJ.** 2000. Construction of a genetic toggle switch in *Escherichia coli*. Nature **403**:339–342.
78. **Chang DE, Leung S, Atkinson MR, Reifler A, Forger D, Ninfa AJ.** 2010. Building biological memory by linking positive feedback loops. Proc. Natl. Acad. Sci. U. S. A. **107**:175–180.
79. **Xiong W, Ferrell JE.** 2003. A positive-feedback-based bistable "memory module" that governs a cell fate decision. Nature **426**:460–465.
80. **Thomas R, Kaufman M.** 2001. Multistationarity, the basis of cell differentiation and memory. ii. logical analysis of regulatory networks in terms of feedback circuits. Chaos **11**:180–195.
81. **Oehler S, Alberti S, Müller-Hill B.** 2006. Induction of the *lac* promoter in the absence of DNA loops and the stoichiometry of induction. Nucleic Acids Res. **34**:606–612.
82. **Ozbudak EM, Thattai M, Lim HN, Shraiman BI, van Oudenaarden A.** 2004. Multistability in the lactose utilization network of *Escherichia coli*. Nature **427**:737–740.
83. **Novick A, Weiner M.** 1957. Enzyme induction as an all-or-none phenomenon. Proc. Natl. Acad. Sci. U. S. A. **43**:553–566.
84. **Dubnau D, Losick R.** 2006. Bistability in bacteria. Mol. Microbiol. **61**:564–572.
85. **Atkinson MR, Savageau MA, Myers JT, Ninfa AJ.** 2003. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. Cell **113**:597–607.
86. **Morelli MJ, Wolde PRT, Allen RJ.** 2009. DNA looping provides stability and robustness to the bacteriophage lambda switch. Proc. Natl. Acad. Sci. U. S. A. **106**:8101–8106.
87. **Little JW, Shepley DP, Wert DW.** 1999. Robustness of a gene regulatory circuit. EMBO J. **18**:4299–4307.
88. **Gama-Castro S, Jiménez-Jacinto V, Peralta-Gil M, Santos-Zavaleta A, Peñaloza-Spinola MI, Contreras-Moreira B, Segura-Salazar J, Muñiz-Rascado L, Martínez-Flores I, Salgado H, Bonavides-Martínez C, Abreu-Goodger C, Rodríguez-Penagos C, Miranda-Ríos J, Morett E, Merino E, Huerta AM, Treviño-Quintanilla L, Collado-Vides J.** 2008. RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. Nucleic Acids Res. **36**:D120–D124.
89. **Müller J, Oehler S, Müller-Hill B.** 1996. Repression of *lac* promoter as a function of distance, phase and quality of an auxiliary *lac* operator. J. Mol. Biol. **257**:21–29.
90. **Hagerman PJ.** 1988. Flexibility of DNA. Annu. Rev. Biophys. Biophys. Chem. **17**:265–286.
91. **Garcia HG, Grayson P, Han L, Inamdar M, Kondev J, Nelson PC, Phillips R, Widom J, Wiggins PA.** 2007. Biological consequences of tightly bent DNA: the other life of a macromolecular celebrity. Biopolymers **85**:115–130.
92. **Geggier S, Vologodskii A.** 2010. Sequence dependence of DNA bending rigidity. Proc. Natl. Acad. Sci. U. S. A. **107**:15421–15426.
93. **Wiggins PA, van der Heijden T, Moreno-Herrero F, Spakowitz A, Phillips R, Widom J, Dekker C, Nelson PC.** 2006. High flexibility of DNA on short length scales probed by atomic force microscopy. Nat. Nanotechnol. **1**:137–141.
94. **Swigon D, Coleman BD, Olson WK.** 2006. Modeling the *lac* repressor-operator assembly: the influence of DNA looping on Lac repressor conformation. Proc. Natl. Acad. Sci. U. S. A. **103**:9879–9884.
95. **Lewis M.** 2005. The *lac* repressor. C. R. Biol. **328**:521–548.
96. **Yamamoto K, Ishihama A.** 2003. Two different modes of transcription repression of the *Escherichia coli* acetate operon by IclR. Mol. Microbiol. **47**:183–194.
97. **Ray WK, Larson TJ.** 2004. Application of AgaR repressor and dominant repressor variants for verification of a gene cluster involved in N-acetylgalactosamine metabolism in *Escherichia coli K-12*. Mol. Microbiol. **51**:813–826.
98. **Plumbridge J, Pellegrini O.** 2004. Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP. Mol. Microbiol. **52**:437–449.
99. **Jubelin G, Vianney A, Beloin C, Ghigo JM, Lazzaroni JC, Lejeune P, Dorel C.** 2005. CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in *Escherichia coli*. J. Bacteriol. **187**:2038–2049.
100. **Prüss BM, Besemann C, Denton A, Wolfe AJ.** 2006. A complex transcription network controls the early stages of biofilm development by *Escherichia coli*. J. Bacteriol. **188**:3731–3739.
101. **Shin S, Park C.** 1995. Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. J. Bacteriol. **177**:4696–4702.
102. **El Qaidi S, Allemand F, Oberto J, Plumbridge J.** 2009. Repression of *galP*, the galactose transporter in *Escherichia coli*, requires the specific regulator of N-acetylglucosamine metabolism. Mol. Microbiol. **71**:146–157.
103. **Stauffer LT, Stauffer GV.** 1998. Spacing and orientation requirements of GcvA-binding sites 3 and 2 and the Lrp-binding region for *gcvT::lacZ* expression in *Escherichia coli*. Microbiology **144**(Pt 5):1417–1422.
104. **Plumbridge J.** 1995. Co-ordinated regulation of amino sugar biosynthesis and degradation: the NagC repressor acts as both an activator and a repressor for the transcription of the *glmUS* operon and requires two separated NagC binding sites. EMBO J. **14**:3958–3965.
105. **Yang B, Larson TJ.** 1996. Action at a distance for negative control of transcription of the *glpD* gene encoding sn-glycerol 3-phosphate dehydrogenase of *Escherichia coli* K-12. J. Bacteriol. **178**:7090–7098.
106. **Weissenborn DL, Wittekindt N, Larson TJ.** 1992. Structure and regulation of the *glpFK* operon encoding glycerol diffusion facilitator and glycerol kinase of *Escherichia coli* K-12. J. Biol. Chem. **267**:6122–6131.
107. **Yang B, Gerhardt SG, Larson TJ.** 1997. Action at a distance for *glp* repressor control of *glpTQ* transcription in *Escherichia coli* K-12. Mol. Microbiol. **24**:511–521.
108. **Peekhaus N, Conway T.** 1998. Positive and negative transcriptional regulation of the *Escherichia coli* gluconate regulon gene *gntT* by GntR and the cyclic AMP (cAMP)-cAMP receptor protein complex. J. Bacteriol. **180**:1777–1785.
109. **Porco A, Peekhaus N, Bausch C, Tong S, Isturiz T, Conway T.** 1997.

Molecular genetic characterization of the *Escherichia coli gntT* gene of *GntI*, the main system for gluconate metabolism. J. Bacteriol. **179**:1584–1590.
110. **Aguilera L, Campos E, Giménez R, Badía J, Aguilar J, Baldoma L.** 2008. Dual role of LldR in regulation of the *lldPRD* operon, involved in L-lactate metabolism in *Escherichia coli*. J. Bacteriol. **190**:2997–3005.
111. **Plumbridge J.** 1998. Control of the expression of the *manXYZ* operon in *Escherichia coli*: Mlc is a negative regulator of the mannose PTS. Mol. Microbiol. **27**:369–380.
112. **Kahramanoglou C, Webster CL, El-Robh MS, Belyaeva TA, Busby SJW.** 2006. Mutational analysis of the *Escherichia coli melR* gene suggests a two-state concerted model to explain transcriptional activation and repression in the melibiose operon. J. Bacteriol. **188**:3199–3207.
113. **Sohanpal BK, El-Labany S, Lahooti M, Plumbridge JA, Blomfield IC.** 2004. Integrated regulatory responses of *fimB* to N-acetylneuraminic (sialic) acid and GlcNAc in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. U. S. A. **101**:16322–16327.
114. **Rampersaud A, Harlocker SL, Inouye M.** 1994. The OmpR protein of *Escherichia coli* binds to sites in the *ompF* promoter region in a hierarchical manner determined by its degree of phosphorylation. J. Biol. Chem. **269**:12559–12566.
115. **Huang KJ, Schieberl JL, Igo MM.** 1994. A distant upstream site involved in the negative regulation of the *Escherichia coli ompF* gene. J. Bacteriol. **176**:1309–1315.
116. **Bergstrom LC, Qin L, Harlocker SL, Egger LA, Inouye M.** 1998. Hierarchical and co-operative binding of OmpR to a fusion construct containing the *ompC* and *ompF* upstream regulatory sequences of *Escherichia coli*. Genes Cells **3**:777–788.
117. **El Qaidi S, Plumbridge J.** 2008. Switching control of expression of *ptsG* from the Mlc regulon to the NagC regulon. J. Bacteriol. **190**:4677–4686.
118. **Martínez-Antonio A, Collado-Vides J.** 2003. Identifying global regulators in transcriptional regulatory networks in bacteria. Curr. Opin. Microbiol. **6**:482–489.
119. **Buchler NE, Gerland U, Hwa T.** 2003. On schemes of combinatorial transcription logic. Proc. Natl. Acad. Sci. U. S. A. **100**:5136–5141.
120. **Dandanell G, Valentin-Hansen P, Larsen JE, Hammer K.** 1987. Long-range cooperativity between gene regulatory sequences in a prokaryote. Nature **325**:823–826.
121. **Képès F.** 2004. Periodic transcriptional organization of the *Escherichia coli* genome. J. Mol. Biol. **340**:957–964.
122. **Jeong KS, Ahn J, Khodursky AB.** 2004. Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. Genome Biol. **5**:R86. doi:10.1186/gb-2004-5-11-r86.
123. **Marenduzzo D, Faro-Trindade I, Cook PR.** 2007. What are the molecular ties that maintain genomic loops? Trends Genet. **23**:126–133.
124. **Marenduzzo D, Micheletti C, Cook PR.** 2006. Entropy-driven genome organization. Biophys. J. **90**:3712–3721.
125. **Cook PR.** 2002. Predicting three-dimensional genome structure from transcriptional activity. Nat. Genet. **32**:347–352.
126. **Cook PR.** 2010. A model for all genomes: the role of transcription factories. J. Mol. Biol. **395**:1–10.
127. **Fritsche M, Li S, Heermann DW, Wiggins PA.** 2012. A model for *Escherichia coli* chromosome packaging supports transcription factor-induced DNA domain formation. Nucleic Acids Res. **40**:972–980.
128. **Wiggins PA, Cheveralls KC, Martin JS, Lintner R, Kondev J.** 2010. Strong intranucleoid interactions organize the *Escherichia coli* chromosome into a nucleoid filament. Proc. Natl. Acad. Sci. U. S. A. **107**:4991–4995.
129. **Dekker J, Rippe K, Dekker M, Kleckner N.** 2002. Capturing chromosome conformation. Science **295**:1306–1311.
130. **Dostie J, Zhan Y, Dekker J.** 2007. Chromosome conformation capture carbon copy technology. Curr. Protoc. Mol. Biol. **Chapter 21**:Unit 21.14. doi:10.1002/0471142727.mb2114s80.
131. **Wang W, Li GW, Chen C, Xie XS, Zhuang X.** 2011. Chromosome organization by a nucleoid-associated protein in live bacteria. Science **333**:1445–1449.
132. **Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, Lee C, Shendure J, Fields S, Blau CA, Noble WS.** 2010. A three-dimensional model of the yeast genome. Nature **465**:363–367.
133. **Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J.** 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science **326**:289–293.
134. **Umbarger MA, Toro E, Wright MA, Porreca GJ, Baù D, Hong SH, Fero MJ, Zhu LJ, Marti-Renom MA, McAdams HH, Shapiro L, Dekker J, Church GM.** 2011. The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. Mol. Cell **44**:252–264.
135. **Qian Z, Dimitriadis EK, Edgar R, Eswaramoorthy P, Adhya S.** 2012. Galactose repressor mediated intersegmental chromosomal connections in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **109**:11336–11340.
136. **Dekker J.** 2008. Gene regulation in the third dimension. Science **319**:1793–1794.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

MINIREVIEW

Transcription Regulation in Archaea

Alexandra M. Gehring, Julie E. Walker, Thomas J. Santangelo

Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado, USA

The known diversity of metabolic strategies and physiological adaptations of archaeal species to extreme environments is extraordinary. Accurate and responsive mechanisms to ensure that gene expression patterns match the needs of the cell necessitate regulatory strategies that control the activities and output of the archaeal transcription apparatus. *Archaea* are reliant on a single RNA polymerase for all transcription, and many of the known regulatory mechanisms employed for archaeal transcription mimic strategies also employed for eukaryotic and bacterial species. Novel mechanisms of transcription regulation have become apparent by increasingly sophisticated *in vivo* and *in vitro* investigations of archaeal species. This review emphasizes recent progress in understanding archaeal transcription regulatory mechanisms and highlights insights gained from studies of the influence of archaeal chromatin on transcription.

RNA polymerase (RNAP) is a well-conserved, multisubunit essential enzyme that transcribes DNA to generate RNA in all cells. Although RNA synthesis is carried out by RNAP, the activities of RNAP during each phase of transcription are subject to basal and regulatory transcription factors. Substantial differences in transcription regulatory strategies exist in the three domains (*Bacteria*, *Archaea*, and *Eukarya*). Only a single transcription factor (NusG or Spt5) is universally conserved (1, 2), and the roles of many archaeon-encoded factors have not been evaluated using *in vivo* and *in vitro* techniques. *Archaea* are reliant on a transcription apparatus that is homologous to the eukaryotic transcription machinery; similarities include additional RNAP subunits that form a discrete subdomain of RNAP (3, 4) as well as basal transcription factors that direct transcription initiation and elongation (5–8). The shared homology of archaeal-eukaryotic transcription components aligns with the shared ancestry of *Archaea* and *Eukarya*, and this homology often is exclusive of *Bacteria*. *Archaea* are prokaryotic, but the transcription apparatus of *Bacteria* differs significantly from that of *Archaea* and *Eukarya*.

The archaeal transcription apparatus is most commonly summarized as a simplified version of the eukaryotic machinery. In some respects this is true, as homologs of only a few eukaryotic transcription factors are encoded in archaeal genomes, and archaeal transcription *in vitro* can be supported by just a few transcription factors. However, much regulatory activity in eukaryotes is devoted to posttranslational modifications of chromatin, RNAP, and transcription factors, and this complexity seemingly does not transfer to the *Archaea*, where few posttranslational modifications or chromatin-imposed regulation events are currently known. The ostensible simplicity of archaeal transcription is under constant revision, as more detailed examinations of archaeon-encoded factors become possible through increasingly sophisticated *in vivo* and *in vitro* techniques. This review will highlight the current understanding of archaeal transcription, emphasizing the roles of factors that regulate archaeal RNAP throughout each stage of the transcription cycle and also highlighting outstanding issues in the field.

THE ARCHAEAL TRANSCRIPTION CYCLE

Transcription is highly regulated, and the transcription cycle is typically demarcated into three phases: initiation, elongation, and termination (9–13) (Fig. 1). An abbreviated and overall introduction to this cycle is presented first, with sections below detailing the activities of RNAP and associated factors during each stage of transcription. Briefly, archaeal transcription initiation requires that RNAP be directed to promoter sequences defined by the binding of TATA binding protein (TBP) and transcription factor B (TFB). TBP, TFB, and RNAP are sufficient to generate a single-stranded section of DNA (the transcription bubble) and feed the template strand into the bipartite active center of RNAP (7, 14). RNAP can initiate transcript synthesis *de novo*, and continued synthesis then competes with favorable promoter and initiation factor contacts until promoter escape can be achieved. Release of RNAP from the initiating factors classically defines the end of initiation, although in reality no clear boundary separates the last stages of initiation from the early stages of elongation. Although TFB and TBP are necessary and sufficient to permit promoter-directed transcription initiation, a third conserved factor, transcription factor E (TFE), can also assist in transcription initiation and leaves the promoter with RNAP during the early stages of transcript elongation (15–18). Transition to a stable, long-lived elongation complex is believed to involve internal rearrangements of RNAP. This transition involves the exchange of initiation factors for stably bound elongation factors that monitor RNA synthesis for accuracy, respond to regulatory DNA sequences, react to regulatory inputs of more transiently associated transcription factors, and influence processivity of RNAP. Elongation is, in general, very stable, but specific sequences can lower the overall energy of the transcription elongation complex, permitting either spontaneous intrinsic or factor-assisted termination (19, 20). Transcription termination results in release of both the transcript and RNAP from the DNA template.

Accepted manuscript posted online 2 May 2016

Citation Gehring AM, Walker JE, Santangelo TJ. 2016. Transcription regulation in archaea. J Bacteriol 198:1906–1917. doi:10.1128/JB.00255-16.

Editor: W. Margolin, University of Texas Medical School at Houston

Address correspondence to Thomas J. Santangelo, thomas.santangelo@colostate.edu.

A.M.G. and J.E.W. contributed equally to this work.

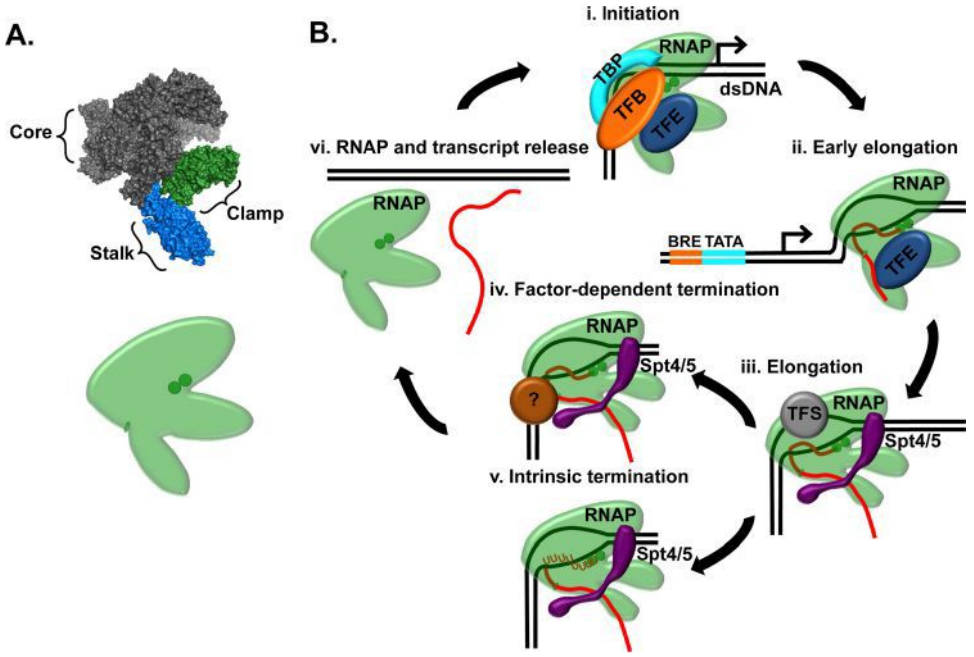


FIG 1 The archaeal transcription cycle. (A) The euryarchaeal RNA polymerase crystal structure from *Thermococcus kodakarensis* (PDB ID no. 4QIW) is shown in a surface representation. The clamp and stalk domains are highlighted. A simplified cartoon structure of RNA polymerase is shown below this in light green; the bipartite active site and RNA exit channel are highlighted in dark green. (B) Steps in the transcription cycle. (i) RNAP is recruited to the promoter by transcription factors TFB, TFE, and TBP during transcription initiation. (ii) RNAP escapes the promoter, and early elongation begins with TFE bound to RNAP. (iii) TFE is replaced by elongation factor Spt5 during elongation. (iv) Factor-dependent termination is predicted to occur in archaea by an unknown factor. (v) Intrinsic termination sequences are characterized by a run of T's on the nontemplate strand. (vi) The transcript is released, and RNAP is recycled for another round of transcription.

REGULATED TRANSCRIPTION INITIATION

Transcription initiation is tightly regulated by both transcription factors and DNA elements. The minimal, necessary proteins and DNA elements for archaeal transcription initiation are now well defined and characterized (21–28). A recent excellent review (29) summarizes the actions of repressors and activators that function during initiation in archaeal species. We focus here on the roles of new DNA elements and newly discovered strategies of basal initiation factors.

BASAL TRANSCRIPTION FACTORS

TBP and TFB are the only transcription factors required for *in vitro* transcription under optimized conditions, and TFE has been shown to assist promoter opening when conditions are suboptimal (16). *In vivo* studies have shown that *Archaea* must retain at least one gene encoding TBP and one gene encoding TFB, although many archaeal species encode multiple TBP and TFB isoforms (6, 21, 30–35). Some differences in promoter sequence preferences and protein pairing have been noted in TBP-TFB isoform pairs (36–41), but these minor differences are not on par with the clear but not always radical promoter sequence differences noted for alternative σ factors in bacterial transcription (39, 42). TFE also appears essential, and it is currently unclear if this essentiality is due to necessary activities during transcription initiation or some other role in the transcription cycle (26, 43, 44).

All three of the aforementioned transcription factors have close eukaryotic homologs: archaeal TBPs are nearly identical to eukaryotic TBPs (45); archaeal TFB proteins are homologous to eukaryotic transcription factor IIB (TFIIB) proteins (46), with homology also seen with the Pol III initiation factor BRF1 (47) and Pol I initiation factor Rrn7/TAF1B (48); and archaeal TFE proteins are homologous to the N-terminal half of the eukaryotic alpha subunit of TFIIE, or TFIIEα, and very recent evidence identified a separate homolog in some lineages to the eukaryotic beta subunit, TFIIEβ (17). TBP is needed to recognize the TATA box, bend the DNA, and recruit TFB (46); its role had therefore been deemed equivalent to the role of eukaryotic TBPs. Recent, sophisticated total internal-reflection fluorescence–fluorescence resonance energy transfer measurements now detail differences in the activities of archaeal and eukaryotic TBPs, despite the nearly identical three-dimensional folds of these factors (6). In some cases, archaeal TBPs require the cobinding of TFB to stably bind and bend the promoter DNA (6, 22, 49, 50). It is tempting to speculate that different promoter sequences may be regulated by different TFB-TBP pairs based on the interdependence, or lack thereof, of cooperative DNA bending for establishing a stable platform for RNAP recruitment. Recent studies suggest that select isoforms of TFB and TBP can result in differences in transcription output, but further studies will be needed to determine if these effects on such preliminary steps of transcription initiation are a direct mode of regulation resulting in phenotypic differences (37, 51).

In contrast to eukaryotic transcription, archaeal promoter opening is not an energy-dependent process (7). Therefore, TBP and TFB alone are capable of assisting RNAP in the formation of the transcription bubble. In all *Archaea*, TFB is responsible for stabilizing the TBP-bound DNA complex and, together, this bipartite protein platform recruits RNAP (52), but how these molecular interactions melt the DNA is still unresolved. Reconstruc-

tions and analyses of open complexes using archaeal components reveal an overall architecture of the open promoter complex and provide the first placement of the nontemplate strand within the complex (52). TBP and TFB are located closer to RNAP than would be the case for eukaryotic promoters, and this proximity may provide more intimate contacts that collectively provide the energy to open the promoter DNA. The tight network of interactions in the archaeal open complex may torsionally strain the DNA, and melting is likely to relieve this strain and result in open complex formation.

Several new insights into TFE activity and evolution have been recently described. The archaeal TFE had previously been characterized as a monomer and as a homologue of the alpha subunit of eukaryotic TFIIE, termed TFIIEα (16, 18, 53). Eukaryotic TFIIE is a heterodimeric complex of TFIIEα and TFIIEβ, but archaeal genomes had previously only been shown to encode a homologue of only the alpha subunit (54, 55). Eukaryotic RNAPs differ in their requirements for initiation, with RNAP III incorporating homologues of several RNAP II initiation factors as core components of RNAP III (56–58). Comparisons of the RNAP III subunit hRCP39 revealed a well-conserved archaeal homolog (termed TFEβ) that directly and extensively interacts with TFE (now named TFEα) (17). Although TFEβ is not conserved in all *Archaea*, TFEβ is essential for some *Crenarchaea*; when employed *in vitro*, TFEα-TFEβ complexes are effective in binding RNAP, stabilizing open complex formation, and stimulating total transcriptional output (17).

The mechanism of TFE recruitment to the initiation complex and its activities during initiation has been partially resolved. TFEα simultaneously binds TBP, RNAP, and downstream DNA and has been shown to stimulate transcription at noncanonical promoter sequences and at reduced temperatures *in vitro* (16, 18, 59). Several studies have identified critical interactions between TFE and the preinitiation complex that have furthered our understanding of TFE function during initiation (2, 15, 26, 53, 59). TFEα consists of two domains: a winged helix (WH) domain and a zinc ribbon domain (60, 61); TFEβ contains a conserved WH domain and an FeS domain (17). The WH domain of TFEα contacts the upstream, nontemplate strand of DNA and helps form the open promoter complex through an unknown mechanism (15, 52). Several studies have shown that the presence of the RNAP stalk domain—unique to archaeoeukaryotic RNAPs and comprised of two subunits, RpoE and RpoF in archaea and Rpo4 and Rpo7 in eukaryotes—is essential for the full activity of TFEα (59, 62, 63). The predicted interaction between TFEα and the stalk domain was bolstered by copurification of TFEα with intact RNAP and the loss of TFEα from RNAP preparations wherein the stalk domain was missing (44). A recent structure-function study identified critical interactions between TFEα and RpoE of the stalk domain (26). TFE may have an essential role in modulating intramolecular movements of RNAP during the transcription cycle, most notably movements of the clamp domain. Interaction of TFEα with both the stalk and clamp domains of RNAP during transcription initiation may retain the clamp domain in an open conformation necessary for initiation and early elongation. Replacement of TFE by Spt4/5 during early elongation may alter clamp positioning and further stabilize the elongation complex (2).

DNA ELEMENTS

Transcription initiation is regulated by DNA elements that are recognized by basal transcription factors and that influence subsequent steps in promoter opening. There are four DNA elements currently known to regulate archaeal transcription initiation: (i) the TATA box located approximately 25 bp upstream of the site of transcription initiation (64–66), (ii) the TFB recognition element (BRE) located immediately upstream of the TATA box (6), (iii) the initiator element (INR) located within the initially transcribed region, and (iv) the promoter proximal element (PPE) located between the TATA box and the site of transcription initiation (67–69). Of these four, only the TATA box and the BRE are required for transcription initiation, although alterations to all four elements can influence the total output of a promoter.

The INR is not a required DNA element for transcription initiation; however, it is a regulatory element that can increase the strength of the promoter in a TATA- and BRE-dependent manner. The INR is a core promoter element located in the 5′ untranslated region, and it has sequence similarity to the TATA box. The INR has been shown to be targeted by some transcriptional activators, and its high AT content may facilitate promoter opening in some instances. Many archaeal transcripts are leaderless, so the INR is not consistently identifiable, and the regulatory influence of INR sequences does not appear to extend to RNA half-life or alter the translational capacity (70). PPEs, centered approximately 10 bps upstream of the site of initiation, have been shown to increase transcription output through recruitment of TFB (67, 68). Additionally, permanganate footprinting data of the preinitiation complex demonstrated that the border of the transcription bubble is at the PPE and that this region is important for the activity of TFEα-TFEβ (17).

REGULATION OF ELONGATION

As transcription transitions from initiation to elongation, RNAP undergoes a conformational change accompanied by the replacement of initiation factors with elongation factors (2, 12, 71–74). It is plausible that the emerging nascent transcript stimulates the swap of regulatory factors and initiates the intramolecular movements that result in stable elongation complex formation (62, 75). Very few transcription elongation factors have been bioinformatically identified within archaeal genomes, and it is probable that archaeon-specific factors await discovery. It is worth noting what is seemingly not encoded in archaeal genomes, given that so much of archaeal and eukaryotic transcription machinery is shared. Archaeal genomes do not appear to encode any coactivator complexes or megacomplexes for chromatin modification or rearrangements. There does not appear to be machinery for regulated posttranslational modifications of the archaeal transcription apparatus nor of chromatin, with the exception of acetylation/deacetylation of the small chromatin-associated protein Alba (76–79). Furthermore, archaeal transcripts are not capped, do not require nuclear export, and, with the exception of self-splicing introns, are intronless; thus, factors responsible for these activities are similarly lacking from archaeal genomes (80–82).

Transcription elongation factors have various roles, including increasing processivity and fidelity of RNAP and/or increasing genome stability. Only two archaeal elongation factors have been experimentally studied: the aforementioned universally conserved elongation factor Spt5, often with a conserved binding partner Spt4 (Spt4/5) (2, 83, 84), and transcription factor S (TFS) (85, 86).

Several recent studies have shed light on the roles of Spt5 during elongation (1, 72, 87, 88). TFS, with homology to the C-terminal domain of eukaryotic TFIIS and functionally analogous to GreA/GreB in *Bacteria* (8, 89–91), can stimulate endonucleolytic cleavage of the RNA from backtracked RNAP complexes (85, 91–93). The finding of multiple TFS homologues in some archaeal lineages offers the possibility of unique regulatory roles of specific isoforms.

TRANSCRIPTION FACTOR Spt5

Archaeal Spt5, homologous to bacterially encoded NusG, consists of two domains: the NusG N-terminal (NGN) domain and a single C-terminal Kyrpides-Ouzounis-Woese (KOW) domain with affinity for single-stranded RNA (83, 84, 87); eukaryotic Spt5 typically contain three to six repeats of the C-terminal KOW domain (94–96). Critical, direct molecular interactions between Spt5 and RNAP have been identified in both *Bacteria* and *Archaea* (83, 84, 87, 88, 95, 97–99), and the conservation of RNAP and Spt5 infers that these same interactions are used in *Eukarya*. Briefly, a hydrophobic depression on the NGN domain interacts with the mobile clamp domain of RNAP, with additional interactions between the NGN domain and RNAP jaw domain likely fixing the location of the clamp domain in a closed configuration (11, 98). Spt5 interaction with RNAP is not necessary for productive and processive elongation *in vitro*, but the interaction does increase the total output of transcription systems (1). It is plausible that Spt5 increases elongation rates and processivity, as NusG in *Escherichia coli* does, and it is further possible that the increased efficiency of transcription results from the stabilization of the clamp domain that in turn stabilizes the DNA-RNA hybrid in place during transcription elongation (87, 100–102). The NGN domain also contacts the upstream strands of DNA, offering protection from backtracking, and, by inference, may reduce pausing of the transcription elongation complex (87, 88, 103, 104). It is of importance to note that NusG/Spt5 can have a positive and/or negative effect on elongation rates and pause events of RNAP. In *Thermus thermophilus*, NusG slows down RNA elongation rather than increases elongation rates (105). In *Bacillus subtilis*, sequence-specific interactions of the NGN and nontemplate DNA strand within the paused transcription bubble stabilize the pause event in the *trp* operon (103, 106). Furthermore, evidence has shown that Spt4/5 induces pauses during early elongation of Pol I but promotes elongation downstream (107). Although NusG can elicit opposite roles on transcription elongation, the NusG-RNAP binding sites remain well conserved across various species. Archaeal and eukaryotic genomes often encode an additional elongation factor, Spt4 (annotated as RpoE″/RpoE2 in *Archaea*), that forms a complex with Spt5 and stabilizes the Spt5-RNAP interaction (1, 84, 95). Spt4 does not appear to be essential; however, the affinity of Spt5 for RNAP decreases in the absence of Spt4 *in vitro* (1).

The primary interacting partners (e.g., RNAP and Spt4) of the Spt5-NGN domain have been established in molecular detail; however, no specific interacting partners of the KOW domain have been identified in archaea. It is possible that the affinity of the KOW domain for RNA leads to nonspecific interactions with the emerging transcript; however, it is tempting to speculate about greater involvement of the KOW domain based on the known activities of the C terminus of bacterial NusG (108). Bacterial NusG can facilitate elongation or termination depending on its binding partner (99–101, 109–111). The bacterial NusG KOW domain can interact with the S10 ribosomal subunit (NusE) during elongation, thereby linking the leading ribosome with the transcription apparatus (110, 111). When not bound to a trailing ribosome, the bacterial NusG-KOW domain can be bound by and stimulate the activity of the transcription termination factor Rho (109, 112, 113). Archaeal transcription and translation are similarly coupled (114, 115), and it is reasonable to venture that archaeal Spt5 can also link the archaeal transcription and translation apparatuses and also potentially interact with termination factors.

INTRAMOLECULAR REARRANGEMENTS OF RNAP MAY INCREASE PROCESSIVITY

The archaeal and three eukaryotic RNAPs can be reduced in complexity to three large domains: the core, the mobile clamp, and the stalk (4, 73, 116). The archaeoeukaryotic stalk, absent from bacterial RNAP, is used by a host of archaeal and eukaryotic transcription factors to bind and regulate the activities of RNAP. Increasing evidence from biochemical, biophysical, and *in vivo* approaches indicate that transcription factor binding often stimulates intramolecular movements of RNAP that appear necessary for transitions between phases of the transcription cycle (2, 4, 26, 88, 97, 117).

Hinge-like movement of the mobile clamp domain has been demonstrated for the bacterial RNAP (71). The movements of the mobile clamp are sufficiently large enough to open the main channel of RNAP, such that double-stranded DNA can easily enter and exit when the clamp is open, whereas double-stranded DNA—or the RNA-DNA hybrid—would be trapped inside RNAP when the clamp is closed. The bacterial RNAP clamp is open during initiation but remains closed during processive elongation (71), leading to a simple model of encapsulation of the nucleic acids to explain the dramatic stability of the elongation complex. It is logical to propose mechanistic actions of transcription factors that may modulate the clamp positioning with respect to the core and stalk domains of RNAP and thus alter the stability and transitions of RNAP throughout the transcription cycle. TFE is predicted to make contacts with both the clamp and stalk domain of RNAP, thereby fixing the clamp into the open conformation critical for initiation (26, 59, 117–119). As transcription transitions into the elongation phase, RNA emerges from the enzyme and interacts with the stalk domain (62, 75), where a predicted steric clash occurs between the RNA and the TFE, likely driving TFE to disengage from RNAP. The disengagement of TFE allows for Spt5 to bind to the clamp and core domains of RNAP and lock the clamp in the closed position, thus ensuring processivity during elongation (87).

RNAP clamp movement is predicted to be universal; however, both the archaeal and the eukaryotic RNAP contain additional subunits, including the stalk domain (2, 73, 116, 118, 119), and previous structural data predicted that the stalk domain would sterically limit or abolish major movements of the clamp domain. Recent crystallographic evidence of the complete euryarchaeal RNAP demonstrated that the clamp is able to open without a steric clash with the stalk domain through a coordinated swing and rotation movement of both the clamp and stalk domains (73). This evidence supports the bacterial mechanism of the clamp opening and closing during initiation/termination or elongation, respectively, thus supporting a universal model of clamp movement.

TERMINATION

Transcription termination occurs when the transcription elongation complex becomes sufficiently unstable and fails to maintain contact between RNAP and the encapsulated nucleic acids. The stability of the transcription elongation complex is derived from (i) contacts between RNAP and the RNA-DNA hybrid, (ii) contacts between RNAP and single-stranded RNA in the exit channel, (iii) contacts between RNAP and the downstream DNA, and (iv) the base pairing of the RNA-DNA hybrid (116, 120–126). The first and last of these contacts are most likely to be altered during the termination process. Transcription through specific DNA sequences can result in stronger or weaker base pairing within the RNA-DNA hybrid, and contacts between RNAP and the nucleic acids are most easily modified by movements of the clamp domain that relieve movements of the hybrid with respect to the core of RNAP (127–129). Release of the nascent RNA may be possible through continued translocation in the absence of synthesis, or the RNA-DNA hybrid could be released in bulk if the clamp domain transitions from a closed to an open position. The gene-dense nature of many archaeal genomes necessitates timely termination of transcription to prevent aberrant transcription of neighboring genes. It is predicted that there are two mechanisms of termination across all domains: intrinsic termination and factor-dependent termination (Fig. 1B).

INTRINSIC TERMINATION

Intrinsic transcription termination is driven primarily by weak base pairing within the RNA-DNA hybrid and occurs independent of the activity of transcription factors (130, 131). Intrinsic transcription termination has been established in all three domains (19, 20, 132, 133), with some differences in sequence and structural requirements (130, 132, 134–136). The archaeal RNAP, like eukaryotic RNAP III, is sensitive to intrinsic termination (19, 133, 137, 138). Eukaryotic RNAP I and RNAP II do respond to DNA sequence context in the form of pauses and arrests but rarely release the transcript at such positions (139–141). Archaeal intrinsic termination is characterized by a run of 5 to 10 thymidine residues in the nontemplate strand, encoding a poly(U) run at the 3′ end of the nascent RNA (19, 20). The weak rU:dA RNA-DNA hybrid at or near the positions of termination is seemingly insufficiently energy rich to maintain the stability of the elongation complex; RNAP III similarly spontaneously dissociates upon transcription of poly(T) nontemplate tracts.

IDENTIFICATION OF FACTOR-DEPENDENT TERMINATION

Transcription factors involved in initiation and elongation have been characterized in all domains, while a transcription termination factor(s) has been characterized only in *Bacteria* and *Eukarya* (142–145). By inference, from known termination factors that are employed in bacterial and eukaryotic systems, it is easily argued that protein factors are encoded in archaeal genomes that have the capacity to direct transcription termination *in vivo*. Bioinformatic analyses reveal some potential targets that remain to be more fully evaluated, but there are no easily identified homologues of known eukaryotic or bacterial termination factors. Two well-studied transcription bacterial termination factors, Rho and Mfd (13, 146–150), lack clear homologues in archaeal genomes, but there are hints that analogous activities may be present in archaeal species. Rho is a homohexamer helicase that represses phage transcription and mediates polar repression of downstream genes when transcription and translation become uncoupled (142, 151–153). *Archaea* demonstrate polar repression of downstream genes in the absence of continued translation, and it is likely that a factor or factors mediate polarity in archaea (115). It is tempting to use the bacterial model of NusG-Rho interactions to conjure a similar picture for Spt5-KOW interactions with an archaeal transcription termination factor; Rho is capable of terminating a stalled archaeal RNAP *in vitro* (19). The bacterial Mfd protein can remove RNAP from sites of DNA damage and initiate transcription-coupled DNA repair (146, 148, 150, 154). Recent evidence that the archaeal RNAP halts synthesis and forms long-lived complexes at the site of lesions in DNA *in vitro* predicts that mechanisms exist to remove RNAP from the site of damage (T. J. Santangelo, unpublished results).

CHROMATIN ARCHITECTURE AFFECTS THE TRANSCRIPTION CYCLE

Archaea employ two seemingly distinct mechanisms to compact, wrap, and condense their genomes to fit within the cell (Fig. 2) (155). Most euryarchaeal species are polyploid (156–160) and encode histone proteins that dominate chromatin architecture (156–160); archaeal histones mimic the core eukaryotic histone fold (161). In contrast, most crenarchaeal species are diploid and are reliant on small, basic nucleoid proteins to organize their genomes (162, 163). Condensation demands organization of the genome and offers regulatory opportunities by controlling the accessibility of promoter sequences, the introduction of local superhelicities that may promote or inhibit promoter opening, and the potential for the introduction of chromatin-based obstacles to transcription elongation. The overall role of genome architecture with respect to archaeal transcription is an emerging area, with several recent studies highlighting the breadth of influences that genome architecture can have on transcription output at the organismal level.

Archaeal histone-based chromatin is composed of nucleosome particles that wrap and condense the genome. The best-described complexes are homo- or hetero-histone tetramers, homologous to the H3/H4 tetramer in eukaryotes, that associate with ~60 bp of double-stranded DNA. Archaeal histones share similar biases with eukaryotic nucleosomes for flexible DNA sequences and are, in general, absent from the core promoters of archaeal genes (164, 165). Archaeal histone proteins share the same core fold as eukaryotic histones but lack the extensions from this fold (i.e., tails) that are highly modified and essential for proper nucleosome dynamics in eukaryotes (166). Higher-order structure has been demonstrated in *Thermococcus kodakarensis* in the form of dynamic histone polymers that have the ability to wrap up to 180 bp (167). Archaeal nucleosomes present a surmountable barrier to the progression of the transcription elongation complex, although traversion does slow the elongation complex (168). The lack of known modifications to archaeal histones, and the lack of known machinery for the repositioning or movement of archaeal nucleosomes, suggests that transcription elongation complexes simply traverse the nucleosomes and that chromatin organization spontaneously reforms when the histones gain access to preferred binding positions following the departure of RNAP. This mechanism of elongation through the histones is similar to the mechanism of Pol III in eukaryotes (168–170).

The activities or stimulatory effects of archaeal elongation factors on transcription through archaeal histone-based chromatin

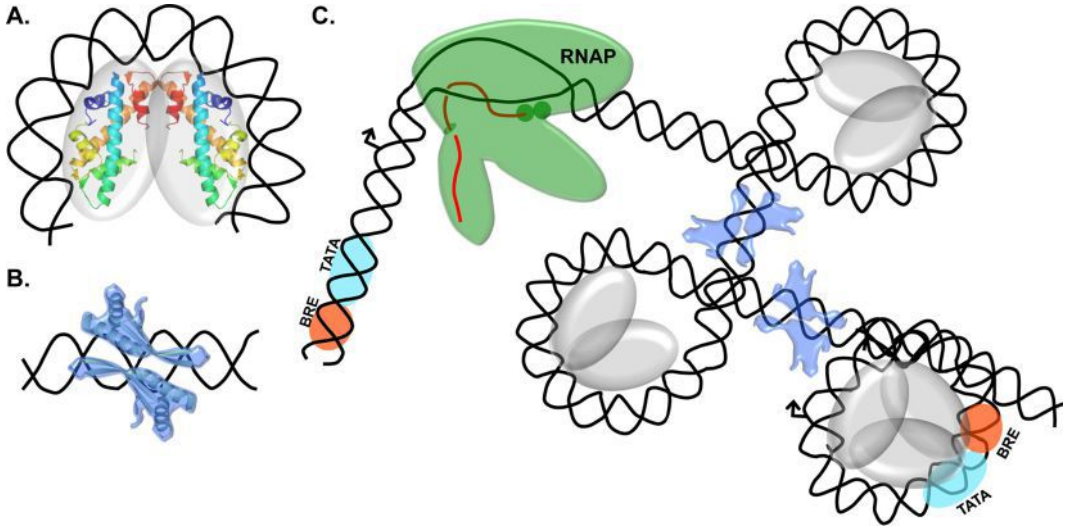


FIG 2 Transcription in the context of archaeal chromatin. (A) The structure of histone A from *Methothermus fervidus* (PDB ID no. 1B67) is overlaid by a cartoon representation of each histone dimer with ~60 bp of DNA wrapping the complex. (B) The crystal structure of an Alba dimer from *Sulfolobus solfataricus* (PDB ID no. 1H0X) bound to DNA is overlaid by a cartoon representation. (C) Transcription elongation continues in a chromatin environment. Accessibility of the TATA box and BRE is altered by localized chromatin structure.

remain to be explored; the substantial pausing and delayed progress of RNAP on chromatinized templates suggest that elongation factors will accelerate progress of the transcription elongation complex. Any role of chromatin architecture in transcription termination is similarly unexplored. The topology of naked DNA templates does influence the positions and efficiencies of intrinsic terminators, suggesting that chromatin templates may also influence termination patterns. Nucleosomes are depleted not only from promoter regions but also from predicted termination regions, suggesting a potential regulatory role for chromatin architecture on termination of transcription (164).

HISTONE-BASED REGULATION OF TRANSCRIPTION

Several genetic studies have addressed the role of archaeal histone-based chromatin on gene expression at the organismal level, with surprisingly different results. In some halophilic species, singular histone-encoding genes are nonessential, and histone proteins appear to function more akin to site-specific transcription factors, moderately influencing the expression of only a few genes (171). These studies contrast the view of histone proteins as general organizational factors with the global influence on gene expression and minimally suggest that the archaeal chromatin of some species is dependent on the activities of many nucleoid-associated proteins. When histone-encoding genes have been deleted, or have been attempted to be deleted from other species, more global disruption of gene expression has been noted (161, 164, 165, 167, 171–176). Some species are reliant on at least one histone protein, and it is unclear at this point whether the noted global changes in gene expression seen in deletion strains stem from reorganization or disorganization of the archaeal genomes or the primary, secondary, and tertiary effects of localized disruptions that lead to additional differences in regulation at remote sites.

NUCLEOSOME OCCUPANCY AT THE PROMOTER

Chromatin architecture at a promoter could influence or prevent transcription initiation by occluding transcription factor binding or inhibiting DNA melting (164, 167, 168, 177). Crenarchaeon-encoded nucleoid-associated proteins have been shown to influence transcription output through the acetylation/deaceytlation of Alba *in vitro* (76), although Alba has not yet been shown to influence transcription *in vivo*. It is possible that Alba regulates transcription, given that Alba proteins can loop, condense, bridge, and even saturate DNA *in vitro*, but the *in vivo* dynamics remain unknown (178–182). In the euryarchaeal organism *Methanococcus voltae* the deletion of the gene encoding Alba resulted in the upregulation of only a small number of genes, implying that Alba-based regulation may be limited in scope (173). Additional research may reveal a clearer picture of transcriptional regulation through the binding of Alba.

The binding preferences and genomic locations of stable euryarchaeal histone proteins interactions have been mapped, and it has been shown that regions directly upstream from the start codon are nucleosome depleted on a global scale (164, 165). The presence of histones bound at the promoter has been correlated with a decrease in total transcription *in vitro* (177), and it was suggested that both steric and torsional effects limited binding of basal transcription factors to the DNA (177). Although most data support the lack of nucleosomes at the promoter, specific promoters can be regulated by nucleosome occupancy. This appears to be a general mechanism of histone-based regulation in some halophiles and a more specialized mechanism of regulation in other species. The transcriptional activator Ptr2 from *Methanocaldococcus jannaschii* must outcompete histones for binding to the promoter to activate transcription of select genes (183).

CONCLUSIONS AND OUTSTANDING ISSUES

Exploration of archaeal transcription and regulation continues to yield a bounty of evolutionary, biophysical, and mechanistic details of transcription mechanisms that are often applicable to all extant life. The ability to reconstitute the complete archaeal transcription apparatus permits biophysical studies not possible with eukaryotic components, and the simplicity and explicit homology

of many factors provide meaningful insights into the mechanistic roles of individual factors and even of specific domains and residues of archaeal transcription components. The development and recent advances in genetic techniques for more archaeal species are now offering complementary *in vivo* studies to probe regulatory strategies and rationally manipulate protein interfaces and activities in the cell. Although discussion of transcriptome mapping of archaeal organisms is outside the scope of this review, the mapping is becoming more frequent (36, 184–189) and offers invaluable insight into noncoding RNA, transcription start site selection and redundancy, and expression levels under various growth conditions (36, 171, 190–193).

There is still much to be learned regarding archaeal transcription regulation and mechanisms. The identification and characterization of additional archaeal elongation and potential termination factors offer the opportunity to examine archaeon-specific mechanisms of regulation. Factors that regulate the organization and dynamics of archaeal chromatin are likely to be identified and should offer contrasting regulatory potential with the network of regulatory strategies employed for eukaryotic chromatin. Continued insightful biophysical probing of shared archaeoeukaryotic factors will surely reveal conserved regulatory strategies for promoter recognition, DNA melting, transcription factor swapping, and elongation through chromatinized templates. Advances in genetic techniques will quickly move studies of archaeal transcription inside the cell, and the application of omics approaches to gene expression in modified strains should answer outstanding question surrounding archaeal responses to external signals and ever-changing environments. Given the extremophilic nature of many experimentally utilized *Archaea*, the evolutionary survival strategies of these remarkable microbes will come into better focus.

ACKNOWLEDGMENTS

We thank Michael Bartlett and Finn Werner as well as the Santangelo laboratory for discussions and edits to improve the manuscript.

This work was supported by NIH (grant GM100329) and Department of Energy (grant 004010-00002) funding to T.J.S.

FUNDING INFORMATION

This work, including the efforts of Thomas J. Santangelo, was funded by Department of Energy (004010-00002). This work, including the efforts of Thomas J. Santangelo, was funded by HHS | National Institutes of Health (NIH) (GM100329).

REFERENCES

1. **Hirtreiter A, Damsma GE, Cheung ACM, Klose D, Grohmann D, Vojnic E, Martin ACR, Cramer P, Werner F.** 2010. Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. Nucleic Acids Res **38:**4040–4051. http://dx.doi.org/10.1093/nar/gkq135.
2. **Grohmann D, Nagy J, Chakraborty A, Klose D, Fielden D, Ebright RH, Michaelis J, Werner F.** 2011. The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. Mol Cell **43:**263–274. http://dx.doi.org/10.1016/j.molcel.2011.05.030.
3. **Armache K-J, Mitterweger S, Meinhart A, Cramer P.** 2005. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J Biol Chem **280:**7131–7134. http://dx.doi.org/10.1074/jbc.M413038200.
4. **Hirata A, Klein BJ, Murakami KS.** 2008. The X-ray crystal structure of RNA polymerase from *Archaea*. Nature **451:**851–854. http://dx.doi.org/10.1038/nature06530.
5. **Rowlands T, Baumann P, Jackson SP.** 1994. The TATA-binding protein: a general transcription factor in eukaryotes and archaebacteria. Science **264:**1326–1329. http://dx.doi.org/10.1126/science.8191287.
6. **Qureshi SA, Bell SD, Jackson SP.** 1997. Factor requirements for transcription in the archaeon *Sulfolobus shibatae*. EMBO J **16:**2927–2936. http://dx.doi.org/10.1093/emboj/16.10.2927.
7. **Hausner W, Thomm M.** 2001. Events during initiation of archaeal transcription: open complex formation and DNA-protein interactions. J Bacteriol **183:**3025–3031. http://dx.doi.org/10.1128/JB.183.10.3025-3031.2001.
8. **Langer D, Hain J, Thuriaux P, Zillig W.** 1995. Transcription in archaea: similarity to that in eucarya. Proc Natl Acad Sci U S A **92:**5768–5772. http://dx.doi.org/10.1073/pnas.92.13.5768.
9. **Decker KB, Hinton DM.** 2013. Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. Annu Rev Microbiol **67:**113–139. http://dx.doi.org/10.1146/annurev-micro-092412-155756.
10. **De Carlo S, Lin S-C, Taatjes DJ, Hoenger A.** Molecular basis of transcription initiation in *Archaea*. Transcription **1:**103–111.
11. **Belogurov GA, Mooney RA, Svetlov V, Landick R, Artsimovitch I.** 2009. Functional specialization of transcription elongation factors. EMBO J **28:**112–122. http://dx.doi.org/10.1038/emboj.2008.268.
12. **Fouqueau T, Zeller ME, Cheung AC, Cramer P, Thomm M.** 2013. The RNA polymerase trigger loop functions in all three phases of the transcription cycle. Nucleic Acids Res **41:**7048–7059. http://dx.doi.org/10.1093/nar/gkt433.
13. **D'Heygère F, Schwartz A, Coste F, Castaing B, Boudvillain M.** 2015. ATP-dependent motor activity of the transcription termination factor Rho from *Mycobacterium tuberculosis*. Nucleic Acids Res **43:**6099–6111. http://dx.doi.org/10.1093/nar/gkv505.
14. **Bartlett MS, Thomm M, Geiduschek EP.** 2000. The orientation of DNA in an archaeal transcription initiation complex. Nat Struct Biol **7:**782–785. http://dx.doi.org/10.1038/79020.
15. **Grünberg S, Bartlett MS, Naji S, Thomm M.** 2007. Transcription factor E is a part of transcription elongation complexes. J Biol Chem **282:**35482–35490. http://dx.doi.org/10.1074/jbc.M707371200.
16. **Bell SD, Brinkman AB, van der Oost J, Jackson SP.** 2001. The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. EMBO Rep **2:**133–138. http://dx.doi.org/10.1093/embo-reports/kve021.
17. **Blombach F, Salvadori E, Fouqueau T, Yan J, Reimann J, Sheppard C, Smollett KL, Albers SV, Kay CW, Thalassinos K, Werner F.** 2015. Archaeal TFEα/β is a hybrid of TFIIE and the RNA polymerase III subcomplex hRPC62/39. eLife **4:**e08378.
18. **Hanzelka BL, Darcy TJ, Reeve JN.** 2001. TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eucaryal transcription factor TFIIEalpha. J Bacteriol **183:**1813–1818. http://dx.doi.org/10.1128/JB.183.5.1813-1818.2001.
19. **Santangelo TJ, Reeve JN.** 2006. Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. J Mol Biol **355:**196–210. http://dx.doi.org/10.1016/j.jmb.2005.10.062.
20. **Santangelo TJ, Cubonová L, Skinner KM, Reeve JN.** 2009. Archaeal intrinsic transcription termination *in vivo*. J Bacteriol **191:**7102–7108. http://dx.doi.org/10.1128/JB.00982-09.
21. **Santangelo TJ, Cubonova L, James CL, Reeve JN.** 2007. TFB1 or TFB2 is sufficient for *Thermococcus kodakaraensis* viability and for basal transcription *in vitro*. J Mol Biol **367:**344–357. http://dx.doi.org/10.1016/j.jmb.2006.12.069.
22. **Gietl A, Holzmeister P, Blombach F, Schulz S, von Voithenberg LV, Lamb DC, Werner F, Tinnefeld P, Grohmann D.** 2014. Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. Nucleic Acids Res **42:**6219–6231. http://dx.doi.org/10.1093/nar/gku273.
23. **Ochs SM, Thumann S, Richau R, Weirauch MT, Lowe TM, Thomm M, Hausner W.** 2012. Activation of archaeal transcription mediated by recruitment of transcription factor B. J Biol Chem **287:**18863–18871. http://dx.doi.org/10.1074/jbc.M112.365742.
24. **Renfrow MB, Naryshkin N, Lewis LM, Chen HT, Ebright RH, Scott RA.** 2004. Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. J Biol Chem **279:**2825–2831. http://dx.doi.org/10.1074/jbc.M311433200.
25. **Goede B, Naji S, von Kampen O, Ilg K, Thomm M.** 2006. Protein-protein interactions in the archaeal transcriptional machinery: binding studies of isolated RNA polymerase subunits and transcription factors. J Biol Chem **281:**30581–30592. http://dx.doi.org/10.1074/jbc.M605209200.

26. **Walker JE, Santangelo TJ.** 2015. Analyses of in vivo interactions between transcription factors and the archaeal RNA polymerase. Methods **86:**73–79. http://dx.doi.org/10.1016/j.ymeth.2015.05.023.
27. **Bell SD, Kosa PL, Sigler PB, Jackson SP.** 1999. Orientation of the transcription preinitiation complex in archaea. Proc Natl Acad Sci U S A **96:**13662–13667. http://dx.doi.org/10.1073/pnas.96.24.13662.
28. **Ouhammouch M, Dewhurst RE, Hausner W, Thomm M, Geiduschek EP.** 2003. Activation of archaeal transcription by recruitment of the TATA-binding protein. Proc Natl Acad Sci U S A **100:**5097–5102. http://dx.doi.org/10.1073/pnas.0837150100.
29. **Karr EA.** 2014. Transcription regulation in the third domain. Adv Appl Microbiol **89:**101–133. http://dx.doi.org/10.1016/B978-0-12-800259-9.00003-2.
30. **Ng WV, Kennedy SP, Mahairas GG, Berquist B, Pan M, Shukla HD, Lasky SR, Baliga NS, Thorsson V, Sbrogna J, Swartzell S, Weir D, Hall J, Dahl TA, Welti R, Goo YA, Leithauser B, Keller K, Cruz R, Danson MJ, Hough DW, Maddocks DG, Jablonski PE, Krebs MP, Angevine CM, Dale H, Isenbarger TA, Peck RF, Pohlschroder M, Spudich JL, Jung KW, Alam M, Freitas T, Hou S, Daniels CJ, Dennis PP, Omer AD, Ebhardt H, Lowe TM, Liang P, Riley M, Hood L, DasSarma S.** 2000. Genome sequence of *Halobacterium* species NRC-1. Proc Natl Acad Sci U S A **97:**12176–12181. http://dx.doi.org/10.1073/pnas.190337797.
31. **Baliga NS, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, Shannon P, Chiu Y, Weng RS, Gan RR, Hung P, Date SV, Marcotte E, Hood L, Ng WV.** 2004. Genome sequence of *Haloarcula marismortui*: a halophilic archaeon from the Dead Sea. Genome Res **14:**2221–2234. http://dx.doi.org/10.1101/gr.2700304.
32. **Goo YA, Yi EC, Baliga NS, Tao WA, Pan M, Aebersold R, Goodlett DR, Hood L, Ng WV.** 2003. Proteomic analysis of an extreme halophilic archaeon, *Halobacterium* sp. NRC-1. Mol Cell Proteomics **2:**506–524.
33. **Micorescu M, Grunberg S, Franke A, Cramer P, Thomm M, Bartlett M.** 2008. Archaeal transcription: function of an alternative transcription factor B from *Pyrococcus furiosus*. J Bacteriol **190:**157–167. http://dx.doi.org/10.1128/JB.01498-07.
34. **Reichlen MJ, Murakami KS, Ferry JG.** 2010. Functional Analysis of the three TATA binding protein homologs in *Methanosarcina acetivorans*. J Bacteriol **192:**1511–1517. http://dx.doi.org/10.1128/JB.01165-09.
35. **Bartlett MS, Thomm M, Geiduschek EP.** 2004. Topography of the euryarchaeal transcription initiation complex. J Biol Chem **279:**5894–5903. http://dx.doi.org/10.1074/jbc.M311429200.
36. **Coker JA, DasSarma S.** 2007. Genetic and transcriptomic analysis of transcription factor genes in the model halophilic archaeon: coordinate action of TbpD and TfbA. BMC Genet **8:**61. http://dx.doi.org/10.1186/1471-2156-8-61.
37. **Hidese R, Nishikawa R, Gao L, Katano M, Imai T, Kato S, Kanai T, Atomi H, Imanaka T, Fujiwara S.** 2014. Different roles of two transcription factor B proteins in the hyperthermophilic archaeon *Thermococcus kodakarensis*. Extremophiles **18:**573–588. http://dx.doi.org/10.1007/s00792-014-0638-9.
38. **Bonneau R, Facciotti MT, Reiss DJ, Schmid AK, Pan M, Kaur A, Thorsson V, Shannon P, Johnson MH, Bare JC, Longabaugh W, Vuthoori M, Whitehead K, Madar A, Suzuki L, Mori T, Chang D-E, Diruggiero J, Johnson CH, Hood L, Baliga NS.** 2007. A predictive model for transcriptional control of physiology in a free living cell. Cell **131:**1354–1365. http://dx.doi.org/10.1016/j.cell.2007.10.053.
39. **Baliga NS, Goo YA, Ng WV, Hood L, Daniels CJ, DasSarma S.** 2000. Is gene expression in *Halobacterium* NRC-1 regulated by multiple TBP and TFB transcription factors? Mol Microbiol **36:**1184–1185. http://dx.doi.org/10.1046/j.1365-2958.2000.01916.x.
40. **Turkarslan S, Reiss DJ, Gibbins G, Su WL, Pan M, Bare JC, Plaisier CL, Baliga NS.** 2011. Niche adaptation by expansion and reprogramming of general transcription factors. Mol Syst Biol **7:**554.
41. **Paytubi S, White MF.** 2009. The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. Mol Microbiol **72:**1487–1499. http://dx.doi.org/10.1111/j.1365-2958.2009.06737.x.
42. **Österberg S, del Peso-Santos T, Shingler V.** 2011. Regulation of alternative sigma factor use. Annu Rev Microbiol **65:**37–55. http://dx.doi.org/10.1146/annurev.micro.112408.134219.
43. **Gehring AM, Santangelo TJ.** 2015. Manipulating archaeal systems to permit analyses of transcription elongation-termination decisions in vitro. Methods Mol Biol **1276:**263–279. http://dx.doi.org/10.1007/978-1-4939-2392-2_15.
44. **Hirata A, Kanai T, Santangelo TJ, Tajiri M, Manabe K, Reeve JN, Imanaka T, Murakami KS.** 2008. Archaeal RNA polymerase subunits E and F are not required for transcription *in vitro*, but a *Thermococcus kodakarensis* mutant lacking subunit F is temperature-sensitive. Mol Microbiol **70:**623–633. http://dx.doi.org/10.1111/j.1365-2958.2008.06430.x.
45. **DeDecker BS, O'Brien R, Fleming PJ, Geiger JH, Jackson SP, Sigler PB.** 1996. The crystal structure of a hyperthermophilic archaeal TATA-box binding protein. J Mol Biol **264:**1072–1084. http://dx.doi.org/10.1006/jmbi.1996.0697.
46. **Littlefield O, Korkhin Y, Sigler PB.** 1999. The structural basis for the oriented assembly of a TBP/TFB/promoter complex. Proc Natl Acad Sci U S A **96:**13668–13673. http://dx.doi.org/10.1073/pnas.96.24.13668.
47. **Kassavetis GA, Joazeiro CA, Pisano M, Geiduschek EP, Colbert T, Hahn S, Blanco JA.** 1992. The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. Cell **71:**1055–1064. http://dx.doi.org/10.1016/0092-8674(92)90399-W.
48. **Knutson BA, Hahn S.** 2011. Yeast Rrn7 and human TAF1B are TFIIB-related RNA polymerase I general transcription factors. Science **333:**1637–1640. http://dx.doi.org/10.1126/science.1207699.
49. **Adachi N, Senda M, Natsume R, Senda T, Horikoshi M.** 2008. Crystal structure of *Methanococcus jannaschii* TATA box-binding protein. Genes Cells **13:**1127–1140.
50. **Qureshi SA, Jackson SP.** 1998. Sequence-specific DNA binding by the *S. shibatae* TFIIB homolog, TFB, and its effect on promoter strength. Mol Cell **1:**389–400. http://dx.doi.org/10.1016/S1097-2765(00)80039-8.
51. **Bleiholder A, Frommherz R, Teufel K, Pfeifer F.** 2012. Expression of multiple *tfb* genes in different *Halobacterium salinarum* strains and interaction of TFB with transcriptional activator GvpE. Arch Microbiol **194:**269–279. http://dx.doi.org/10.1007/s00203-011-0756-z.
52. **Nagy J, Grohmann D, Cheung ACM, Schulz S, Smollett K, Werner F, Michaelis J.** 2015. Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. Nat Commun **6:**6161. http://dx.doi.org/10.1038/ncomms7161.
53. **Werner F, Weinzierl ROJ.** 2005. Direct modulation of RNA polymerase core functions by basal transcription factors. Mol Cell Biol **25:**8344–8355. http://dx.doi.org/10.1128/MCB.25.18.8344-8355.2005.
54. **Ohkuma Y, Sumimoto H, Horikoshi M, Roeder RG.** 1990. Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE. Proc Natl Acad Sci U S A **87:**9163–9167. http://dx.doi.org/10.1073/pnas.87.23.9163.
55. **Gregory Peterson M, Inostroza J, Maxon ME, Flores O, Admon A, Reinberg D, Tjian R.** 1991. Structure and functional properties of human general transcription factor IIE. Nature **354:**369–373. http://dx.doi.org/10.1038/354369a0.
56. **Vannini A, Cramer P.** 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol Cell **45:**439–446. http://dx.doi.org/10.1016/j.molcel.2012.01.023.
57. **Lefèvre S, Dumay-Odelot H, El-Ayoubi L, Budd A, Legrand P, Pinaud N, Teichmann M, Fribourg S.** 2011. Structure-function analysis of hRPC62 provides insights into RNA polymerase III transcription initiation. Nat Struct Mol Biol **18:**352–358. http://dx.doi.org/10.1038/nsmb.1996.
58. **Carter R, Drouin G.** 2010. The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. Mol Biol Evol **27:**1035–1043. http://dx.doi.org/10.1093/molbev/msp316.
59. **Naji S, Grunberg S, Thomm M.** 2007. The RPB7 orthologue E′ is required for transcriptional activity of a reconstituted archaeal core enzyme at low temperatures and stimulates open complex formation. J Biol Chem **282:**11047–11057. http://dx.doi.org/10.1074/jbc.M611674200.
60. **Meinhart A, Blobel J, Cramer P.** 2003. An extended winged helix domain in general transcription factor E/IIE alpha. J Biol Chem **278:**48267–48274. http://dx.doi.org/10.1074/jbc.M307874200.
61. **Okuda M, Tanaka A, Arai Y, Satoh M, Okamura H, Nagadoi A, Hanaoka F, Ohkuma Y, Nishimura Y.** 2004. A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. J Biol Chem **279:**51395–51403. http://dx.doi.org/10.1074/jbc.M404722200.

62. **Hirtreiter A, Grohmann D, Werner F.** 2010. Molecular mechanisms of RNA polymerase: the F/E (RPB4/7) complex is required for high processivity *in vitro*. Nucleic Acids Res **38:**585–596. http://dx.doi.org/10.1093/nar/gkp928.
63. **Ouhammouch M, Werner F, Weinzierl ROJ, Geiduschek EP.** 2004. A fully recombinant system for activator-dependent archaeal transcription. J Biol Chem **279:**51719–51721. http://dx.doi.org/10.1074/jbc.C400446200.
64. **Thomm M, Wich G.** 1988. An archaebacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. Nucleic Acids Res **16:**151–163. http://dx.doi.org/10.1093/nar/16.1.151.
65. **Reiter WD, Hüdepohl U, Zillig W.** 1990. Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. Proc Natl Acad Sci U S A **87:**9509–9513. http://dx.doi.org/10.1073/pnas.87.24.9509.
66. **Reiter WD, Palm P, Zillig W.** 1988. Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. Nucleic Acids Res **16:**1–19. http://dx.doi.org/10.1093/nar/16.1.1.
67. **Hain J, Reiter WD, Hüdepohl U, Zillig W.** 1992. Elements of an archaeal promoter defined by mutational analysis. Nucleic Acids Res **20:**5423–5428. http://dx.doi.org/10.1093/nar/20.20.5423.
68. **Peng N, Xia Q, Chen Z, Liang YX, She Q.** 2009. An upstream activation element exerting differential transcriptional activation on an archaeal promoter. Mol Microbiol **74:**928–939. http://dx.doi.org/10.1111/j.1365-2958.2009.06908.x.
69. **Brenneis M, Hering O, Lange C, Soppa J.** 2007. Experimental characterization of Cis-acting elements important for translation and transcription in halophilic archaea. PLoS Genet **3:**e229. http://dx.doi.org/10.1371/journal.pgen.0030229.
70. **Ao X, Li Y, Wang F, Feng M, Lin Y, Zhao S, Liang Y, Peng N.** 2013. The *Sulfolobus* initiator element is an important contributor to promoter strength. J Bacteriol **195:**5216–5222. http://dx.doi.org/10.1128/JB.00768-13.
71. **Chakraborty A, Wang D, Ebright YW, Korlann Y, Kortkhonjia E, Kim T, Chowdhury S, Wigneshweraraj S, Irschik H, Jansen R, Nixon BT, Knight J, Weiss S, Ebright RH.** 2012. Opening and closing of the bacterial RNA polymerase clamp. Science **337:**591–595. http://dx.doi.org/10.1126/science.1218716.
72. **Blombach F, Daviter T, Fielden D, Grohmann D, Smollett K, Werner F.** 2013. Archaeology of RNA polymerase: factor swapping during the transcription cycle. Biochem Soc Trans **41:**362–367. http://dx.doi.org/10.1042/BST20120274.
73. **Jun SH, Hirata A, Kanai T, Santangelo TJ, Imanaka T, Murakami KS.** 2014. The X-ray crystal structure of the euryarchaeal RNA polymerase in an open-clamp configuration. Nat Commun **5:**5132. http://dx.doi.org/10.1038/ncomms6132.
74. **Hartzog G a. Kaplan CD.** 2011. Competing for the clamp: promoting RNA polymerase processivity and managing the transition from initiation to elongation. Mol Cell **43:**161–163. http://dx.doi.org/10.1016/j.molcel.2011.07.002.
75. **Grohmann D, Klose D, Klare JP, Kay CW, Steinhoff HJ, Werner F.** 2010. RNA-binding to archaeal RNA polymerase subunits F/E: a DEER and FRET study. J Am Chem Soc **132:**5954–5955. http://dx.doi.org/10.1021/ja101663d.
76. **Bell SD, Botting CH, Wardleworth BN, Jackson SP, White MF.** 2002. The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. Science **296:**148–151. http://dx.doi.org/10.1126/science.1070506.
77. **Marsh VL, Peak-Chew SY, Bell SD.** 2005. Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. J Biol Chem **280:**21122–21128. http://dx.doi.org/10.1074/jbc.M501280200.
78. **Wardleworth BN, Russell RJM, Bell SD, Taylor GL, White MF.** 2002. Structure of Alba: an archaeal chromatin protein modulated by acetylation. EMBO J **21:**4654–4662. http://dx.doi.org/10.1093/emboj/cdf465.
79. **Goyal M, Banerjee C, Nag S, Bandyopadhyay U.** 2016. The Alba protein family: structure and function. Biochim Biophys Acta **1864:**570–583. http://dx.doi.org/10.1016/j.bbapap.2016.02.015.
80. **Tang TH, Rozhdestvensky TS, d'Orval BC, Bortolin M-L, Huber H, Charpentier B, Branlant C, Bachellerie J-P, Brosius J, Hüttenhofer A.** 2002. RNomics in *Archaea* reveals a further link between splicing of archaeal introns and rRNA processing. Nucleic Acids Res **30:**921–930. http://dx.doi.org/10.1093/nar/30.4.921.
81. **Tocchini-Valentini GD, Fruscoloni P, Tocchini-Valentini GP.** 2005. Coevolution of tRNA intron motifs and tRNA endonuclease architecture in *Archaea*. Proc Natl Acad Sci U S A **102:**15418–15422. http://dx.doi.org/10.1073/pnas.0506750102.
82. **Kjems J, Jensen J, Olesen T, Garrett RA.** 1989. Comparison of transfer RNA and ribosomal RNA intron splicing in the extreme thermophile and archaebacterium *Desulfurococcus mobilis*. Can J Microbiol **35:**210–214. http://dx.doi.org/10.1139/m89-033.
83. **Martinez-Rucobo FW, Sainsbury S, Cheung ACM, Cramer P.** 2011. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. EMBO J **30:**1302–1310. http://dx.doi.org/10.1038/emboj.2011.64.
84. **Zhou H, Liu Q, Gao Y, Teng M, Niu L.** 2009. Crystal structure of NusG N-terminal (NGN) domain from *Methanocaldococcus jannaschii* and its interaction with rpoE. Proteins **76:**787–793. http://dx.doi.org/10.1002/prot.22465.
85. **Hausner W, Lange U, Musfeldt M.** 2000. Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. J Biol Chem **275:**12393–12399. http://dx.doi.org/10.1074/jbc.275.17.12393.
86. **Langer D, Zillig W.** 1993. Putative tfIIs gene of *Sulfolobus acidocaldarius* encoding an archaeal transcription elongation factor is situated directly downstream of the gene for a small subunit of DNA-dependent RNA polymerase. Nucleic Acids Res **21:**2251. http://dx.doi.org/10.1093/nar/21.9.2251.
87. **Klein BJ, Bose D, Baker KJ, Yusoff ZM, Zhang X, Murakami KS.** 2011. RNA polymerase and transcription elongation factor Spt4/5 complex structure. Proc Natl Acad Sci U S A **108:**546–550. http://dx.doi.org/10.1073/pnas.1013828108.
88. **Guo G, Gao Y, Zhu Z, Zhao D, Liu Z, Zhou H, Niu L, Teng M.** 2015. Structural and biochemical insights into the DNA-binding mode of MjSpt4p:Spt5 complex at the exit tunnel of RNAPII. J Struct Biol **192:**418–425. http://dx.doi.org/10.1016/j.jsb.2015.09.023.
89. **Marr MT, Roberts JW.** 2000. Function of transcription cleavage factors GreA and GreB at a regulatory pause site. Mol Cell **6:**1275–1285. http://dx.doi.org/10.1016/S1097-2765(00)00126-X.
90. **Erie DA, Hajiseyedjavadi O, Young MC, von Hippel PH.** 1993. Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. Science **262:**867–873. http://dx.doi.org/10.1126/science.8235608.
91. **Grünberg S, Reich C, Zeller ME, Bartlett MS, Thomm M.** 2010. Rearrangement of the RNA polymerase subunit H and the lower jaw in archaeal elongation complexes. Nucleic Acids Res **38:**1950–1963. http://dx.doi.org/10.1093/nar/gkp1190.
92. **Lange U, Hausner W.** 2004. Transcriptional fidelity and proofreading in archaea and implications for the mechanism of TFS-induced RNA cleavage. Mol Microbiol **52:**1133–1143. http://dx.doi.org/10.1111/j.1365-2958.2004.04039.x.
93. **Schweikhard V, Meng C, Murakami K, Kaplan CD, Kornberg RD, Block SM.** 2014. Transcription factors TFIIF and TFIIS promote transcript elongation by RNA polymerase II by synergistic and independent mechanisms. Proc Natl Acad Sci U S A **111:**6642–6647. http://dx.doi.org/10.1073/pnas.1405181111.
94. **Meyer PA, Li S, Zhang M, Yamada K, Takagi Y, Hartzog GA, Fu J.** 2015. Structures and functions of the multiple KOW domains of transcription elongation factor Spt5. Mol Cell Biol **35:**3354–3369. http://dx.doi.org/10.1128/MCB.00520-15.
95. **Guo M, Xu F, Yamada J, Egelhofer T, Gao Y, Hartzog GA, Teng M, Niu L.** 2008. Core structure of the yeast Spt4-Spt5 complex: a conserved module for regulation of transcription elongation. Structure **16:**1649–1658. http://dx.doi.org/10.1016/j.str.2008.08.013.
96. **Yakhnin AV, Babitzke P.** 2014. NusG/Spt5: are there common functions of this ubiquitous transcription elongation factor? Curr Opin Microbiol **18:**68–71. http://dx.doi.org/10.1016/j.mib.2014.02.005.
97. **Drögemüller J, Strauß M, Schweimer K, Jurk M, Rösch P, Knauer SH.** 2015. Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy. Sci Rep **5:**16428. http://dx.doi.org/10.1038/srep16428.
98. **Sevostyanova A, Belogurov Ga, Mooney Ra, Landick R, Artsimovitch I.** 2011. The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Mol Cell **43:**253–262. http://dx.doi.org/10.1016/j.molcel.2011.05.026.
99. **Mooney RA, Schweimer K, Rösch P, Gottesman M, Landick R.** 2009. Two structurally independent domains of *E. coli* NusG create regulatory

plasticity via distinct interactions with RNA polymerase and regulators. J Mol Biol **391:**341–358. http://dx.doi.org/10.1016/j.jmb.2009.05.078.
100. **Burova E, Hung SC, Sagitov V, Stitt BL, Gottesman ME.** 1995. *Escherichia coli* NusG protein stimulates transcription elongation rates *in vivo* and *in vitro*. J Bacteriol **177:**1388–1392.
101. **Herbert KM, Zhou J, Mooney Ra, Porta A, La, Landick R, Block SM.** 2010. *E. coli* NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. J Mol Biol **399:**17–30. http://dx.doi.org/10.1016/j.jmb.2010.03.051.
102. **Belogurov GA, Artsimovitch I.** 2015. Regulation of transcript elongation. Annu Rev Microbiol **69:**49–69. http://dx.doi.org/10.1146/annurev-micro-091014-104047.
103. **Yakhnin AV, Murakami KS, Babitzke P.** 2016. NusG is a sequence-specific RNA polymerase pause factor that binds to the nontemplate DNA within the paused transcription bubble. J Biol Chem **291:**5299–5308. http://dx.doi.org/10.1074/jbc.M115.704189.
104. **Crickard JB, Fu J, Reese JC.** 2016. Biochemical Analysis of yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II arrest. J Biol Chem **291:**9853–9870.
105. **Sevostyanova A, Artsimovitch I.** 2010. Functional analysis of *Thermus thermophilus* transcription factor NusG. Nucleic Acids Res **38:**7432–7445. http://dx.doi.org/10.1093/nar/gkq623.
106. **Yakhnin AV, Yakhnin H, Babitzke P.** 2008. Function of the *Bacillus subtilis* transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader. Proc Natl Acad Sci U S A **105:**16131–16136. http://dx.doi.org/10.1073/pnas.0808842105.
107. **Anderson SJ, Sikes ML, Zhang Y, French SL, Salgia S, Beyer AL, Nomura M, Schneider DA.** 2011. The transcription elongation factor Spt5 influences transcription by RNA polymerase I positively and negatively. J Biol Chem **286:**18816–18824. http://dx.doi.org/10.1074/jbc.M110.202101.
108. **Tomar SK, Artsimovitch I.** 2013. NusG-Spt5 proteins: universal tools for transcription modification and communication. Chem Rev **113:**8604–8619. http://dx.doi.org/10.1021/cr400064k.
109. **Chalissery J, Muteeb G, Kalarickal NC, Mohan S, Jisha V, Sen R.** 2011. Interaction surface of the transcription terminator Rho required to form a complex with the C-terminal domain of the antiterminator NusG. J Mol Biol **405:**49–64. http://dx.doi.org/10.1016/j.jmb.2010.10.044.
110. **Burmann BM, Schweimer K, Luo X, Wahl MC, Stitt BL, Gottesman ME, Rösch P.** 2010. A NusE:NusG complex links transcription and translation. Science **501:**501–504.
111. **Proshkin S, Rahmouni AR, Mironov A, Nudler E.** 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science **328:**504–508. http://dx.doi.org/10.1126/science.1184939.
112. **Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, Landick R.** 2012. Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. Genes Dev **26:**2621–2633. http://dx.doi.org/10.1101/gad.196741.112.
113. **Burns CM, Richardson JP.** 1995. NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. Proc Natl Acad Sci U S A **92:**4738–4742. http://dx.doi.org/10.1073/pnas.92.11.4738.
114. **French SL, Santangelo TJ, Beyer AL, Reeve JN.** 2007. Transcription and translation are coupled in archaea. Mol Biol Evol **24:**893–895. http://dx.doi.org/10.1093/molbev/msm007.
115. **Santangelo TJ, Cubonová L, Matsumi R, Atomi H, Imanaka T, Reeve JN.** 2008. Polarity in archaeal operon transcription in *Thermococcus kodakaraensis*. J Bacteriol **190:**2244–2248. http://dx.doi.org/10.1128/JB.01811-07.
116. **Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD.** 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science **292:**1876–1882. http://dx.doi.org/10.1126/science.1059495.
117. **Schulz S, Gietl A, Smollett K, Tinnefeld P, Werner F, Grohmann D.** 2016. TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. Proc Natl Acad Sci U S A **113:**E1816–E1825.
118. **Grohmann D, Klose D, Fielden D, Werner F.** 2011. FRET (fluorescence resonance energy transfer) sheds light on transcription. Biochem Soc Trans **39:**122–127. http://dx.doi.org/10.1042/BST0390122.
119. **Tanaka A, Akimoto Y, Kobayashi S, Hisatake K, Hanaoka F, Ohkuma Y.** 2015. Association of the winged helix motif of the TFIIEα subunit of TFIIE with either the TFIIEβ subunit or TFIIB distinguishes its functions in transcription. Genes Cells **20:**203–216. http://dx.doi.org/10.1111/gtc.12212.
120. **Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I.** 2007. Structural basis for transcription elongation by bacterial RNA polymerase. Nature **448:**157–162. http://dx.doi.org/10.1038/nature05932.
121. **Kettenberger H, Armache K-J, Cramer P.** 2004. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol Cell **16:**955–965. http://dx.doi.org/10.1016/j.molcel.2004.11.040.
122. **Kireeva ML, Komissarova N, Waugh DS, Kashlev M.** 2000. The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J Biol Chem **275:**6530–6536. http://dx.doi.org/10.1074/jbc.275.9.6530.
123. **Andrecka J, Treutlein B, Arcusa MAI, Muschielok A, Lewis R, Cheung ACM, Cramer P, Michaelis J.** 2009. Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. Nucleic Acids Res **37:**5803–5809. http://dx.doi.org/10.1093/nar/gkp601.
124. **Reeder TC, Hawley DK.** 1996. Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. Cell **87:**767–777. http://dx.doi.org/10.1016/S0092-8674(00)81395-1.
125. **Nudler E, Avetissova E, Markovtsov V, Goldfarb A.** 1996. Transcription processivity: protein-DNA interactions holding together the elongation complex. Science **273:**211–217. http://dx.doi.org/10.1126/science.273.5272.211.
126. **Sidorenkov I, Komissarova N, Kashlev M.** 1998. Crucial role of the RNA:DNA hybrid in the processivity of transcription. Mol Cell **2:**55–64. http://dx.doi.org/10.1016/S1097-2765(00)80113-6.
127. **Komissarova N, Becker J, Solter S, Kireeva M, Kashlev M.** 2002. Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Mol Cell **10:**1151–1162. http://dx.doi.org/10.1016/S1097-2765(02)00738-4.
128. **Hein PP, Kolb KE, Windgassen T, Bellecourt MJ, Darst SA, Mooney RA, Landick R.** 2014. RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. Nat Struct Mol Biol **21:**794–802. http://dx.doi.org/10.1038/nsmb.2867.
129. **Bochkareva A, Yuzenkova Y, Tadigotla VR, Zenkin N.** 2012. Factor-independent transcription pausing caused by recognition of the RNA-DNA hybrid sequence. EMBO J **31:**630–639. http://dx.doi.org/10.1038/emboj.2011.432.
130. **Gusarov I, Nudler E.** 1999. The mechanism of intrinsic transcription termination. Mol Cell **3:**495–504. http://dx.doi.org/10.1016/S1097-2765(00)80477-3.
131. **Martin FH, Tinoco I.** 1980. DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. Nucleic Acids Res **8:**2295–2300. http://dx.doi.org/10.1093/nar/8.10.2295.
132. **Czyz A, Mooney RA, Iaconi A, Landick R.** 2014. Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. mBio **5:**e00931.
133. **Arimbasseri AG, Rijal K, Maraia RJ.** Transcription termination by the eukaryotic RNA polymerase III. Biochim Biophys Acta **1829:**318–330.
134. **Reiter WD, Palm P, Zillig W.** 1988. Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. Nucleic Acids Res **16:**2445–2459. http://dx.doi.org/10.1093/nar/16.6.2445.
135. **Santangelo TJ, Roberts JW.** 2004. Forward translocation is the natural pathway of RNA release at an intrinsic terminator. Mol Cell **14:**117–126. http://dx.doi.org/10.1016/S1097-2765(04)00154-6.
136. **Reynolds R, Chamberlin MJ.** 1992. Parameters affecting transcription termination by *Escherichia coli* RNA. II: construction and analysis of hybrid terminators. J Mol Biol **224:**53–63.
137. **Dieci G, Sentenac A.** 1996. Facilitated recycling pathway for RNA polymerase III. Cell **84:**245–252. http://dx.doi.org/10.1016/S0092-8674(00)80979-4.
138. **Spitalny P, Thomm M.** 2008. A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. Mol Microbiol **67:**958–970. http://dx.doi.org/10.1111/j.1365-2958.2007.06084.x.
139. **Dedrick RL, Kane CM, Chamberlin MJ.** 1987. Purified RNA polymerase II recognizes specific termination sites during transcription in vitro. J Biol Chem **262:**9098–9108.

140. **Palangat M, Hittinger CT, Landick R.** 2004. Downstream DNA selectively affects a paused conformation of human RNA polymerase II. J Mol Biol **341:**429–442. http://dx.doi.org/10.1016/j.jmb.2004.06.009.
141. **Kerppola TK, Kane CM.** 1988. Intrinsic sites of transcription termination and pausing in the *c-myc* gene. Mol Cell Biol **8:**4389–4394. http://dx.doi.org/10.1128/MCB.8.10.4389.
142. **Roberts JW.** 1969. Termination factor for RNA synthesis. Nature **224:**1168–1174. http://dx.doi.org/10.1038/2241168a0.
143. **Park J-S, Marr MT, Roberts JW.** 2002. E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. Cell **109:**757–767. http://dx.doi.org/10.1016/S0092-8674(02)00769-9.
144. **Fong N, Brannan K, Erickson B, Kim H, Cortazar MA, Sheridan RM, Nguyen T, Karp S, Bentley DL.** 2015. Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. Mol Cell **60:**256–267. http://dx.doi.org/10.1016/j.molcel.2015.09.026.
145. **Merkl P, Perez-Fernandez J, Pilsl M, Reiter A, Williams L, Gerber J, Böhm M, Deutzmann R, Griesenbeck J, Milkereit P, Tschochner H.** 2014. Binding of the termination factor Nsi1 to its cognate DNA site is sufficient to terminate RNA polymerase I transcription *in vitro* and to induce termination *in vivo*. Mol Cell Biol **34:**3817–3827. http://dx.doi.org/10.1128/MCB.00395-14.
146. **Howan K, Smith AJ, Westblade LF, Joly N, Grange W, Zorman S, Darst SA, Savery NJ, Strick TR.** 2012. Initiation of transcription-coupled repair characterized at single-molecule resolution. Nature **490:**431–434. http://dx.doi.org/10.1038/nature11430.
147. **Leela JK, Syeda AH, Anupama K, Gowrishankar J.** 2013. Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. Proc Natl Acad Sci U S A **110:**258–263. http://dx.doi.org/10.1073/pnas.1213123110.
148. **Smith AJ, Pernstich C, Savery NJ.** 2012. Multipartite control of the DNA translocase, Mfd. Nucleic Acids Res **40:**10408–10416. http://dx.doi.org/10.1093/nar/gks775.
149. **Park J-S, Roberts JW.** 2006. Role of DNA bubble rewinding in enzymatic transcription termination. Proc Natl Acad Sci U S A **103:**4870–4875. http://dx.doi.org/10.1073/pnas.0600145103.
150. **Haines NM, Kim Y-IT, Smith AJ, Savery NJ.** 2014. Stalled transcription complexes promote DNA repair at a distance. Proc Natl Acad Sci U S A **111:**4037–4042. http://dx.doi.org/10.1073/pnas.1322350111.
151. **Adhya S, Gottesman M, De Crombrugghe B.** 1974. Release of polarity in *Escherichia coli* by gene *N* of phage lambda: termination and antitermination of transcription. Proc Natl Acad Sci U S A **71:**2534–2538. http://dx.doi.org/10.1073/pnas.71.6.2534.
152. **Menouni R, Champ S, Espinosa L, Boudvillain M, Ansaldi M.** 2013. Transcription termination controls prophage maintenance in *Escherichia coli* genomes. Proc Natl Acad Sci U S A **110:**14414–14419. http://dx.doi.org/10.1073/pnas.1303400110.
153. **Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, Nudler E.** 2008. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. Science **320:**935–938. http://dx.doi.org/10.1126/science.1152763.
154. **Witkin EM.** 1966. Radiation-induced mutations and their repair. Science **152:**1345–1353. http://dx.doi.org/10.1126/science.152.3727.1345.
155. **Peeters E, Driessen RPC, Werner F, Dame RT.** 2015. The interplay between nucleoid organization and transcription in archaeal genomes. Nat Rev Microbiol **13:**333–341. http://dx.doi.org/10.1038/nrmicro3467.
156. **Zerulla K, Soppa J.** 2014. Polyploidy in haloarchaea: advantages for growth and survival. Front Microbiol **5:**274.
157. **Spaans SK, van der Oost J, Kengen SWM.** 2015. The chromosome copy number of the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1. Extremophiles **19:**741–750. http://dx.doi.org/10.1007/s00792-015-0750-5.
158. **Jaakkola ST, Zerulla K, Guo Q, Liu Y, Ma H, Yang C, Bamford DH, Chen X, Soppa J, Oksanen HM.** 2014. Halophilic archaea cultivated from surface sterilized middle-late Eocene rock salt are polyploid. PLoS One **9:**e110533. http://dx.doi.org/10.1371/journal.pone.0110533.
159. **Hildenbrand C, Stock T, Lange C, Rother M, Soppa J.** 2011. Genome copy numbers and gene conversion in methanogenic archaea. J Bacteriol **193:**734–743. http://dx.doi.org/10.1128/JB.01016-10.
160. **Breuert S, Allers T, Spohn G, Soppa J.** 2006. Regulated polyploidy in halophilic archaea. PLoS One **1:**e92. http://dx.doi.org/10.1371/journal.pone.0000092.
161. **Pereira SL, Grayling RA, Lurz R, Reeve JN.** 1997. Archaeal nucleosomes. Proc Natl Acad Sci U S A **94:**12633–12637. http://dx.doi.org/10.1073/pnas.94.23.12633.
162. **Driessen RPC, Dame RT.** 2013. Structure and dynamics of the crenarchaeal nucleoid. Biochem Soc Trans **41:**321–325. http://dx.doi.org/10.1042/BST20120336.
163. **Driessen RPC, Dame RT.** 2011. Nucleoid-associated proteins in *Crenarchaea*. Biochem Soc Trans **39:**116–121. http://dx.doi.org/10.1042/BST0390116.
164. **Ammar R, Torti D, Tsui K, Gebbia M, Durbic T, Bader GD, Giaever G, Nislow C.** 2012. Chromatin is an ancient innovation conserved between Archaea and Eukarya. Elife **1:**e00078.
165. **Nalabothula N, Xi L, Bhattacharyya S, Widom J, Wang JP, Reeve JN, Santangelo TJ, Fondufe-Mittendorf YN.** 2013. Archaeal nucleosome positioning *in vivo* and *in vitro* is directed by primary sequence motifs. BMC Genomics **14:**391. http://dx.doi.org/10.1186/1471-2164-14-391.
166. **Forbes AJ, Patrie SM, Taylor GK, Kim Y-B, Jiang L, Kelleher NL.** 2004. Targeted analysis and discovery of posttranslational modifications in proteins from methanogenic archaea by top-down MS. Proc Natl Acad Sci U S A **101:**2678–2683. http://dx.doi.org/10.1073/pnas.0306575101.
167. **Maruyama H, Harwood JC, Moore KM, Paszkiewicz K, Durley SC, Fukushima H, Atomi H, Takeyasu K, Kent NA.** 2013. An alternative beads-on-a-string chromatin architecture in *Thermococcus kodakarensis*. EMBO Rep **14:**711–717. http://dx.doi.org/10.1038/embor.2013.94.
168. **Xie Y, Reeve JN.** 2004. Transcription by an archaeal RNA polymerase is slowed but not blocked by an archaeal nucleosome. J Bacteriol **186:**3492–3498. http://dx.doi.org/10.1128/JB.186.11.3492-3498.2004.
169. **Studitsky VM.** 1997. Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. Science **278:**1960–1963. http://dx.doi.org/10.1126/science.278.5345.1960.
170. **Clark DJ, Felsenfeld G.** 1992. A nucleosome core is transferred out of the path of a transcribing polymerase. Cell **71:**11–22. http://dx.doi.org/10.1016/0092-8674(92)90262-B.
171. **Dulmage KA, Todor H, Schmid AK.** 2015. Growth-phase-specific modulation of cell morphology and gene expression by an archaeal histone protein. mBio **6:**e00649–15.
172. **Weidenbach K, Glöer J, Ehlers C, Sandman K, Reeve JN, Schmitz RA.** 2008. Deletion of the archaeal histone in *Methanosarcina mazei* Gö1 results in reduced growth and genomic transcription. Mol Microbiol **67:**662–671. http://dx.doi.org/10.1111/j.1365-2958.2007.06076.x.
173. **Heinicke I, Müller J, Pittelkow M, Klein A.** 2004. Mutational analysis of genes encoding chromatin proteins in the archaeon *Methanococcus voltae* indicates their involvement in the regulation of gene expression. Mol Genet Genomics **272:**76–87.
174. **Čuboňováa L, Katano M, Kanai T, Atomi H, Reeve JN, Santangelo TJ.** 2012. An archaeal histone is required for transformation of *Thermococcus kodakarensis*. J Bacteriol **194:**6864–6874. http://dx.doi.org/10.1128/JB.01523-12.
175. **Dinger ME, Baillie GJ, Musgrave DR.** 2000. Growth phase-dependent expression and degradation of histones in the thermophilic archaeon *Thermococcus zilligii*. Mol Microbiol **36:**876–885. http://dx.doi.org/10.1046/j.1365-2958.2000.01904.x.
176. **Sandman K, Grayling RA, Dobrinski B, Lurz R, Reeve JN.** 1994. Growth-phase-dependent synthesis of histones in the archaeon *Methanothermus fervidus*. Proc Natl Acad Sci U S A **91:**12624–12628. http://dx.doi.org/10.1073/pnas.91.26.12624.
177. **Soares D, Dahlke I, Li W-T, Sandman K, Hethke C, Thomm M, Reeve JN.** 1998. Archaeal histone stability, DNA binding, and transcription inhibition above 90°C. Extremophiles **2:**75–81. http://dx.doi.org/10.1007/s007920050045.
178. **Jelinska C, Petrovic-Stojanovska B, Ingledew WJ, White MF.** 2010. Dimer-dimer stacking interactions are important for nucleic acid binding by the archaeal chromatin protein Alba. Biochem J **427:**49–55. http://dx.doi.org/10.1042/BJ20091841.
179. **Lurz R, Grote M, Dijk J, Reinhardt R, Dobrinski B.** 1986. Electron microscopic study of DNA complexes with proteins from the archaebacterium *Sulfolobus acidocaldarius*. EMBO J **5:**3715–3721.
180. **Jelinska C, Conroy MJ, Craven CJ, Hounslow AM, Bullough PA, Waltho JP, Taylor GL, White MF.** 2005. Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control

of DNA packaging. Structure **13:**963–971. http://dx.doi.org/10.1016/j.str.2005.04.016.
181. **Laurens N, Driessen RPC, Heller I, Vorselen D, Noom MC, Hol FJH, White MF, Dame RT, Wuite GJL.** 2012. Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA. Nat Commun **3:**1328. http://dx.doi.org/10.1038/ncomms2330.
182. **Tanaka T, Padavattan S, Kumarevel T.** 2012. Crystal structure of archaeal chromatin protein Alba2-double-stranded DNA complex from *Aeropyrum pernix* K1. J Biol Chem **287:**10394–10402. http://dx.doi.org/10.1074/jbc.M112.343210.
183. **Wilkinson SP, Ouhammouch M, Geiduschek EP.** 2010. Transcriptional activation in the context of repression mediated by archaeal histones. Proc Natl Acad Sci U S A **107:**6777–6781. http://dx.doi.org/10.1073/pnas.1002360107.
184. **Jäger D, Pernitzsch SR, Richter AS, Backofen R, Sharma CM, Schmitz RA.** 2012. An archaeal sRNA targeting cis-and transencoded mRNAs via two distinct domains. Nucleic Acids Res **40:**10964–10979. http://dx.doi.org/10.1093/nar/gks847.
185. **Jäger D, Förstner KU, Sharma CM, Santangelo TJ, Reeve JN.** 2014. Primary transcriptome map of the hyperthermophilic archaeon *Thermococcus kodakarensis*. BMC Genomics **15:**684. http://dx.doi.org/10.1186/1471-2164-15-684.
186. **Wurtzel O, Sapra R, Chen F, Zhu Y, Simmons BA, Sorek R.** 2010. A single-base resolution map of an archaeal transcriptome. Genome Res **20:**133–141. http://dx.doi.org/10.1101/gr.100396.109.
187. **Toffano-Nioche C, Ott A, Crozat E, Nguyen AN, Zytnicki M, Leclerc F, Forterre P, Bouloc P, Gautheret D.** 2013. RNA at 92°C: the noncoding transcriptome of the hyperthermophilic archaeon *Pyrococcus abyssi*. RNA Biol **10:**1211–1220. http://dx.doi.org/10.4161/rna.25567.
188. **Heyer R, Dörr M, Jellen-Ritter A, Späth B, Babski J, Jaschinski K, Soppa J, Marchfelder A.** 2012. High throughput sequencing reveals a plethora of small RNAs, including tRNA derived fragments in *Haloferax volcanii*. RNA Biol **9:**1011–1018. http://dx.doi.org/10.4161/rna.20826.
189. **López Muñoz MM, Schönheit P, Metcalf WW.** 2015. Genetic, genomic, and transcriptomic studies of pyruvate metabolism in *Methanosarcina barkeri* fusaro. J Bacteriol **197:**3592–3600. http://dx.doi.org/10.1128/JB.00551-15.
190. **Jäger D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA.** 2009. Deep sequencing analysis of the *Methanosarcina mazei* Go1 transcriptome in response to nitrogen availability. Proc Natl Acad Sci U S A **106:**21878–21882. http://dx.doi.org/10.1073/pnas.0909051106.
191. **Vannier P, Michoud G, Oger P, Marteinsson VÞ Jebbar M.** 2015. Genome expression of *Thermococcus barophilus* and *Thermococcus kodakarensis* in response to different hydrostatic pressure conditions. Res Microbiol **166:**717–725. http://dx.doi.org/10.1016/j.resmic.2015.07.006.
192. **Esclapez J, Pire C, Camacho M, Bautista V, Martínez-Espinosa RM, Zafrilla B, Vegara A, Alcaraz LA, Bonete MJ.** 2015. Transcriptional profiles of *Haloferax mediterranei* based on nitrogen availability. J Biotechnol **193:**100–107. http://dx.doi.org/10.1016/j.jbiotec.2014.11.018.
193. **Lagorce A, Fourçans A, Dutertre M, Bouyssiere B, Zivanovic Y, Confalonieri F.** 2012. Genome-wide transcriptional response of the archaeon *Thermococcus gammatolerans* to cadmium. PLoS One **7:**e41935. http://dx.doi.org/10.1371/journal.pone.0041935.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

MINIREVIEW

Mycobacterium tuberculosis Transcription Machinery: Ready To Respond to Host Attacks

Kelly Flentie, Ashley L. Garner, Christina L. Stallings

Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, USA

Regulating responses to stress is critical for all bacteria, whether they are environmental, commensal, or pathogenic species. For pathogenic bacteria, successful colonization and survival in the host are dependent on adaptation to diverse conditions imposed by the host tissue architecture and the immune response. Once the bacterium senses a hostile environment, it must enact a change in physiology that contributes to the organism's survival strategy. Inappropriate responses have consequences; hence, the execution of the appropriate response is essential for survival of the bacterium in its niche. Stress responses are most often regulated at the level of gene expression and, more specifically, transcription. This minireview focuses on mechanisms of regulating transcription initiation that are required by *Mycobacterium tuberculosis* to respond to the arsenal of defenses imposed by the host during infection. In particular, we highlight how certain features of *M. tuberculosis* physiology allow this pathogen to respond swiftly and effectively to host defenses. By enacting highly integrated and coordinated gene expression changes in response to stress, *M. tuberculosis* is prepared for battle against the host defense and able to persist within the human population.

The survival of any organism relies on its ability to sense and respond to changes in its environment. For bacteria, stress responses are primarily mediated through the regulation of gene expression. By integrating multiple molecular approaches to gene regulation, pathogenic bacteria are able to orchestrate condition-specific patterns that promote survival and pathogenesis in the face of a strong immune response. This minireview focuses on mechanisms of transcription regulation required for stress responses in one of the most successful and deadly pathogens in the world, *Mycobacterium tuberculosis*. *M. tuberculosis* has coexisted with humans for >50,000 years (1) and continues to cause more than 1.5 million deaths a year (2). The coevolution of *M. tuberculosis* with the human host response to infection has resulted in a pathogen that is specialized for long-term infection in people. Tuberculosis is a complex disease that requires the bacteria to multiply within phagocytes, survive extracellularly in hypoxic and necrotic granulomas, and endure a robust immune response to persist in the host. During infection, the host immune response restrains *M. tuberculosis* from proliferating by imposing a battery of defenses, including reactive oxygen and nitrogen stress, hypoxia, acid stress, genotoxic stress, cell surface stress, and starvation (3). Despite this onslaught of attacks, *M. tuberculosis* is able to persist for the lifetime of the host, indicating that this pathogen has highly effective molecular mechanisms to resist host-inflicted damage. In order to enact these defenses and facilitate this specialized lifestyle, *M. tuberculosis* executes a complex, interconnected web of stress responses that rely on changes in gene expression. In fact, *M. tuberculosis* is well suited to respond quickly to diverse stresses in a coordinated fashion. For instance, the RNA polymerase (RNAP) bears kinetic properties that allow it to be easily modulated by accessory factors. Compared to other obligate human pathogens, *M. tuberculosis* encodes the highest ratio of σ factors to genome size (4), which allows the bacterium to tailor its expression profile in response to a given environment. Even during exponential growth in culture, traditionally thought of as a relatively stress-free environment, *M. tuberculosis* expresses its entire complement of σ factors (5–7), indicating that *M. tuberculosis* is poised to quickly respond to stress. *M. tuberculosis* also integrates stress responses into basic cellular processes; as a result, some stress-associated transcriptional regulators are essential in *M. tuberculosis*. In this minireview, we discuss features of the mycobacterial transcription apparatus that position *M. tuberculosis* to be ready to respond to host attacks, the networks of factors that contribute to these responses, and how this culminates in a successful pathogenic strategy. The general strategies to be discussed are illustrated in Fig. 1, and individual factors touched on in this minireview are summarized in Fig. 2.

THE MYCOBACTERIAL RNA POLYMERASE—READY TO RESPOND

Transcription is achieved in all bacteria by a single core RNAP enzyme, consisting of the essential subunits β and β′ and 2 α subunits along with the nonessential ω subunit (8, 9). To recognize and bind promoter sequences upstream from genes, the core RNAP associates with a σ subunit to form an RNAP holoenzyme. Most transcriptional regulation occurs at the level of initiation (10), and transcription factors (TFs) can mediate this regulation by directly affecting the polymerase-promoter interaction, manipulating the equilibrium between closed and open RNAP-promoter complexes (RP_c and RP_o, respectively), or affecting rates of promoter escape (11, 12). The majority of studies on the mechanisms of transcription initiation and its regulation have used *Escherichia coli* as a model system. However, multiple groups have recently shown that *Mycobacterium bovis* RNAP, which differs from the *M. tuberculosis* RNAP by only one amino acid (aa), exhibits an inherently unstable RP_o complex compared to *E. coli* RNAP on the same promoter (13, 14). In these reports, saturating

Accepted manuscript posted online 16 February 2016

Citation Flentie K, Garner AL, Stallings CL. 2016. *Mycobacterium tuberculosis* transcription machinery: ready to respond to host attacks. J Bacteriol 198:1360–1373. doi:10.1128/JB.00935-15.

Editor: W. Margolin

Address correspondence to Christina L. Stallings, stallings@wusm.wustl.edu.

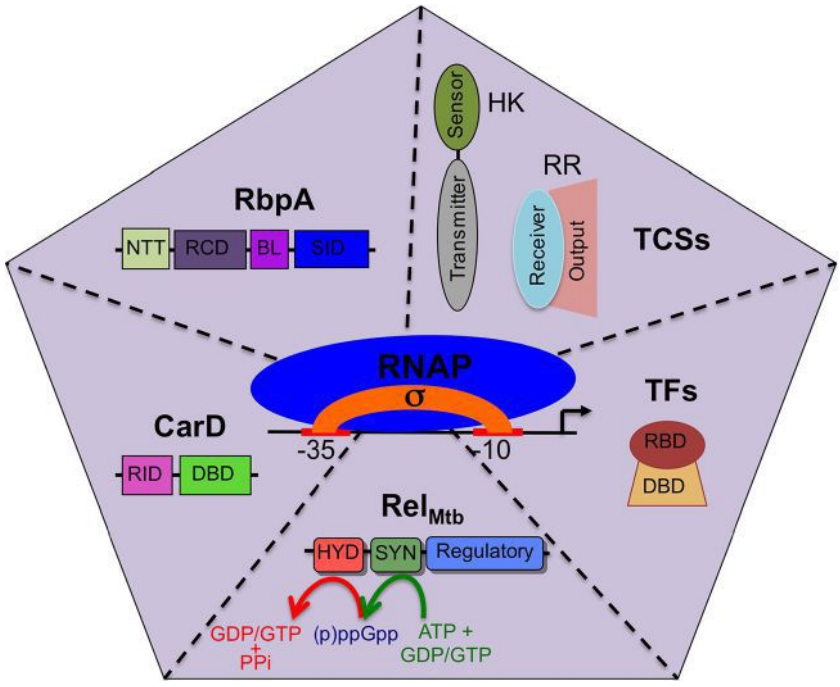


FIG 1 Summary of the branches of transcriptional regulation that are discussed in this minireview. The illustration shows 6 types of factors (σ factors, CarD, RbpA, TCSs, TFs, and Rel_{Mtb}) that modulate RNAP activity at promoters to mediate reprogramming of the expression profile in *M. tuberculosis* in response to different environments. A σ factor associates with the core RNAP to form the RNAP holoenzyme, which is then modified by the other factors shown in the sections of the pentagon. Domains of each protein are shown. For CarD, RID is the RNAP interaction domain and DBD is the DNA binding domain. For RbpA, NTT is the N-terminal tail, RCD is the RbpA core domain, BL is the basic linker, and SID is the sigma interaction domain. For Rel_{Mtb}, HYD is the (p)ppGpp hydrolase domain and SYN is the (p)ppGpp synthetase domain. For TFs, RBD is the RNAP binding domain and DBD is the DNA binding domain. In the presence of a given stress, these factors coordinate their responses to effectively respond to host attacks.

concentrations of *M. bovis* RNAP σ^A holoenzyme were found to be incapable of opening a large percentage of the promoters, leaving the majority of bound complexes in the closed state. It has been proposed (14) that the presence or absence of lineage-specific insertions within RNAP could contribute to the inherent differences in stability of the promoter complexes formed by *M. bovis* versus *E. coli* RNAP. Notably, RNAPs from *Bacillus subtilis*, *Thermus aquaticus*, and *Thermus thermophilus* have also been found to generate relatively unstable open promoter complexes (15–17). Based on these observations, it is worth considering that the properties of *E. coli* RNAP may not be representative of most bacterial RNAPs and that there may be significant lineage-specific variation in enzyme kinetics. The inherent instability of RNAP-promoter complexes would allow the mycobacterial RNAP to be poised to respond to changes in the environment by being easily modified in activity by additional factors.

σ FACTORS: THE GENERALS OF STRESS RESPONSES

The first determinant of gene expression in response to different conditions is the activity of the σ factor repertoire. Each σ factor binds a specific promoter sequence, thus determining what promoters are targeted by the RNAP holoenzyme for transcription. Changes in σ factor activity in response to different stresses and conditions are able to shift a bacterium's expression profile. The σ factor network of *M. tuberculosis* includes one essential housekeeping group 1 σ factor (σ^A), one stress-responsive group 2 σ factor (σ^B), and 11 group 3 and 4 alternative σ factors that also function as environmentally responsive regulators (σ^C to σ^M) (4, 6, 18). This broad panel of σ factors allows *M. tuberculosis* to tune its transcriptional response for a large and diverse set of conditions. All of the σ factors in *M. tuberculosis* belong to the σ^{70} family, whose members in *E. coli* recognize two sequences in the promoter DNA, the −10 element (recognized by sigma region 2.4) and the −35 element (recognized by sigma region 4.2) (19). *M. tuberculosis* promoters contain a conserved −10 sequence that is essential and sometimes sufficient for transcription, while the −35 sequences are less conserved (19–21). The spacer region between the −10 and −35 elements in *M. tuberculosis* also varies dramatically compared to *E. coli* promoters (19, 22, 23). These differences in promoter elements may reflect the sigma diversity in *M. tuberculosis* (19, 23).

The activity of σ factors in *M. tuberculosis* is most often regulated by anti-σ factors that inactivate their cognate σ factors until a signal is received to liberate the σ factor for action. Specifically, σ^B, σ^D, σ^E, σ^F, σ^H, σ^K, σ^L, and σ^M are all regulated by a cognate anti-σ factor (24–32). A putative anti-σ factor has also been proposed for σ^G (33). To investigate under which conditions a particular σ factor is active, the expression levels of σ factors have been studied *in vitro* under many physiologically relevant conditions, but transcriptional upregulation of a given σ factor does not necessarily equate to σ factor activity. Therefore, σ factor gene deletion or overexpression strains have been used to determine the functional role of individual σ factors in response to stress. These data are summarized here and together paint a picture of an intricate circuitry of transcriptional regulation that integrates multiple σ factor regulons under many conditions (Fig. 3 and 4), allowing *M. tuberculosis* to respond to the arsenal of attacks from the host.

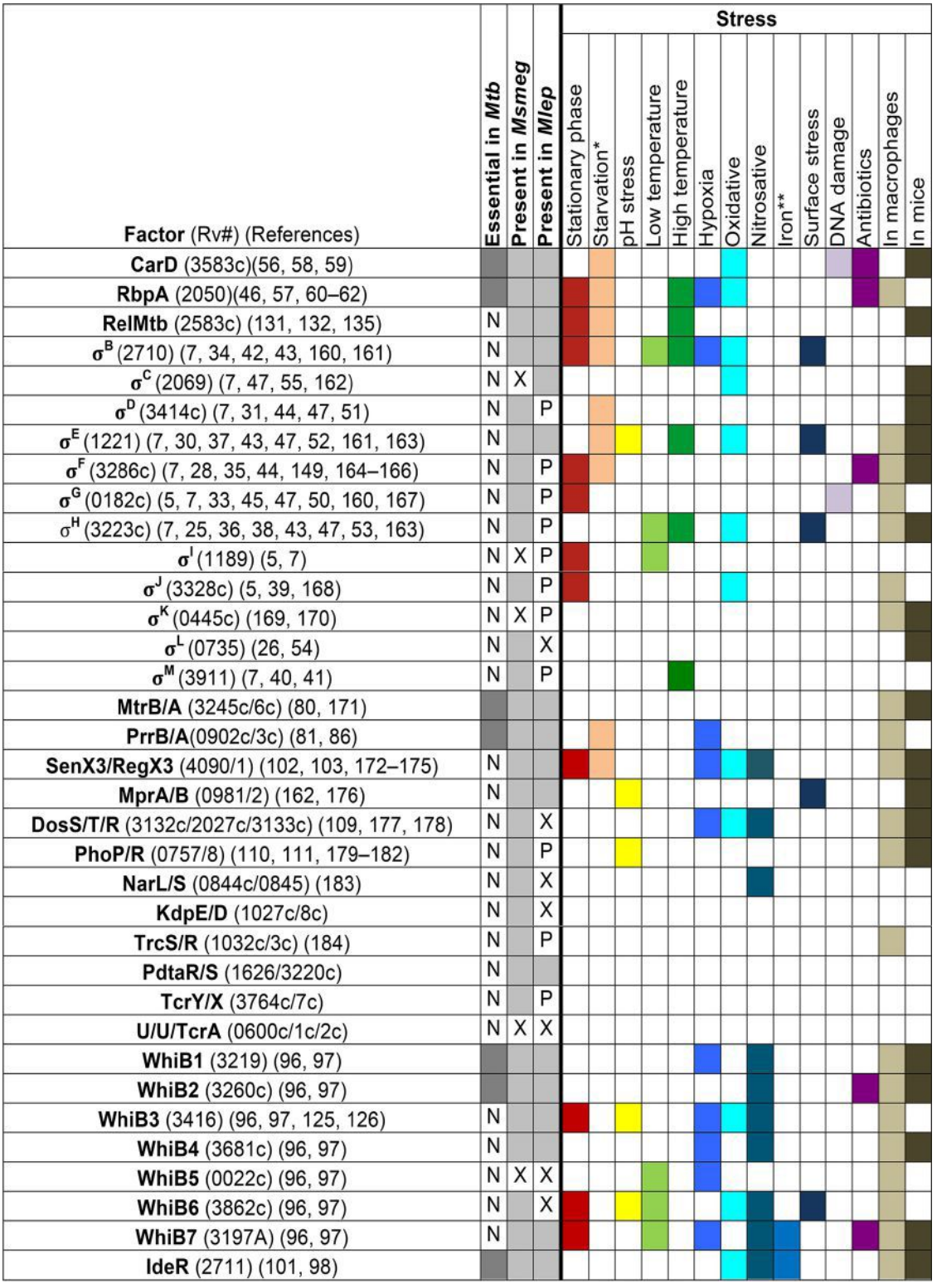


FIG 2 Conservation of *M. tuberculosis* regulatory factors and the stresses that the factors are associated with in *M. tuberculosis* (160–184). The left side of the table designates whether the gene for a transcriptional regulator is essential (shaded) or not essential (N) in *M. tuberculosis* (*Mtb*) and whether that gene is conserved (shaded) or not conserved (X) or exists as a pseudogene (P) in the environmental saprophytic *M. smegmatis* (*Msmeg*) or the obligate pathogen *M. leprae* (*Mlep*). The right side of the table indicates whether a particular stress condition has been associated with a given transcriptional regulator. Involvement in the response to a particular stress is designated by shading of the box and may represent expression profiling data or phenotypic analysis of mutants. An unshaded square indicates that the factor is not induced, is not important for survival, or has not been studied under that particular condition. U, unnamed factor; *, starvation (including nutrient, phosphate, and nitrogen starvation); **, iron-depleted or iron-replete conditions. See specific references for more information.

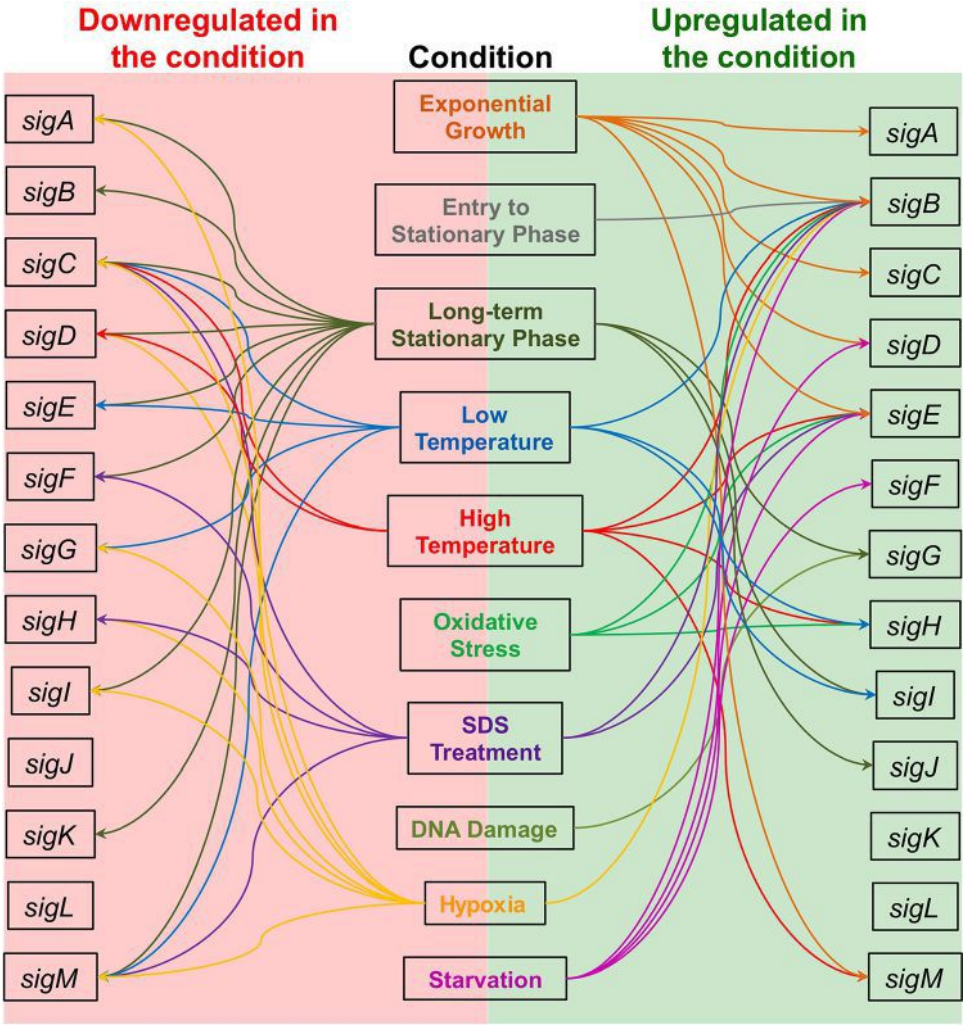


FIG 3 Transcriptional regulation of *M. tuberculosis* σ factor genes in response to various stresses. Transcriptional responses of σ factor genes of *M. tuberculosis* include responses to entry to stationary phase, the long-term stationary phase, mild cold shock (room temperature), heat shock (45°C), oxidative stress, exposure to SDS, DNA damage, hypoxia, and starvation. The σ factor genes that are transcriptionally upregulated in response to a stress are diagramed with arrows to the right, and the σ factor genes that are transcriptionally downregulated are shown with arrows to the left. The σ factor genes that are highly expressed during exponential growth in culture are shown as being upregulated under this condition. Where no arrow is present to connect a σ factor gene to a particular stress, this indicates that expression of the σ factor gene is not significantly changed during exposure to that stress or has not been studied under that particular condition. References are available in the text.

During exponential growth of *M. tuberculosis* in culture, *sigA*, *sigB*, *sigC*, *sigD*, *sigE*, and *sigM* are the most highly expressed σ factor genes (7). Upon entry into stationary phase, levels of *sigB* transcripts increase (34). Strains with a disrupted *sigF* gene grow to a density three times greater than that seen with wild-type cultures in stationary phase, suggesting that σ^F may have a key role in regulating this transition (35). Later in stationary phase, there is a global change in regulation of σ factors resulting in downregulation of most of the σ factor genes, with the exception of *sigG*, *sigI*, and *sigJ*, which are upregulated in long-term stationary cultures (5, 7). σ^H is a central regulator of the response of *M. tuberculosis* to both heat and oxidative stress through regulation of *sigE*, *sigB*, heat shock proteins, thioredoxin reductase/thioredoxin, and synthesis of mycothiol precursors (36). In addition to σ^B, σ^E, and σ^H, survival during oxidative stress is also dependent on σ^C and σ^J (6, 36–39). *sigM* is also induced during exposure to heat in the *M. tuberculosis* CDC1551 strain but not in *M. tuberculosis* H37Rv, indicating strain-specific regulation of σ factor expression (24, 40, 41). Cold temperatures induce expression of *sigB*, *sigH*, and *sigI* while repressing transcription of *sigC*, *sigE*, *sigG*, and *sigM* (7). σ^I is the most highly induced σ factor during cold shock and has been proposed to be important for the bacterium's survival in aerosol particles between hosts (7). Deletion of *sigB*, *sigE*, or *sigH* has been shown to increase *M. tuberculosis*'s sensitivity to cell surface stress (6, 37, 42, 43). Expression of *sigB* is also upregulated under hypoxic conditions (7) and σ^B is the only σ factor shown to impact the sensitivity of *M. tuberculosis* to hypoxia (42). Deletion of *sigF* induces permeability changes in the cell envelope, although this does not affect sensitivity to tested surface stresses (35, 44). *In vitro* studies have shown that *sigG* is induced upon DNA damage but that deletion of *sigG* does not sensitize strains to DNA damage (45). *sigB*, *sigD*, *sigE*, and *sigF* have all been shown to be upregulated during prolonged nutrient starvation (46).

Evidence that alternative σ factors are important in *M. tuber-*

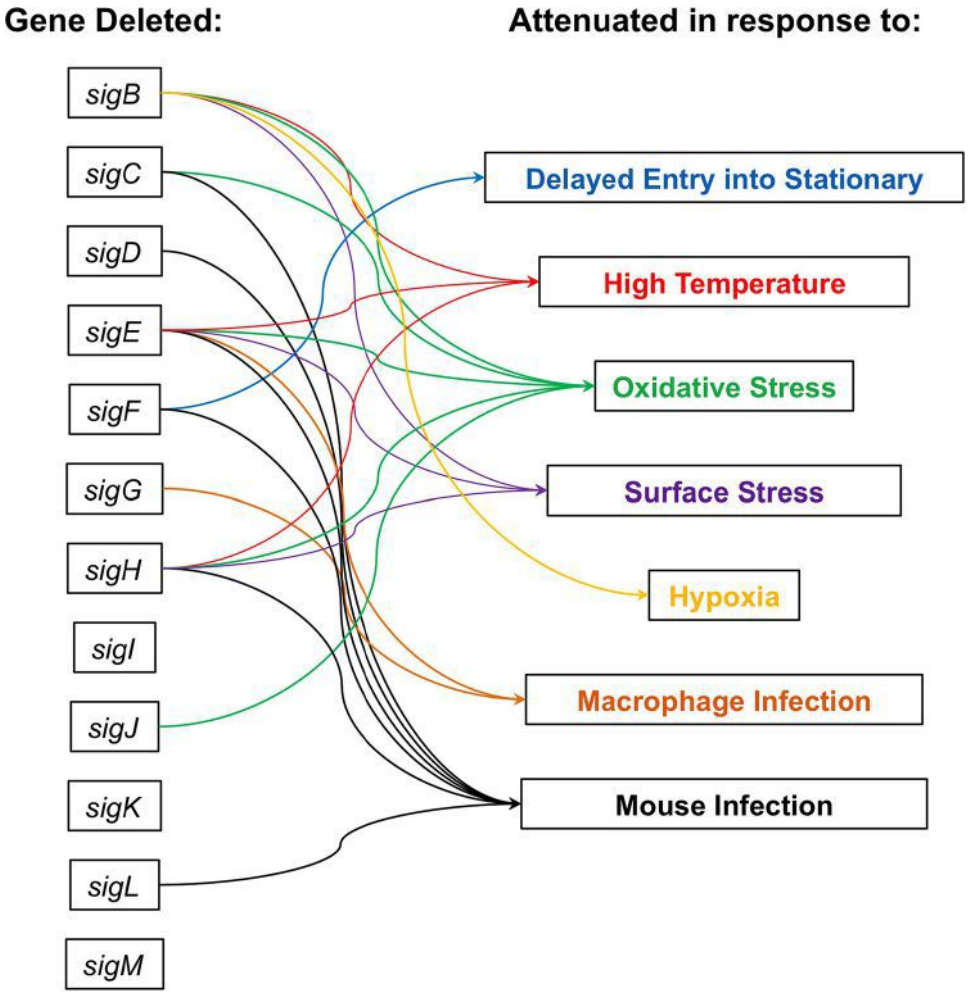


FIG 4 Effects of σ factor gene deletions on stress responses in *M. tuberculosis*. Arrows indicate whether deletion of a σ factor gene causes delayed entry into stationary phase, decreased survival during heat shock, decreased survival during oxidative stress, decreased survival during surface stress or changes in cell permeability, decreased survival during hypoxia, decreased survival in macrophages, or decreased immunopathology during mouse infection. Where no arrow is present to connect a σ factor gene to a particular stress, this indicates that deletion of that σ factor gene did not significantly change survival during exposure to that stress or has not been studied under that particular condition. References are available in the text.

culosis during infection has come from cell culture and animal infection models. *sigE*, *sigF*, *sigG*, *sigH*, and *sigJ* are upregulated during infection of macrophages (47, 48), and both *sigE* and *sigG* are necessary for survival within macrophages (37, 49, 50). Deletion of *sigB*, *sigG*, *sigJ*, or *sigM* has no effect in animal models (6, 39, 40, 50). Deletion of *sigD*, *sigE*, *sigH*, or *sigL* results in a delayed time to death without affecting bacterial burden (51–54), while deletion of *sigC* or *sigF* results in a delayed time to death and a decrease in bacterial burden during acute (*sigC*) or chronic (*sigF*) infection (35, 44, 55). The importance of individual σ factors during infection and for survival under stressful conditions highlights both their central role in guiding *M. tuberculosis*'s stress response and the diverse adverse conditions encountered by *M. tuberculosis* during infection.

CarD AND RbpA—MAINTAINING THE PEACE, BUT READY TO DEFEND

The next branch of transcriptional regulation during stress responses involves RNAP-binding proteins that further modify gene expression from a given holoenzyme. CarD and RbpA are RNAP-binding proteins in *M. tuberculosis* that were each originally identified in experiments looking for genes upregulated in response to stress (56, 57). *carD* expression is upregulated in response to oxidative stress, starvation, and a broad panel of antibiotics. CarD activity is required for survival under the same conditions as well as for virulence in a mouse model of infection (56, 58, 59). *rbpA* is upregulated during oxidative stress, stationary phase, starvation, hypoxia, high temperatures, and treatment with antibiotics and during infection in macrophages (46, 57, 60–62). Overexpression of *rbpA* in mycobacteria also improves resistance to the antibiotic rifampin (63). CarD and RbpA both act by stabilizing the inherently unstable mycobacterial RNAP-promoter complexes, albeit by different mechanisms. While the presence of RbpA is limited to actinobacteria, CarD is present in members of numerous other bacterial phyla (56, 64, 65), including *Bacillus* and *Thermus*, where purified RNAPs also generate relatively unstable open promoter complexes (15–17), but not in *E. coli*, where RNAP generally forms stable open complexes (13, 14) (Fig. 5). *carD* and *rbpA* are essential in *M. tuberculosis* even during growth in nutrient-rich cultures (56, 66–68), indicating a general role in

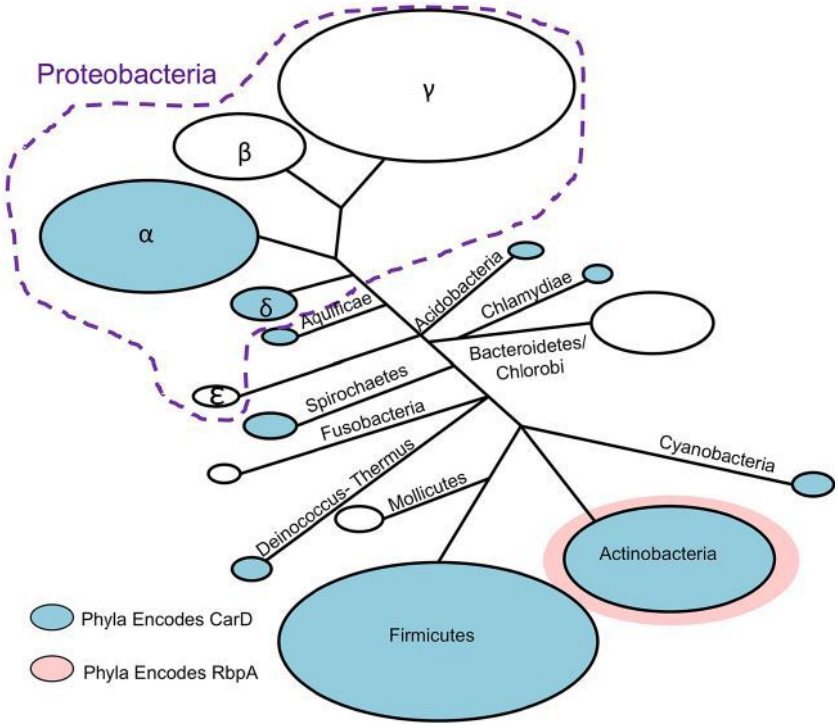


FIG 5 Phylogenetic distribution of CarD and RbpA. The BLAST database of completed genomes was searched for homologs of *M. tuberculosis* CarD and RbpA. Homologs of each protein were schematically drawn on a phylogenetic tree using a previously calculated phylogenetic distribution of bacteria based on the sequence conservation of RNAP subunits (185). Blue-shaded phyla have members that encode CarD homologs, members of the pink-shaded phylum (actinobacteria) encode RbpA, and phyla that are not shaded do not encode CarD or RbpA.

promoting efficient gene expression that also allows the RNAP to optimally respond to stress.

CarD interacts with the RNAP β-subunit β1-lobe through an N-terminal RNAP interaction domain (RID) and with DNA via a C-terminal basic patch (56, 58, 59, 65, 69, 70). In mycobacteria cultured under nutrient-rich conditions, CarD associates with RNAP-promoter complexes throughout the genome to enhance RP_o stability (14, 58, 59, 65). Using a bulk fluorescence assay to measure the effects of CarD on transcription initiation kinetics, it was shown that CarD associates with RP_o with high affinity and slows the rate of DNA closing by preventing bubble collapse and that CarD associates with RP_c with lower affinity and increases the rate of DNA opening (13). Importantly, the concentration of CarD in cells is sufficient for both of these activities to be physiologically relevant (13). These two activities of CarD change the kinetics of open complex formation such that the *M. bovis* RNAP more closely mirrors the *E. coli* RNAP (13, 14). The interactions between CarD and both DNA and RNAP are required for CarD activity (13). In addition, a conserved tryptophan within the C-terminal basic patch is also important for CarD's effects on RNAP-promoter complex stability and, based on structural studies, has been proposed to serve as a wedge at the upstream edge of the transcription bubble that prevents bubble collapse (59, 64, 65). Taken together, the inherently weak transcription initiation activity of *M. bovis* RNAP and CarD's global promoter localization suggest that CarD may be a general member of the mycobacterial transcription machinery.

RbpA consists of a central RbpA core domain (RCD) flanked by an unstructured 26-aa N-terminal tail and a C-terminal σ interaction domain (SID) linked to the RCD by a 15-aa basic linker (BL) (68, 71, 72). RbpA forms a stable binary complex with the σ2-domain of group 1 (σ^A in *M. tuberculosis*) and certain group 2 (σ^B in *M. tuberculosis*) σ factors through its SID (68, 71, 72), with additional contacts made between the N terminus and the σ factor (68). Based on structural modeling, the RbpA BL domain and adjacent residues interact with the DNA phosphate backbone of the nontemplate strand upstream of the −10 promoter element in the RP_o conformation (68). Additional contacts between RbpA and RNAP β have been proposed based on cross-linking experiments (63, 73, 74), but the recent structural modeling of RbpA onto an RNAP-promoter open complex would be incompatible with these interactions (71), suggesting that further analysis will be needed to resolve these inconsistencies. RbpA has been shown to increase the affinity of the σ factor to the core RNAP, increase the affinity of RNAP holoenzyme to promoter DNA, and facilitate the formation of RP_o (71, 75, 76), all of which could contribute to the ability of RbpA to promote RNAP-promoter complex formation and stability. The housekeeping σ factor σ^A has been reported to have an affinity for *M. tuberculosis* RNAP core enzyme similar to that of the alternative σ factor σ^F (74), in which case RbpA may be necessary to improve σ^A affinity and competitiveness for RNAP under conditions that require the activity of σ^A. In *E. coli*, in contrast, σ^{70} has a very high affinity to the RNAP core enzyme and thus can outcompete other σ factors under conditions where it is required without accessary factors such as RbpA. The RbpA SID and BL are important and sufficient to partially activate transcription *in vitro* (71, 72), but full activation of transcription requires the full-length protein, although the function of the N terminus of RbpA remains elusive.

Based on structural modeling performed with the information

currently available, association of CarD and RbpA with the same RNAP holoenzyme is feasible (71), but why *M. tuberculosis* requires both CarD and RbpA activities is unknown. CarD and RbpA transcriptional regulatory activities have thus far been analyzed only on limited promoters under limited conditions. However, their roles and effects at individual promoters likely depend on the kinetic properties of individual RNAP-promoter complexes and the presence of additional transcriptional regulators. The roles for CarD and RbpA during stress responses indicate that their effects on RP_o stability also provide a mechanism for adjusting gene expression during the switch between different physiological states in response to stress. Indeed, RP_o formation and stability comprise a commonly regulated step of transcription initiation during stress responses in bacteria, including during the stringent response (77, 78). It is possible that CarD and RbpA are important in stabilizing transcription complexes activated by stress-responsive transcription factors or alternative σ factors. While the functions of CarD and RbpA in stress responses remain unclear, the diversity of the stresses that they respond to suggests that they are acting at a common point shared among numerous stress responses.

ESSENTIAL TCSs AND TFs—ALWAYS ON THE LOOKOUT FOR HOSTILITY

M. tuberculosis encodes 12 complete two-component systems (TCSs), which are classically recognized as bacterial systems that sense and respond to stress and changes in the environment (79). Each TCS consists of at least one sensor histidine kinase (HK) that responds to specific environmental conditions by autophosphorylation and phosphotransfer to its cognate response regulator (RR), which then binds DNA and activates transcription of a specific regulon (79). Two TCSs in *M. tuberculosis*, MtrAB (80) and PrrAB (81), are essential for growth under unstressed culture conditions and have been integrated into the basic physiology of the bacteria. The HK MtrB colocalizes with cell division machinery at the bacterial septa and poles (82). Upon stimulation by an unknown signal, MtrB phosphorylates its cognate RR MtrA, which then binds DNA and activates transcription of a regulon that includes essential replication and cell division genes *dnaA* and *ripA* as well as the *fbpB* and *rpfB* genes that encode proteins with roles during infection (82–84). Integration of a TCS with the cell division machinery could allow these slowly replicating bacteria to sense environmental stress and abort cell division if unfavorable conditions surface. The second essential TCS in *M. tuberculosis* is the PrrAB system. The RR PrrA can bind DNA in the unphosphorylated state, but its binding affinity increases once phosphorylated by HK PrrB (85). The stimulus that results in activation of the PrrAB TCS has not been characterized, but expression of the *prrAB* operon is induced by nitrogen limitation and growth inside macrophages (81, 86), suggesting a possible role for this TCS under these conditions.

M. tuberculosis also encodes a series of essential iron-binding transcription factors (TF). *M. tuberculosis* does not contain functional homologues of the common redox-sensing TFs, FNR, SoxR, and OxyR, that allow other bacteria to sense and respond to redox state and reactive nitrogen and oxygen species (87–91). Instead, *M. tuberculosis* encodes a 7-member family of WhiB iron-sulfur (Fe-S) cluster TFs that sense the redox state in the cell and regulate gene expression accordingly (92). Of these, *whiB1* and *whiB2* are predicted to be essential, although their regulons have yet to be defined (93–95). *whiB1* is also upregulated during hypoxia and within infected mouse lungs (96, 97). WhiB2 may play a role in cell cycle progression, as a conditional *whiB2* mutant in *Mycobacterium smegmatis* was filamentous during depletion (95). The iron-binding TF IdeR is also essential for *M. tuberculosis* viability (98). IdeR dimerizes when bound to iron (99) and binds DNA as a dimer to inhibit transcription of genes involved in iron uptake and storage in order to promote adaptation to changing levels of iron (100, 101). By reducing levels of intracellular iron that can catalyze formation of reactive oxygen species, IdeR protects *M. tuberculosis* from oxidative and nitrosative stress and is important for survival in macrophages and mice (98, 100, 101). The essentiality of *whiB1*, *whiB2*, and *ideR* indicates a particular need for *M. tuberculosis* to couple redox sensing and iron availability with basic cellular processes to maintain homeostasis.

NONESSENTIAL TCSs AND TFs: SPECIAL FORCES OF THE STRESS RESPONSE TEAM

In addition to the essential TCSs and TFs mentioned above, *M. tuberculosis* maintains 10 nonessential TCSs and a number of nonessential TFs that are not required for bacterial growth *in vitro* but respond to particular stresses.

- The SenX3/RegX3 TCS is activated under low-phosphate conditions to regulate expression of genes encoding proteins involved in phosphate uptake, translation, lipid metabolism, DNA replication, and DNA repair (102, 103). The SenX3/RegX3 TCS is important for optimal *M. tuberculosis* growth during phosphate starvation and for survival in macrophages and mice where the bacteria encounter low phosphate levels (102).
- The DosRST system responds to nitric oxide and hypoxia to activate the "dormancy regulon" in *M. tuberculosis* (104). This TCS contains 2 separate HKs, DosS and DosT, that are both capable of activating the DosR RR. DosS acts as a redox sensor and DosT as a hypoxia sensor, illustrating the integration and differentiation of *M. tuberculosis* stress responses (105). Genetic disruption of the *dosRST* TCS results in reduced bacterial survival under low-oxygen conditions, in mouse models that develop hypoxic lesions, and in a nonhuman primate macaque model of infection (106–109).
- The PhoPR TCS is stimulated by low pH (110). The PhoP regulon includes multiple genes involved in cellular lipid synthesis, *dosR*, *dosS*, and genes involved in the ESX1 secretion system (111, 112). *M. tuberculosis* strains deficient in PhoPR activity display defects in replication in mice and macrophages (111–113). Supporting the idea of a role in *M. tuberculosis* virulence, mutations in *phoPR* in *M. bovis* and *Mycobacterium africanum* are associated with reduced mycobacterial virulence (114). In addition, *M. tuberculosis phoPR* mutants have defects in cell morphology and lipid production in the absence of stress, suggesting that PhoPR is required to maintain normal cell physiology under all growth conditions (113).
- The MprAB TCS regulates expression of a subset of genes in the DosR regulon, the stress-responsive chaperone *pepD*, and the *espA* operon, which encodes ESX-1 substrates (115–118). The MprAB TCS also activates expression of *sigB* and *sigE* in response to envelope stress and indirectly regulates

the stringent response mediator *M. tuberculosis rel* gene (rel_{Mtb}) through σ^E activity (119, 120). Deletion of this TCS compromises *M. tuberculosis* viability during a persistent infection in mice but renders *M. tuberculosis* hypervirulent in macrophages, suggesting a role for this TCS in allowing the bacteria to appropriately respond to their specific *in vivo* niche (80, 121).

- Genes encoding six additional TCSs, KdpDE, TrcRS, TcrXY, NarLS, PtdaRS, and Rv0600c/Rv0601c/TcrA, have been identified in the *M. tuberculosis* genome but have yet to be investigated in detail (79).
- *M. tuberculosis* encodes five nonessential Fe-S cluster WhiB TF family members that have been implicated in a variety of cellular responses (96, 97). In particular, WhiB3, WhiB4, and WhiB5 impact *M. tuberculosis* virulence (122–124). Of these, WhiB3 has been studied in the most detail. WhiB3 promotes mycobacterial lipid regulation, and *whiB3* mutants demonstrate altered macrophage cytokine release and reduced pathology *in vivo*, without directly impacting bacterial titers (125, 126). A model has been proposed in which WhiB3 senses the intracellular redox state and redirects lipid synthesis pathways to cope with reductive stress generated by host lipid catabolism during infection (125).
- *M. tuberculosis* encodes a number of other known and predicted TFs not highlighted in this review. Recently, researchers overexpressed 200 predicted TFs in *M. tuberculosis* and performed chromatin immunoprecipitation sequencing experiments and microarray analyses to catalogue a genome-wide characterization of TF binding events and target gene expression (127–129). These reports describe 16,000 binding sites for 154 TFs and identify regulatory routes for ~70% of the genome. The complex regulatory circuits that were uncovered highlight how much remains to be investigated regarding how *M. tuberculosis* regulates transcription to integrate precise stress responses.

THE STRINGENT RESPONSE: WHEN RATIONS RUN LOW

The stringent response is a conserved global stress response in bacteria that provides an additional layer of gene regulation in harsh environments. The stringent response is best characterized during amino acid starvation, when the Rel_{Mtb} enzyme senses uncharged tRNAs in ribosomes and responds by transferring the pyrophosphate (PPi) group from ATP to GDP and GTP to synthesize hyperphosphorylated guanine nucleotides ppGpp and pppGpp [collectively called (p)ppGpp] (130). (p)ppGpp then coordinates downstream regulation of bacterial physiology and mediates changes in the transcriptional profile to support survival during stress. Deletion of rel_{Mtb} led to differential expression of 159 genes during starvation, including genes involved in coordinating metabolic rate reduction, production of mycobacterial cell wall and lipids, secreted proteins, and cell division machinery (131). (p)ppGpp synthesis by Rel_{Mtb} is required for survival under low-nutrient conditions, in long-term culture, and during infection in animal models, all indicative of a strict requirement for Rel_{Mtb} during exposure to stress (131–135). In *E. coli*, (p)ppGpp directly affects transcription initiation by binding the RNAP (136, 137). In contrast, in a number of Gram-positive bacteria, (p)ppGpp inhibits GTP biosynthesis by directly interacting with GTP synthesis enzymes, which impacts gene expression by altering initiating nucleotide levels (137–140). Although (p)ppGpp has not been demonstrated to directly bind *M. tuberculosis* RNAP or GTP synthesis enzymes, (p)ppGpp has been reported to influence mycobacterial RNAP activity *in vitro*, suggesting that the mechanism of (p)ppGpp action in *M. tuberculosis* transcriptional modulation requires further investigation (136, 138, 141).

Rel_{Mtb} also encodes a second distinct catalytic domain that hydrolyzes (p)ppGpp into PPi and GDP or GTP (142). It was recently shown that (p)ppGpp hydrolysis by Rel_{Mtb} is important for growth and normal physiology in culture and during infection (135). These observations suggest that Rel_{Mtb} constitutively produces (p)ppGpp independently of activation during nutrient limitation and may act continuously to maintain *M. tuberculosis* homeostasis under all growth conditions in addition to its role in survival during stress.

FINAL THOUGHTS

In order to respond to host-derived stresses, *M. tuberculosis* has evolved a complex network of strategies to modify gene expression and promote survival. The responses to different stresses are integrated and coordinated, often resulting in overlapping regulons and stress responders (Fig. 2, 3, and 4). Not only do these highly effective stress response strategies protect *M. tuberculosis* from host immunity, but the resulting changes in physiology also contribute to antibiotic tolerance, which precludes eradication of the infection (143–148). The recalcitrance of *M. tuberculosis* in response to antibiotic therapy has led to an increase in drug-resistant *M. tuberculosis* infections to the point that we are not equipped to successfully battle the *M. tuberculosis* epidemic (2). Therefore, new therapeutic strategies that target *M. tuberculosis* stress responses could increase the susceptibility of the bacteria to both the immune system and antibiotic treatment.

As an obligate pathogen, *M. tuberculosis* is specialized for survival in a mammalian host. Analysis of the conservation of transcriptional regulators across different mycobacterial species reveals some interesting patterns that reflect their respective lifestyles (Fig. 2). *Mycobacterium leprae* is an even more specialized pathogen than *M. tuberculosis* and has undergone a drastic reduction in genetic material to the point that this degenerate genome has retained only 4 functional σ factor genes (*sigA*, *sigB*, *sigC*, and *sigE*) and 5 TCSs. On the other end of the spectrum, environmental mycobacteria such as *Mycobacterium smegmatis* must adapt to a larger diversity of conditions within a larger range of environments. As such, *M. smegmatis* encodes 28 σ factors to facilitate a more versatile lifestyle. In addition, even when a transcriptional regulator is conserved across mycobacterial species, it can be coopted to perform a function specific for a particular species. For example, σ^F homologs are differentially regulated and activated in *M. tuberculosis*, *M. smegmatis*, and *M. bovis* (7, 29, 149).

Finally, this minireview is in no way exhaustive in terms of all of the mechanisms of transcriptional regulation that *M. tuberculosis* employs to respond to stress. In particular, there is a growing area of research into the roles of nucleoid-associated proteins and small RNAs (150–155). *M. tuberculosis* also contains 11 serine/threonine protein kinases (STPKs) that, like TCSs, are involved in signal transduction pathways that aid *M. tuberculosis* in adaptation to its environment (156). However, unlike TCSs that consist of HKs that activate RRs to directly modulate *M. tuberculosis* transcription, STPKs are single proteins that phosphorylate numerous

downstream targets (156). Although STPKs do not directly affect *M. tuberculosis* transcription, they do influence gene expression by modifying the activity of other *M. tuberculosis* proteins with more-direct roles in transcription, such as σ factors, nucleoid-associated proteins, anti-anti-σ factors, and TCSs (24, 154, 157–159). These and other aspects of gene regulation further add to the complexity of stress responses in *M. tuberculosis*.

ACKNOWLEDGMENTS

We thank Michael Caparon, Katherine Mann, and Jerome Prusa for careful reading of the manuscript and providing their insightful comments.

C.L.S. is supported by a Beckman Young Investigator Award from the Arnold and Mabel Beckman Foundation, an Interdisciplinary Research Initiative grant from the Children's Discovery Institute of Washington University and St. Louis Children's Hospital, and grants GM107544 and AI111696-01 from the National Institutes of Health. K.F. is supported by a pilot award from the Center for Women's Infectious Disease Research at Washington University School of Medicine. A.L.G. is supported by a National Institute of General Medical Sciences (NIGMS) Cell and Molecular Biology Training Grant (grant GM007067) and the Stephen I. Morse Graduate Fellowship.

FUNDING INFORMATION

This work, including the efforts of Christina L. Stallings, was funded by HHS | National Institutes of Health (NIH) (GM107544 and AI111696-01). This work, including the efforts of Ashley L. Garner, was funded by HHS | National Institutes of Health (NIH) (GM007067). This work, including the efforts of Christina L. Stallings, was funded by Arnold and Mabel Beckman Foundation (BYI).

REFERENCES

1. **Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, Roach JC, Kremer K, Petrov DA, Feldman MW, Gagneux S.** 2008. High functional diversity in Mycobacterium tuberculosis driven by genetic drift and human demography. PLoS Biol **6:**e311. http://dx.doi.org/10.1371/journal.pbio.0060311.
2. **WHO.** 2015. Global tuberculosis report 2015. WHO, Geneva, Switzerland.
3. **Stallings CL, Glickman MS.** 2010. Is Mycobacterium tuberculosis stressed out? A critical assessment of the genetic evidence. Microbes Infect **12:**1091–1101.
4. **Rodrigue S, Provvedi R, Jacques PÉ Gaudreau L, Manganelli R.** 2006. The σ factors of Mycobacterium tuberculosis. FEMS Microbiol Rev **30:**926–941. http://dx.doi.org/10.1111/j.1574-6976.2006.00040.x.
5. **Hu Y, Coates ARM.** 2001. Increased levels of sigJ mRNA in late stationary phase cultures of Mycobacterium tuberculosis detected by DNA array hybridisation. FEMS Microbiol Lett **202:**59–65. http://dx.doi.org/10.1111/j.1574-6968.2001.tb10780.x.
6. **Manganelli R, Provvedi R, Rodrigue S, Beaucher J, Gaudreau L, Smith I.** 2004. Sigma factors and global gene regulation in Mycobacterium tuberculosis. J Bacteriol **186:**895–902. http://dx.doi.org/10.1128/JB.186.4.895-902.2004.
7. **Manganelli R, Dubnau E, Tyagi S, Kramer FR, Smith I.** 1999. Differential expression of 10 sigma factor genes in Mycobacterium tuberculosis. Mol Microbiol **31:**715–724. http://dx.doi.org/10.1046/j.1365-2958.1999.01212.x.
8. **Murakami KS, Darst SA.** 2003. Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol **13:**31–39. http://dx.doi.org/10.1016/S0959-440X(02)00005-2.
9. **Saecker RM, Record MT, Dehaseth PL.** 2011. Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. J Mol Biol **412:**754–771. http://dx.doi.org/10.1016/j.jmb.2011.01.018.
10. **Browning DF, Busby SJ.** 2004. The regulation of bacterial transcription initiation. Nat Rev Microbiol **2:**57–65. http://dx.doi.org/10.1038/nrmicro787.
11. **Rojo F.** 2001. Mechanisms of transcriptional repression. Curr Opin Microbiol **4:**145–151. http://dx.doi.org/10.1016/S1369-5274(00)00180-6.
12. **Lee DJ, Minchin SD, Busby SJW.** 2012. Activating transcription in bacteria. Annu Rev Microbiol **66:**125–152. http://dx.doi.org/10.1146/annurev-micro-092611-150012.
13. **Rammohan J, Ruiz Manzano A, Garner AL, Stallings CL, Galburt EA.** 2015. CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. Nucleic Acids Res **43:**3272–3285. http://dx.doi.org/10.1093/nar/gkv078.
14. **Davis E, Chen J, Leon K, Darst SA, Campbell EA.** 2015. Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. Nucleic Acids Res **43:**433–445. http://dx.doi.org/10.1093/nar/gku1231.
15. **Whipple FW, Sonenshein AL.** 1992. Mechanism of initiation of transcription by Bacillus subtilis RNA polymerase at several promoters. J Mol Biol **223:**399–414. http://dx.doi.org/10.1016/0022-2836(92)90660-C.
16. **Xue Y, Hogan BP, Erie DA.** 2000. Purification and initial characterization of RNA polymerase from Thermus thermophilus strain HB8. Biochemistry **39:**14356–14362. http://dx.doi.org/10.1021/bi0012538.
17. **Miropolskaya N, Ignatov A, Bass I, Zhilina E, Pupov D, Kulbachinskiy A.** 2012. Distinct functions of regions 1.1 and 1.2 of RNA polymerase subunits from Escherichia coli and Thermus aquaticus in transcription initiation. J Biol Chem **287:**23779–23789. http://dx.doi.org/10.1074/jbc.M112.363242.
18. **Wösten M.** 1998. Eubacterial sigma-factors. FEMS Microbiol Rev **22:**127–150.
19. **Newton-Foot M, Gey van Pittius NC.** 2013. The complex architecture of mycobacterial promoters. Tuberculosis **93:**60–74. http://dx.doi.org/10.1016/j.tube.2012.08.003.
20. **Bashyam MD, Kaushal D, Dasgupta SK, Tyagi AK.** 1996. A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements. J Bacteriol **178:**4847–4853.
21. **Agarwal N, Tyagi AK.** 2006. Mycobacterial transcriptional signals: requirements for recognition by RNA polymerase and optimal transcriptional activity. Nucleic Acids Res **34:**4245–4257. http://dx.doi.org/10.1093/nar/gkl521.
22. **Kremer L, Baulard A, Estaquier J, Content J, Capron A, Locht C.** 1995. Analysis of the Mycobacterium tuberculosis 85A antigen promoter region. J Bacteriol **177:**642–653.
23. **Bashyam MD, Tyagi AK.** 1998. Identification and analysis of "extended −10" promoters from mycobacteria. J Bacteriol **180:**2568–2573.
24. **Sachdeva P, Misra R, Tyagi AK, Singh Y.** 2010. The sigma factors of Mycobacterium tuberculosis: regulation of the regulators. FEBS J **277:**605–626. http://dx.doi.org/10.1111/j.1742-4658.2009.07479.x.
25. **Song T, Dove SL, Lee KH, Husson RN.** 2003. RshA, an anti-sigma factor that regulates the activity of the mycobacterial stress response sigma factor SigH. Mol Microbiol **50:**949–959. http://dx.doi.org/10.1046/j.1365-2958.2003.03739.x.
26. **Hahn M, Raman S, Anaya M, Husson RN.** 2005. The Mycobacterium tuberculosis extracytoplasmic-function sigma factor SigL regulates polyketide synthases and secreted or membrane proteins and is required for virulence. J Bacteriol **187:**7062–7071. http://dx.doi.org/10.1128/JB.187.20.7062-7071.2005.
27. **Saïd-Salim B, Mostowy S, Kristof AS, Behr MA.** 2006. Mutations in Mycobacterium tuberculosis Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in Mycobacterium bovis. Mol Microbiol **62:**1251–1263. http://dx.doi.org/10.1111/j.1365-2958.2006.05455.x.
28. **DeMaio J, Zhang Y, Ko C, Bishai WR.** 1997. Mycobacterium tuberculosis sigF is part of a gene cluster with similarities to the Bacillus subtilis sigF and sigB operons. Tuber Lung Dis **78:**3–12. http://dx.doi.org/10.1016/S0962-8479(97)90010-1.
29. **Singh AK, Singh BN.** 2008. Conservation of sigma F in mycobacteria and its expression in Mycobacterium smegmatis. Curr Microbiol **56:**574–580. http://dx.doi.org/10.1007/s00284-008-9126-8.
30. **Donà V, Rodrigue S, Dainese E, Palù G, Gaudreau L, Manganelli R, Provvedi R.** 2008. Evidence of complex transcriptional, translational, and posttranslational regulation of the extracytoplasmic function sigma factor σE in Mycobacterium tuberculosis. J Bacteriol **190:**5963–5971. http://dx.doi.org/10.1128/JB.00622-08.
31. **Schneider JS, Sklar JG, Glickman MS.** 2014. The Rip1 protease of mycobacterium tuberculosis controls the SigD regulon. J Bacteriol **196:**2638–2645. http://dx.doi.org/10.1128/JB.01537-14.

32. **Sklar JG, Makinoshima H, Schneider JS, Glickman MS.** 2010. M. tuberculosis intramembrane protease Rip1 controls transcription through three anti-sigma factor substrates. Mol Microbiol **77:**605–617. http://dx.doi.org/10.1111/j.1365-2958.2010.07232.x.
33. **Gaudion AE.** 2011. The role of the ECF sigma factor SigG in Mycobacterium tuberculosis. Ph.D. thesis. MRC National Institute for Medical Research, London, United Kingdom.
34. **Hu Y, Coates ARM.** 1999. Transcription of two sigma 70 homologue genes, sigA and sigB, in stationary-phase Mycobacterium tuberculosis. J Bacteriol **181:**469–476.
35. **Chen P, Ruiz RE, Li Q, Silver RF, Bishai WR.** 2000. Construction and characterization of a Mycobacterium tuberculosis mutant lacking the alternate sigma factor gene, sigF. Infect Immun **68:**5575–5580. http://dx.doi.org/10.1128/IAI.68.10.5575-5580.2000.
36. **Raman S, Song T, Puyang X, Jacobs WRJ, Husson RN.** 2001. The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in Mycobacterium tuberculosis. J Bacteriol **183:**6119–6125. http://dx.doi.org/10.1128/JB.183.20.6119-6125.2001.
37. **Manganelli R, Voskuil MI, Schoolnik GK, Smith I.** 2001. The Mycobacterium tuberculosis ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. Mol Microbiol **41:**423–437. http://dx.doi.org/10.1046/j.1365-2958.2001.02525.x.
38. **Manganelli R, Voskuil MI, Schoolnik GK, Dubnau E, Gomez M, Smith I.** 2002. Role of the extracytoplasmic-function sigma factor sigma(H) in Mycobacterium tuberculosis global gene expression. Mol Microbiol **45:**365–374. http://dx.doi.org/10.1046/j.1365-2958.2002.03005.x.
39. **Hu Y, Kendall S, Stoker NG, Coates ARM.** 2004. The Mycobacterium tuberculosis sigJ gene controls sensitivity of the bacterium to hydrogen peroxide. FEMS Microbiol Lett **237:**415–423. http://dx.doi.org/10.1111/j.1574-6968.2004.tb09725.x.
40. **Raman S, Puyang X, Cheng TY, Young DC, Moody DB, Husson RN.** 2006. Mycobacterium tuberculosis SigM positively regulates Esx secreted protein and nonribosomal peptide synthetase genes and down regulates virulence-associated surface lipid synthesis. J Bacteriol **188:**8460–8468. http://dx.doi.org/10.1128/JB.01212-06.
41. **Agarwal N, Woolwine SC, Tyagi S, Bishai WR.** 2007. Characterization of the Mycobacterium tuberculosis sigma factor SigM by assessment of virulence and identification of SigM-dependent genes. Infect Immun **75:**452–461. http://dx.doi.org/10.1128/IAI.01395-06.
42. **Fontán PA, Voskuil MI, Gomez M, Tan D, Pardini M, Manganelli R, Fattorini L, Schoolnik GK, Smith I.** 2009. The Mycobacterium tuberculosis sigma factor B is required for full response to cell envelope stress and hypoxia in vitro, but it is dispensable for in vivo growth. J Bacteriol **191:**5628–5633. http://dx.doi.org/10.1128/JB.00510-09.
43. **Dutta NK, Mehra S, Kaushal D.** 2010. A Mycobacterium tuberculosis sigma factor network responds to cell-envelope damage by the promising anti-mycobacterial thioridazine. PLoS One **5:**e10069. http://dx.doi.org/10.1371/journal.pone.0010069.
44. **Geiman D, Kaushal D, Ko C.** 2004. Attenuation of late-stage disease in mice infected by the Mycobacterium tuberculosis mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis. Infect Immun **72:**1733–1745. http://dx.doi.org/10.1128/IAI.72.3.1733-1745.2004.
45. **Smollett KL, Dawson LF, Davis EO.** 2011. SigG does not control gene expression in response to DNA damage in Mycobacterium tuberculosis H37Rv. J Bacteriol **193:**1007–1011. http://dx.doi.org/10.1128/JB.01241-10.
46. **Betts JC, Lukey PT, Robb LC, McAdam RA, Duncan K.** 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. Mol Microbiol **43:**717–731. http://dx.doi.org/10.1046/j.1365-2958.2002.02779.x.
47. **Graham JE, Clark-Curtiss JE.** 1999. Identification of Mycobacterium tuberculosis RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). Proc Natl Acad Sci U S A **96:**11554–11559. http://dx.doi.org/10.1073/pnas.96.20.11554.
48. **Cappelli G, Volpe E, Grassi M, Liseo B, Colizzi V, Mariani F.** 2006. Profiling of Mycobacterium tuberculosis gene expression during human macrophage infection: upregulation of the alternative sigma factor G, a group of transcriptional regulators, and proteins with unknown function. Res Microbiol **157:**445–455. http://dx.doi.org/10.1016/j.resmic.2005.10.007.
49. **Lee J-H, Geiman DE, Bishai WR.** 2008. Role of stress response sigma factor SigG in Mycobacterium tuberculosis. J Bacteriol **190:**1128–1133. http://dx.doi.org/10.1128/JB.00511-07.
50. **Gaudion A, Dawson L, Davis E, Smollett K.** 2013. Characterisation of the mycobacterium tuberculosis alternative sigma factor SigG: its operon and regulon. Tuberculosis **93:**482–491. http://dx.doi.org/10.1016/j.tube.2013.05.005.
51. **Calamita H, Ko C, Tyagi S, Yoshimatsu T, Morrison NE, Bishai WR.** 2005. The Mycobacterium tuberculosis SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence. Cell Microbiol **7:**233–244.
52. **Ando M, Yoshimatsu T, Ko C, Converse PJ, Bishai WR.** 2003. Deletion of Mycobacterium tuberculosis sigma factor E results in delayed time to death with bacterial persistence in the lungs of aerosol-infected mice. Infect Immun **71:**7170–7172. http://dx.doi.org/10.1128/IAI.71.12.7170-7172.2003.
53. **Kaushal D, Schroeder BG, Tyagi S, Yoshimatsu T, Scott C, Ko C, Carpenter L, Mehrotra J, Manabe YC, Fleischmann RD, Bishai WR.** 2002. Reduced immunopathology and mortality despite tissue persistence in a Mycobacterium tuberculosis mutant lacking alternative sigma factor, SigH. Proc Natl Acad Sci U S A **99:**8330–8335. http://dx.doi.org/10.1073/pnas.102055799.
54. **Dainese E, Rodrigue S, Delogu G, Provvedi R, Laflamme L, Brzezinski R, Fadda G, Smith I, Gaudreau L, Palù G, Manganelli R.** 2006. Posttranslational regulation of Mycobacterium tuberculosis extracytoplasmic-function sigma factor sigma L and roles in virulence and in global regulation of gene expression. Infect Immun **74:**2457–2461. http://dx.doi.org/10.1128/IAI.74.4.2457-2461.2006.
55. **Sun R, Converse PJ, Ko C, Tyagi S, Morrison NE, Bishai WR.** 2004. Mycobacterium tuberculosis ECF sigma factor sigC is required for lethality in mice and for the conditional expression of a defined gene set. Mol Microbiol **52:**25–38. http://dx.doi.org/10.1111/j.1365-2958.2003.03958.x.
56. **Stallings CL, Stephanou NC, Chu L, Hochschild A, Nickels BE, Glickman MS.** 2009. CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. Cell **138:**146–159. http://dx.doi.org/10.1016/j.cell.2009.04.041.
57. **Paget MS, Molle V, Cohen G, Aharonowitz Y, Buttner MJ.** 2001. Defining the disulphide stress response in Streptomyces coelicolor A3(2): identification of the sigmaR regulon. Mol Microbiol **42:**1007–1020. http://dx.doi.org/10.1046/j.1365-2958.2001.02675.x.
58. **Weiss LA, Harrison PG, Nickels BE, Glickman MS, Campbell EA, Darst SA, Stallings CL.** 2012. Interaction of CarD with RNA polymerase mediates Mycobacterium tuberculosis viability, rifampin resistance, and pathogenesis. J Bacteriol **194:**5621–5631. http://dx.doi.org/10.1128/JB.00879-12.
59. **Garner AL, Weiss LA, Ruiz Manzano A, Galburt EA, Stallings CL.** 2014. CarD integrates three functional modules to promote efficient transcription, antibiotic tolerance, and pathogenesis in mycobacteria. Mol Microbiol **93:**682–697. http://dx.doi.org/10.1111/mmi.12681.
60. **Stewart GR, Wernisch L, Stabler R, Mangan JA, Hinds J, Laing KG, Young DB, Butcher PD.** 2002. Dissection of the heat-shock response in Mycobacterium tuberculosis using mutants and microarrays. Microbiology **148:**3129–3138. http://dx.doi.org/10.1099/00221287-148-10-3129.
61. **Provvedi R, Boldrin F, Falciani F, Palù G, Manganelli R.** 2009. Global transcriptional response to vancomycin in Mycobacterium tuberculosis. Microbiology **155:**1093–1102. http://dx.doi.org/10.1099/mic.0.024802-0.
62. **Murphy DJ, Brown JR.** 2007. Identification of gene targets against dormant phase Mycobacterium tuberculosis infections. BMC Infect Dis **7:**84. http://dx.doi.org/10.1186/1471-2334-7-84.
63. **Dey A, Verma AK, Chatterji D.** 2010. Role of an RNA polymerase interacting protein, MsRbpA, from Mycobacterium smegmatis in phenotypic tolerance to rifampicin. Microbiology **156:**873–883. http://dx.doi.org/10.1099/mic.0.033670-0.
64. **Bae B, Chen J, Davis E, Leon K, Darst SA, Campbell EA.** 2015. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. Elife **4:**e08505. http://dx.doi.org/10.7554/eLife.08505.
65. **Srivastava DB, Leon K, Osmundson J, Garner AL, Weiss LA, Westblade LF, Glickman MS, Landick R, Darst SA, Stallings CL, Campbell EA.** 2013. Structure and function of CarD, an essential mycobacterial

transcription factor. Proc Natl Acad Sci U S A **110:**12619–12624. http://dx.doi.org/10.1073/pnas.1308270110.
66. **Forti F, Mauri V, Dehò G, Ghisotti D.** 2011. Isolation of conditional expression mutants in Mycobacterium tuberculosis by transposon mutagenesis. Tuberculosis (Edinb) **91:**569–578. http://dx.doi.org/10.1016/j.tube.2011.07.004.
67. **Sassetti CM, Boyd DH, Rubin EJ.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77–84. http://dx.doi.org/10.1046/j.1365-2958.2003.03425.x.
68. **Bortoluzzi A, Muskett FW, Waters LC, Addis PW, Rieck B, Munder T, Schleier S, Forti F, Ghisotti D, Carr MD, O'Hare HM.** 2013. Mycobacterium tuberculosis RNA polymerase-binding protein A (RbpA) and its interactions with sigma factors. J Biol Chem **288:**14438–14450. http://dx.doi.org/10.1074/jbc.M113.459883.
69. **García-Moreno D, Abellón-Ruiz J, García-Heras F, Murillo FJ, Padmanabhan S, Elías-Arnanz M.** 2010. CdnL, a member of the large CarD-like family of bacterial proteins, is vital for Myxococcus xanthus and differs functionally from the global transcriptional regulator CarD. Nucleic Acids Res **38:**4586–4598. http://dx.doi.org/10.1093/nar/gkq214.
70. **Gulten G, Sacchettini JC.** 2013. Structure of the Mtb CarD/RNAP β-lobes complex reveals the molecular basis of interaction and presents a distinct DNA-binding domain for Mtb CarD. Structure **21:**1859–1869. http://dx.doi.org/10.1016/j.str.2013.08.014.
71. **Hubin EA, Tabib-Salazar A, Humphrey LJ, Flack JE, Olinares PDB, Darst SA, Campbell EA, Paget MS.** 2015. Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. Proc Natl Acad Sci U S A **112:**7171–7176. http://dx.doi.org/10.1073/pnas.1504942112.
72. **Tabib-Salazar A, Liu B, Doughty P, Lewis RA, Ghosh S, Parsy M-L, Simpson PJ, O'Dwyer K, Matthews SJ, Paget MS.** 2013. The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase. Nucleic Acids Res **41:**5679–5691. http://dx.doi.org/10.1093/nar/gkt277.
73. **Dey A, Verma AK, Chatterji D.** 2011. Molecular insights into the mechanism of phenotypic tolerance to rifampicin conferred on mycobacterial RNA polymerase by MsRbpA. Microbiology **157:**2056–2071. http://dx.doi.org/10.1099/mic.0.047480-0.
74. **Hu Y, Morichaud Z, Chen S, Leonetti J-P, Brodolin K.** 2012. Mycobacterium tuberculosis RbpA protein is a new type of transcriptional activator that stabilizes the σ A-containing RNA polymerase holoenzyme. Nucleic Acids Res **40:**6547–6557. http://dx.doi.org/10.1093/nar/gks346.
75. **Hu Y, Morichaud Z, Perumal AS, Roquet-Baneres F, Brodolin K.** 2014. Mycobacterium RbpA cooperates with the stress-response σB subunit of RNA polymerase in promoter DNA unwinding. Nucleic Acids Res **42:**10399–10408. http://dx.doi.org/10.1093/nar/gku742.
76. **Verma AK, Chatterji D.** 2014. Dual role of MsRbpA: transcription activation and rescue of transcription from the inhibitory effect of rifampicin. Microbiology **160:**2018–2029. http://dx.doi.org/10.1099/mic.0.079186-0.
77. **Potrykus K, Cashel M.** 2008. (p)ppGpp: still magical? Annu Rev Microbiol **62:**35–51. http://dx.doi.org/10.1146/annurev.micro.62.081307.162903.
78. **Hauryliuk V, Atkinson GC, Murakami KS, Tenson T, Gerdes K.** 2015. Recent functional insights into the role of (p)ppGpp in bacterial physiology. Nat Rev Microbiol **13:**298–309. http://dx.doi.org/10.1038/nrmicro3448.
79. **Parish T.** 2014. Two-component regulatory systems of Mycobacteria, p 209–223. *In* Hatfull GF, Jacobs WR (ed), Molecular genetics of Mycobacteria, 2nd ed. ASM Press, Washington, DC. http://dx.doi.org/10.1128/microbiolspec.MGM2-0010-2013.
80. **Zahrt TC, Deretic V.** 2000. An essential two-component signal transduction system in Mycobacterium tuberculosis. J Bacteriol **182:**3832–3838. http://dx.doi.org/10.1128/JB.182.13.3832-3838.2000.
81. **Haydel SE, Malhotra V, Cornelison GL, Clark-Curtiss JE.** 2012. The prrAB two-component system is essential for Mycobacterium tuberculosis viability and is induced under nitrogen-limiting conditions. J Bacteriol **194:**354–361. http://dx.doi.org/10.1128/JB.06258-11.
82. **Plocinska R, Purushotham G, Sarva K, Vadrevu IS, Pandeeti EVP, Arora N, Plocinski P, Madiraju MV, Rajagopalan M.** 2012. Septal localization of the Mycobacterium tuberculosis MtrB sensor kinase promotes MtrA regulon expression. J Biol Chem **287:**23887–23899. http://dx.doi.org/10.1074/jbc.M112.346544.
83. **Rajagopalan M, Dziedzic R, Al Zayer M, Stankowska D, Ouimet M-C, Bastedo DP, Marczynski GT, Madiraju MV.** 2010. Mycobacterium tuberculosis origin of replication and the promoter for immunodominant secreted antigen 85B are the targets of MtrA, the essential response regulator. J Biol Chem **285:**15816–15827. http://dx.doi.org/10.1074/jbc.M109.040097.
84. **Sharma AK, Chatterjee A, Gupta S, Banerjee R, Mandal S, Mukhopadhyay J, Basu J, Kundu M.** 2015. MtrA, an essential response regulator of the MtrAB two-component system, regulates the transcription of resuscitation-promoting factor B of Mycobacterium tuberculosis. Microbiology **161:**1271–1281. http://dx.doi.org/10.1099/mic.0.000087.
85. **Ewann F, Locht C, Supply P.** 2004. Intracellular autoregulation of the Mycobacterium tuberculosis PrrA response regulator. Microbiology **150:**241–246. http://dx.doi.org/10.1099/mic.0.26516-0.
86. **Ewann F.** 2002. Transient requirement of the PrrA-PrrB two-component system for early intracellular multiplication of Mycobacterium tuberculosis. Infect Immun **70:**2256–2263. http://dx.doi.org/10.1128/IAI.70.5.2256-2263.2002.
87. **Crack JC, Green J, Cheesman MR, Le Brun NE, Thomson AJ.** 2007. Superoxide-mediated amplification of the oxygen-induced switch from [4Fe-4S] to [2Fe-2S] clusters in the transcriptional regulator FNR. Proc Natl Acad Sci U S A **104:**2092–2097. http://dx.doi.org/10.1073/pnas.0609514104.
88. **Gu M, Imlay JA.** 2011. The SoxRS response of Escherichia coli is directly activated by redox-cycling drugs rather than by superoxide. Mol Microbiol **79:**1136–1150. http://dx.doi.org/10.1111/j.1365-2958.2010.07520.x.
89. **Hausladen A, Privalle CT, Keng T, DeAngelo J, Stamler JS.** 1996. Nitrosative stress: activation of the transcription factor OxyR. Cell **86:**719–729. http://dx.doi.org/10.1016/S0092-8674(00)80147-6.
90. **Poole RK, Anjum MF, Membrillo-Hernández J, Kim SO, Hughes MN, Stewart V.** 1996. Nitric oxide, nitrite, and Fnr regulation of hmp (flavohemoglobin) gene expression in Escherichia coli K-12. J Bacteriol **178:**5487–5492.
91. **Wu J, Weiss B.** 1992. Two-stage induction of the soxRS (superoxide response) regulon of Escherichia coli. J Bacteriol **174:**3915–3920.
92. **Saini V, Farhana A, Glasgow JN, Steyn AJC.** 2012. Iron sulfur cluster proteins and microbial regulation: implications for understanding tuberculosis. Curr Opin Chem Biol **16:**45–53. http://dx.doi.org/10.1016/j.cbpa.2012.03.004.
93. **Smith LJ, Stapleton MR, Fullstone GJM, Crack JC, Thomson AJ, Le Brun NE, Hunt DM, Harvey E, Adinolfi S, Buxton RS, Green J.** 2010. Mycobacterium tuberculosis WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron-sulfur cluster. Biochem J **432:**417–427. http://dx.doi.org/10.1042/BJ20101440.
94. **Raghunand TR, Bishai WR.** 2006. Mycobacterium smegmatis whmD and its homologue Mycobacterium tuberculosis whiB2 are functionally equivalent. Microbiology **152:**2735–2747. http://dx.doi.org/10.1099/mic.0.28911-0.
95. **Gomez JE, Bishai WR.** 2000. whmD is an essential mycobacterial gene required for proper septation and cell division. Proc Natl Acad Sci U S A **97:**8554–8559. http://dx.doi.org/10.1073/pnas.140225297.
96. **Larsson C, Luna B, Ammerman NC, Maiga M, Agarwal N, Bishai WR.** 2012. Gene expression of Mycobacterium tuberculosis putative transcription factors whiB1–7 in redox environments. PLoS One **7:**e37516. http://dx.doi.org/10.1371/journal.pone.0037516.
97. **Geiman DE, Raghunand TR, Agarwal N, Bishai WR.** 2006. Differential gene expression in response to exposure to antimycobacterial agents and other stress conditions among seven Mycobacterium tuberculosis whiB-like genes. Antimicrob Agents Chemother **50:**2836–2841. http://dx.doi.org/10.1128/AAC.00295-06.
98. **Rodriguez GM, Voskuil MI, Gold B, Schoolnik GK, Smith I.** 2002. ideR, an essential gene in Mycobacterium tuberculosis: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. Infect Immun **70:**3371–3381. http://dx.doi.org/10.1128/IAI.70.7.3371-3381.2002.
99. **Wisedchaisri G, Holmes RK, Hol WGJ.** 2004. Crystal structure of an IdeR-DNA complex reveals a conformational change in activated IdeR for base-specific interactions. J Mol Biol **342:**1155–1169. http://dx.doi.org/10.1016/j.jmb.2004.07.083.
100. **Gold B, Rodriguez GM, Marras SA, Pentecost M, Smith I.** 2001. The Mycobacterium tuberculosis IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. Mol Microbiol **42:**851–865.

101. **Pandey R, Rodriguez GM.** 2014. IdeR is required for iron homeostasis and virulence in *Mycobacterium tuberculosis*. Mol Microbiol **91:**98–109. http://dx.doi.org/10.1111/mmi.12441.
102. **Parish T, Smith DA, Roberts G, Betts J, Stoker NG.** 2003. The senX3-regX3 two-component regulatory system of Mycobacterium tuberculosis is required for virulence. Microbiology **149:**1423–1435. http://dx.doi.org/10.1099/mic.0.26245-0.
103. **Rifat D, Bishai WR, Karakousis PC.** 2009. Phosphate depletion: a novel trigger for Mycobacterium tuberculosis persistence. J Infect Dis **200:**1126–1135. http://dx.doi.org/10.1086/605700.
104. **Voskuil MI, Schnappinger D, Visconti KC, Harrell MI, Dolganov GM, Sherman DR, Schoolnik GK.** 2003. Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. J Exp Med **198:**705–713. http://dx.doi.org/10.1084/jem.20030205.
105. **Kumar A, Toledo JC, Patel RP, Lancaster JR, Steyn AJC.** 2007. Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor. Proc Natl Acad Sci U S A **104:**11568–11573. http://dx.doi.org/10.1073/pnas.0705054104.
106. **Mehra S, Foreman TW, Didier PJ, Ahsan MH, Hudock TA, Kissee R, Golden NA, Gautam US, Johnson A-M, Alvarez X, Russell-Lodrigue KE, Doyle LA, Roy CJ, Niu T, Blanchard JL, Khader SA, Lackner AA, Sherman DR, Kaushal D.** 2015. The DosR regulon modulates adaptive immunity and is essential for Mycobacterium tuberculosis persistence. Am J Respir Crit Care Med **191:**1185–1196. http://dx.doi.org/10.1164/rccm.201408-1502OC.
107. **Voskuil MI, Schlesinger LS.** 2015. Toward resolving the paradox of the critical role of the DosR regulon in Mycobacterium tuberculosis persistence and active disease. Am J Respir Crit Care Med **191:**1103–1105. http://dx.doi.org/10.1164/rccm.201503-0424ED.
108. **Gautam US, Sikri K, Vashist A, Singh V, Tyagi JS.** 2014. Essentiality of DevR/DosR interaction with SigA for the dormancy survival program in Mycobacterium tuberculosis. J Bacteriol **196:**790–799. http://dx.doi.org/10.1128/JB.01270-13.
109. **Gautam US, McGillivray A, Mehra S, Didier PJ, Midkiff CC, Kissee RS, Golden NA, Alvarez X, Niu T, Rengarajan J, Sherman DR, Kaushal D.** 2015. DosS Is required for the complete virulence of mycobacterium tuberculosis in mice with classical granulomatous lesions. Am J Respir Cell Mol Biol **52:**708–716. http://dx.doi.org/10.1165/rcmb.2014-0230OC.
110. **Abramovitch RB, Rohde KH, Hsu F-F, Russell DG.** 2011. aprABC: a Mycobacterium tuberculosis complex-specific locus that modulates pH-driven adaptation to the macrophage phagosome. Mol Microbiol **80:**678–694. http://dx.doi.org/10.1111/j.1365-2958.2011.07601.x.
111. **Frigui W, Bottai D, Majlessi L, Monot M, Josselin E, Brodin P, Garnier T, Gicquel B, Martin C, Leclerc C, Cole ST, Brosch R.** 2008. Control of M. tuberculosis ESAT-6 secretion and specific T cell recognition by PhoP. PLoS Pathog **4:**e33. http://dx.doi.org/10.1371/journal.ppat.0040033.
112. **Gonzalo-Asensio J, Mostowy S, Harders-Westerveen J, Huygen K, Hernández-Pando R, Thole J, Behr M, Gicquel B, Martín C.** 2008. PhoP: a missing piece in the intricate puzzle of Mycobacterium tuberculosis virulence. PLoS One **3:**e3496. http://dx.doi.org/10.1371/journal.pone.0003496.
113. **Walters SB, Dubnau E, Kolesnikova I, Laval F, Daffe M, Smith I.** 2006. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. Mol Microbiol **60:**312–330. http://dx.doi.org/10.1111/j.1365-2958.2006.05102.x.
114. **Gonzalo-Asensio J, Malaga W, Pawlik A, Astarie-Dequeker C, Passemar C, Moreau F, Laval F, Daffe M, Martin C, Brosch R, Guilhot C.** 2014. Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. Proc Natl Acad Sci U S A **111:**11491–11496. http://dx.doi.org/10.1073/pnas.1406693111.
115. **White MJ, He H, Penoske RM, Twining SS, Zahrt TC.** 2010. PepD participates in the mycobacterial stress response mediated through MprAB and SigE. J Bacteriol **192:**1498–1510. http://dx.doi.org/10.1128/JB.01167-09.
116. **Bretl DJ, He H, Demetriadou C, White MJ, Penoske RM, Salzman NH, Zahrt TC.** 2012. MprA and DosR coregulate a Mycobacterium tuberculosis virulence operon encoding Rv1813c and Rv1812c. Infect Immun **80:**3018–3033. http://dx.doi.org/10.1128/IAI.00520-12.
117. **He H, Bretl DJ, Penoske RM, Anderson DM, Zahrt TC.** 2011. Components of the Rv0081-Rv0088 locus, which encodes a predicted formate hydrogenlyase complex, are coregulated by Rv0081, MprA, and DosR in Mycobacterium tuberculosis. J Bacteriol **193:**5105–5118. http://dx.doi.org/10.1128/JB.05562-11.
118. **Pang X, Samten B, Cao G, Wang X, Tvinnereim A, Chen RX-L, Howard ST.** 2013. MprAB regulates the espA operon in Mycobacterium tuberculosis and modulates ESX-1 function and host cytokine response. J Bacteriol **195:**66–75. http://dx.doi.org/10.1128/JB.01067-12.
119. **Pang X, Vu P, Byrd TF, Ghanny S, Soteropoulos P, Mukamolova GV, Wu S, Samten B, Howard ST.** 2007. Evidence for complex interactions of stress-associated regulons in an mprAB deletion mutant of Mycobacterium tuberculosis. Microbiology **153:**1229–1242. http://dx.doi.org/10.1099/mic.0.29281-0.
120. **Sureka K, Dey S, Datta P, Singh AK, Dasgupta A, Rodrigue S, Basu J, Kundu M.** 2007. Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria. Mol Microbiol **65:**261–276. http://dx.doi.org/10.1111/j.1365-2958.2007.05814.x.
121. **Zahrt TC, Wozniak C, Jones D, Trevett A.** 2003. Functional analysis of the Mycobacterium tuberculosis MprAB two-component signal transduction system. Infect Immun **71:**6962–6970. http://dx.doi.org/10.1128/IAI.71.12.6962-6970.2003.
122. **Burian J, Ramon-Garcia S, Howes CG, Thompson CJ.** 2012. WhiB7, a transcriptional activator that coordinates physiology with intrinsic drug resistance in Mycobacterium tuberculosis. Expert Rev Anti Infect Ther **10:**1037–1047. http://dx.doi.org/10.1586/eri.12.90.
123. **Casonato S, Cervantes Sánchez A, Haruki H, Rengifo González M, Provvedi R, Dainese E, Jaouen T, Gola S, Bini E, Vicente M, Johnsson K, Ghisotti D, Palù G, Hernández-Pando R, Manganelli R.** 2012. WhiB5, a transcriptional regulator that contributes to Mycobacterium tuberculosis virulence and reactivation. Infect Immun **80:**3132–3144. http://dx.doi.org/10.1128/IAI.06328-11.
124. **Chawla M, Parikh P, Saxena A, Munshi M, Mehta M, Mai D, Srivastava AK, Narasimhulu KV, Redding KE, Vashi N, Kumar D, Steyn AJC, Singh A.** 2012. *Mycobacterium tuberculosis* WhiB4 regulates oxidative stress response to modulate survival and dissemination *in vivo*. Mol Microbiol **85:**1148–1165. http://dx.doi.org/10.1111/j.1365-2958.2012.08165.x.
125. **Singh A, Crossman DK, Mai D, Guidry L, Voskuil MI, Renfrow MB, Steyn AJC.** 2009. Mycobacterium tuberculosis WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. PLoS Pathog **5:**e1000545. http://dx.doi.org/10.1371/journal.ppat.1000545.
126. **Steyn AJC, Collins DM, Hondalus MK, Jacobs WR, Kawakami RP, Bloom BR.** 2002. Mycobacterium tuberculosis WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. Proc Natl Acad Sci U S A **99:**3147–3152. http://dx.doi.org/10.1073/pnas.052705399.
127. **Turkarslan S, Peterson EJR, Rustad TR, Minch KJ, Reiss DJ, Morrison R, Ma S, Price ND, Sherman DR, Baliga NS.** 2015. A comprehensive map of genome-wide gene regulation in Mycobacterium tuberculosis. Sci Data **2:**150010. http://dx.doi.org/10.1038/sdata.2015.10.
128. **Minch KJ, Rustad TR, Peterson EJR, Winkler J, Reiss DJ, Ma S, Hickey M, Brabant W, Morrison B, Turkarslan S, Mawhinney C, Galagan JE, Price ND, Baliga NS, Sherman DR.** 2015. The DNA-binding network of Mycobacterium tuberculosis. Nat Commun **6:**5829. http://dx.doi.org/10.1038/ncomms6829.
129. **Baloni P, Chandra N.** 2015. Architectural plan of transcriptional regulation in Mycobacterium tuberculosis. Trends Microbiol **23:**123–125. http://dx.doi.org/10.1016/j.tim.2015.02.002.
130. **Avarbock D, Avarbock A, Rubin H.** 2000. Differential regulation of opposing RelMtb activities by the aminoacylation state of a tRNA.ribosome.mRNA.RelMtb complex. Biochemistry **39:**11640–11648.
131. **Dahl JL, Kraus CN, Boshoff HIM, Doan B, Foley K, Avarbock D, Kaplan G, Mizrahi V, Rubin H, Barry CE.** 2003. The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of Mycobacterium tuberculosis in mice. Proc Natl Acad Sci U S A **100:**10026–10031. http://dx.doi.org/10.1073/pnas.1631248100.
132. **Primm TP, Andersen SJ, Mizrahi V, Avarbock D, Rubin H, Barry CE.** 2000. The stringent response of Mycobacterium tuberculosis is required for long-term survival. J Bacteriol **182:**4889–4898. http://dx.doi.org/10.1128/JB.182.17.4889-4898.2000.
133. **Klinkenberg LG, Lee J-H, Bishai WR, Karakousis PC.** 2010. The stringent response is required for full virulence of Mycobacterium tuberculosis in guinea pigs. J Infect Dis **202:**1397–1404. http://dx.doi.org/10.1086/656524.
134. **Karakousis PC, Yoshimatsu T, Lamichhane G, Woolwine SC, Nuerm-**

berger EL, Grosset J, Bishai WR. 2004. Dormancy phenotype displayed by extracellular Mycobacterium tuberculosis within artificial granulomas in mice. J Exp Med **200:**647–657. http://dx.doi.org/10.1084/jem.20040646.

135. **Weiss LA, Stallings CL.** 2013. Essential roles for Mycobacterium tuberculosis Rel beyond the production of (p)ppGpp. J Bacteriol **195:**5629–5638. http://dx.doi.org/10.1128/JB.00759-13.
136. **Ross W, Vrentas CE, Sanchez-Vazquez P, Gaal T, Gourse RL.** 2013. The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. Mol Cell **50:**420–429. http://dx.doi.org/10.1016/j.molcel.2013.03.021.
137. **Vrentas CE, Gaal T, Berkmen MB, Rutherford ST, Haugen SP, Vassylyev DG, Ross W, Gourse RL.** 2008. Still looking for the magic spot: the crystallographically defined binding site for ppGpp on RNA polymerase is unlikely to be responsible for rRNA transcription regulation. J Mol Biol **377:**551–564. http://dx.doi.org/10.1016/j.jmb.2008.01.042.
138. **Liu K, Myers AR, Pisithkul T, Claas KR, Satyshur KA, Amador-Noguez D, Keck JL, Wang JD.** 2015. Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp. Mol Cell **57:**735–749. http://dx.doi.org/10.1016/j.molcel.2014.12.037.
139. **Kriel A, Bittner AN, Kim SH, Liu K, Tehranchi AK, Zou WY, Rendon S, Chen R, Tu BP, Wang JD.** 2012. Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance. Mol Cell **48:**231–241. http://dx.doi.org/10.1016/j.molcel.2012.08.009.
140. **Krásný L, Gourse RL.** 2004. An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation. EMBO J **23:**4473–4483. http://dx.doi.org/10.1038/sj.emboj.7600423.
141. **Tare P, Mallick B, Nagaraja V.** 2013. Co-evolution of specific amino acid in sigma 1.2 region and nucleotide base in the discriminator to act as sensors of small molecule effectors of transcription initiation in mycobacteria. Mol Microbiol **90:**569–583. http://dx.doi.org/10.1111/mmi.12384.
142. **Avarbock A, Avarbock D, Teh J-S, Buckstein M, Wang Z, Rubin H.** 2005. Functional regulation of the opposing (p)ppGpp synthetase/hydrolase activities of RelMtb from Mycobacterium tuberculosis. Biochemistry **44:**9913–9923. http://dx.doi.org/10.1021/bi0505316.
143. **Wayne LG, Hayes LG.** 1996. An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. Infect Immun **64:**2062–2069.
144. **Deb C, Lee C-M, Dubey VS, Daniel J, Abomoelak B, Sirakova TD, Pawar S, Rogers L, Kolattukudy PE.** 2009. A novel in vitro multiple-stress dormancy model for Mycobacterium tuberculosis generates a lipid-loaded, drug-tolerant, dormant pathogen. PLoS One **4:**e6077. http://dx.doi.org/10.1371/journal.pone.0006077.
145. **Sarathy J, Dartois V, Dick T, Gengenbacher M.** 2013. Reduced drug uptake in phenotypically resistant nutrient-starved nonreplicating Mycobacterium tuberculosis. Antimicrob Agents Chemother **57:**1648–1653. http://dx.doi.org/10.1128/AAC.02202-12.
146. **Baek S-H, Li AH, Sassetti CM.** 2011. Metabolic regulation of mycobacterial growth and antibiotic sensitivity. PLoS Biol **9:**e1001065. http://dx.doi.org/10.1371/journal.pbio.1001065.
147. **Cunningham-Bussel A, Bange FC, Nathan CF.** 2013. Nitrite impacts the survival of Mycobacterium tuberculosis in response to isoniazid and hydrogen peroxide. Microbiologyopen **2:**901–911. http://dx.doi.org/10.1002/mbo3.126.
148. **Franzblau SG, DeGroote MA, Cho SH, Andries K, Nuermberger E, Orme IM, Mdluli K, Angulo-Barturen I, Dick T, Dartois V, Lenaerts AJ.** 2012. Comprehensive analysis of methods used for the evaluation of compounds against Mycobacterium tuberculosis. Tuberculosis (Edinb) **92:**453–488. http://dx.doi.org/10.1016/j.tube.2012.07.003.
149. **Michele TM, Ko C, Bishai WR.** 1999. Exposure to antibiotics induces expression of the Mycobacterium tuberculosis sigF gene: implications for chemotherapy against mycobacterial persistors. Antimicrob Agents Chemother **43:**218–225.
150. **Arnvig KB, Comas I, Thomson NR, Houghton J, Boshoff HI, Croucher NJ, Rose G, Perkins TT, Parkhill J, Dougan G, Young DB.** 2011. Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. PLoS Pathog **7:**e1002342. http://dx.doi.org/10.1371/journal.ppat.1002342.
151. **Haning K, Cho SH, Contreras LM.** 2014. Small RNAs in mycobacteria: an unfolding story. Front Cell Infect Microbiol **4:**96. http://dx.doi.org/10.3389/fcimb.2014.00096.
152. **Pelly S, Bishai WR, Lamichhane G.** 2012. A screen for non-coding RNA in Mycobacterium tuberculosis reveals a cAMP-responsive RNA that is expressed during infection. Gene **500:**85–92. http://dx.doi.org/10.1016/j.gene.2012.03.044.
153. **DiChiara JM, Contreras-Martinez LM, Livny J, Smith D, McDonough KA, Belfort M.** 2010. Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. Nucleic Acids Res **38:**4067–4078. http://dx.doi.org/10.1093/nar/gkq101.
154. **Gupta M, Sajid A, Sharma K, Ghosh S, Arora G, Singh R, Nagaraja V, Tandon V, Singh Y.** 2014. HupB, a nucleoid-associated protein of Mycobacterium tuberculosis, is modified by serine/threonine protein kinases in vivo. J Bacteriol **196:**2646–2657. http://dx.doi.org/10.1128/JB.01625-14.
155. **Gordon BRG, Li Y, Wang L, Sintsova A, Van Bakel H, Tian S, Navarre WW, Xia B, Liu J.** 2010. Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A **107:**5154–5159. http://dx.doi.org/10.1073/pnas.0913551107.
156. **Prisic S, Husson RN.** 2014. Mycobacterium tuberculosis serine/threonine protein kinases, p 681–708. *In* Hatfull GF, Jacobs WR (ed), Molecular genetics of Mycobacteria, 2nd ed. ASM Press, Washington, DC. http://dx.doi.org/10.1128/microbiolspec.MGM2-0006-2013.
157. **Park ST, Kang C-M, Husson RN.** 2008. Regulation of the SigH stress response regulon by an essential protein kinase in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A **105:**13105–13110. http://dx.doi.org/10.1073/pnas.0801143105.
158. **Greenstein AE, MacGurn JA, Baer CE, Falick AM, Cox JS, Alber T.** 2007. M. tuberculosis Ser/Thr protein kinase D phosphorylates an anti-anti-sigma factor homolog. PLoS Pathog **3:**e49. http://dx.doi.org/10.1371/journal.ppat.0030049.
159. **Chao JD, Papavinasasundaram KG, Zheng X, Chávez-Steenbock A, Wang X, Lee GQ, Av-Gay Y.** 2010. Convergence of Ser/Thr and two-component signaling to coordinate expression of the dormancy regulon in Mycobacterium tuberculosis. J Biol Chem **285:**29239–29246. http://dx.doi.org/10.1074/jbc.M110.132894.
160. **Lee JH, Karakousis PC, Bishai WR.** 2008. Roles of SigB and SigF in the Mycobacterium tuberculosis sigma factor network. J Bacteriol **190:**699–707. http://dx.doi.org/10.1128/JB.01273-07.
161. **He H, Hovey R, Kane J, Singh V, Zahrt TC.** 2006. MprAB is a stress-responsive two-component system that directly regulates expression of sigma factors SigB and SigE in Mycobacterium tuberculosis. J Bacteriol **188:**2134–2143. http://dx.doi.org/10.1128/JB.188.6.2134-2143.2006.
162. **Abdul-Majid K-B, Ly LH, Converse PJ, Geiman DE, McMurray DN, Bishai WR.** 2008. Altered cellular infiltration and cytokine levels during early Mycobacterium tuberculosis sigC mutant infection are associated with late-stage disease attenuation and milder immunopathology in mice. BMC Microbiol **8:**151. http://dx.doi.org/10.1186/1471-2180-8-151.
163. **Dubnau E, Fontán P, Manganelli R, Soares-Appel S, Smith I.** 2002. Mycobacterium tuberculosis genes induced during infection of human macrophages. Infect Immun **70:**2787–2795. http://dx.doi.org/10.1128/IAI.70.6.2787-2795.2002.
164. **DeMaio J, Zhang Y, Ko C, Young DB, Bishai WR.** 1996. A stationary-phase stress-response sigma factor from Mycobacterium tuberculosis. Proc Natl Acad Sci U S A **93:**2790–2794. http://dx.doi.org/10.1073/pnas.93.7.2790.
165. **Hartkoorn RC, Sala C, Uplekar S, Busso P, Rougemont J, Cole ST.** 2012. Genome-wide definition of the SigF regulon in Mycobacterium tuberculosis. J Bacteriol **194:**2001–2009. http://dx.doi.org/10.1128/JB.06692-11.
166. **Williams EP, Lee J-H, Bishai WR, Colantuoni C, Karakousis PC.** 2007. Mycobacterium tuberculosis SigF regulates genes encoding cell wall-associated proteins and directly regulates the transcriptional regulatory gene phoY1. J Bacteriol **189:**4234–4242. http://dx.doi.org/10.1128/JB.00201-07.
167. **Rand L, Hinds J, Springer B, Sander P, Buxton RS, Davis EO.** 2003. The majority of inducible DNA repair genes in Mycobacterium tuberculosis are induced independently of RecA. Mol Microbiol **50:**1031–1042. http://dx.doi.org/10.1046/j.1365-2958.2003.03765.x.
168. **Dutta NK, Mehra S, Didier PJ, Roy CJ, Doyle LA, Alvarez X, Ratterree M, Be NA, Lamichhane G, Jain SK, Lacey MR, Lackner AA, Kaushal D.** 2010. Genetic requirements for the survival of tubercle bacilli in primates. J Infect Dis **201:**1743–1752. http://dx.doi.org/10.1086/652497.
169. **Veyrier F, Saïd-Salim B, Behr MA.** 2008. Evolution of the mycobacterial SigK regulon. J Bacteriol **190:**1891–1899. http://dx.doi.org/10.1128/JB.01452-07.
170. **Talaat AM, Lyons R, Howard ST, Johnston SA.** 2004. The temporal

expression profile of Mycobacterium tuberculosis infection in mice. Proc Natl Acad Sci U S A **101:**4602–4607. http://dx.doi.org/10.1073/pnas.0306023101.
171. **Fol M, Chauhan A, Nair NK, Maloney E, Moomey M, Jagannath C, Madiraju MVVS, Rajagopalan M.** 2006. Modulation of Mycobacterium tuberculosis proliferation by MtrA, an essential two-component response regulator. Mol Microbiol **60:**643–657. http://dx.doi.org/10.1111/j.1365-2958.2006.05137.x.
172. **Rickman L, Saldanha JW, Hunt DM, Hoar DN, Colston MJ, Millar JBA, Buxton RS.** 2004. A two-component signal transduction system with a PAS domain-containing sensor is required for virulence of Mycobacterium tuberculosis in mice. Biochem Biophys Res Commun **314:**259–267. http://dx.doi.org/10.1016/j.bbrc.2003.12.082.
173. **James JN, Hasan Z, Ioerger TR, Brown AC, Personne Y, Carroll P, Ikeh M, Tilston-Lunel NL, Palavecino C, Sacchettini JC, Parish T.** 2012. Deletion of SenX3-RegX3, a key two-component regulatory system of Mycobacterium smegmatis, results in growth defects under phosphate-limiting conditions. Microbiology **158:**2724–2731. http://dx.doi.org/10.1099/mic.0.060319-0.
174. **Singh N, Kumar A.** 2015. Virulence factor SenX3 is the oxygen-controlled replication switch of Mycobacterium tuberculosis. Antioxid Redox Signal **22:**603–613. http://dx.doi.org/10.1089/ars.2014.6020.
175. **Rifat D, Belchis DA, Karakousis PC.** 2014. senX3-independent contribution of regX3 to Mycobacterium tuberculosis virulence. BMC Microbiol **14:**265. http://dx.doi.org/10.1186/s12866-014-0265-8.
176. **Zahrt TC, Deretic V.** 2001. Mycobacterium tuberculosis signal transduction system required for persistent infections. Proc Natl Acad Sci U S A **98:**12706–12711. http://dx.doi.org/10.1073/pnas.221272198.
177. **Kendall SL, Movahedzadeh F, Rison SCG, Wernisch L, Parish T, Duncan K, Betts JC, Stoker NG.** 2004. The Mycobacterium tuberculosis dosRS two-component system is induced by multiple stresses. Tuberculosis (Edinb) **84:**247–255. http://dx.doi.org/10.1016/j.tube.2003.12.007.
178. **Converse PJ, Karakousis PC, Klinkenberg LG, Kesavan AK, Ly LH, Allen SS, Grosset JH, Jain SK, Lamichhane G, Manabe YC, McMurray DN, Nuermberger EL, Bishai WR.** 2009. Role of the dosR-dosS two-component regulatory system in Mycobacterium tuberculosis virulence in three animal models. Infect Immun **77:**1230–1237. http://dx.doi.org/10.1128/IAI.01117-08.
179. **Ferrer NL, Gomez AB, Neyrolles O, Gicquel B, Martin C.** 2010. Interactions of attenuated Mycobacterium tuberculosis phoP mutant with human macrophages. PLoS One **5:**e12978. http://dx.doi.org/10.1371/journal.pone.0012978.
180. **Baker JJ, Johnson BK, Abramovitch RB.** 2014. Slow growth of Mycobacterium tuberculosis at acidic pH is regulated by phoPR and host-associated carbon sources. Mol Microbiol **94:**56–69. http://dx.doi.org/10.1111/mmi.12688.
181. **Pérez E, Samper S, Bordas Y, Guilhot C, Gicquel B, Martín C.** 2001. An essential role for phoP in Mycobacterium tuberculosis virulence. Mol Microbiol **41:**179–187. http://dx.doi.org/10.1046/j.1365-2958.2001.02500.x.
182. **Tan S, Sukumar N, Abramovitch RB, Parish T, Russell DG.** 2013. Mycobacterium tuberculosis responds to chloride and pH as synergistic cues to the immune status of its host cell. PLoS Pathog **9:**e1003282. http://dx.doi.org/10.1371/journal.ppat.1003282.
183. **Malhotra V, Agrawal R, Duncan TR, Saini DK, Clark-Curtiss JE.** 2015. Mycobacterium tuberculosis response regulators, DevR and NarL, interact in vivo and co-regulate gene expression during aerobic nitrate metabolism. J Biol Chem **290:**8294–8309. http://dx.doi.org/10.1074/jbc.M114.591800.
184. **Haydel SE, Benjamin WH, Dunlap NE, Clark-Curtiss JE.** 2002. Expression, autoregulation, and DNA binding properties of the Mycobacterium tuberculosis TrcR response regulator. J Bacteriol **184:**2192–2203. http://dx.doi.org/10.1128/JB.184.8.2192-2203.2002.
185. **Lane WJ, Darst SA.** 2010. Molecular evolution of multisubunit RNA polymerases: sequence analysis. J Mol Biol **395:**671–685. http://dx.doi.org/10.1016/j.jmb.2009.10.062.

Kelly Flentie is currently a postdoctoral associate in the laboratory of Christina Stallings at Washington University in St. Louis. She obtained her B.S. in microbiology from the University of Kansas in 2004. As an undergraduate, she completed her honors research studying the pathogenesis of *Shigella flexneri* under the guidance of William Picking. She completed her Ph.D. at Washington University in St. Louis in 2011 in the laboratory of David Piwnica-Worms. In her doctoral thesis research, she investigated interactions between *Salmonella* and cancer cells, with an emphasis on characterizing bacterial behaviors that could be coopted for novel cancer treatment or diagnostic strategies. Her current research focuses on interrogating mechanisms of stress tolerance in *Mycobacterium tuberculosis* and identifying new ways to target this pathogen.

Ashley L. Garner is a Ph.D. student at Washington University in St. Louis. In 2008, she received a B.S. in microbiology at the University of California, Davis, where she interned in the laboratory of Michele Igo, studying an autotransporter in the plant pathogen *Xylella fastidiosa*. After graduating, she worked for the R&D Department of Novozymes, studying cellulosic ethanol production. She joined the laboratory of Christina Stallings in 2010, where she currently researches the mycobacterial transcription regulator CarD.

Christina L. Stallings is an assistant professor in the Department of Molecular Microbiology at Washington University in St. Louis. She received her Ph.D. with distinction from Columbia University College of Physicians and Surgeons, where she performed her thesis work on alphaherpesviruses in the laboratory of Saul Silverstein. She then transitioned to another fascinating and chronic pathogen, *Mycobacterium tuberculosis*, for her postdoctoral research in Michael Glickman's laboratory at the Sloan-Kettering Institute. She started in her faculty position at Washington University in St. Louis in 2010, and research in her laboratory seeks to dissect the molecular mechanisms involved in *M. tuberculosis* pathogenesis and stress responses.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

MINIREVIEW

Two-Component Signal Transduction Systems That Regulate the Temporal and Spatial Expression of *Myxococcus xanthus* Sporulation Genes

Zaara Sarwar,[a] Anthony G. Garza[b]

Department of Chemistry, State University of New York College of Environmental Science and Forestry, Syracuse, New York, USA[a]; Department of Biology, Syracuse University, Syracuse, New York, USA[b]

When starved for nutrients, *Myxococcus xanthus* produces a biofilm that contains a mat of rod-shaped cells, known as peripheral rods, and aerial structures called fruiting bodies, which house thousands of dormant and stress-resistant spherical spores. Because rod-shaped cells differentiate into spherical, stress-resistant spores and spore differentiation occurs only in nascent fruiting bodies, many genes and multiple levels of regulation are required. Over the past 2 decades, many regulators of the temporal and spatial expression of *M. xanthus* sporulation genes have been uncovered. Of these sporulation gene regulators, two-component signal transduction circuits, which typically contain a histidine kinase sensor protein and a transcriptional regulator known as response regulator, are among the best characterized. In this review, we discuss prototypical two-component systems (Nla6S/Nla6 and Nla28S/Nla28) that regulate an early, preaggregation phase of sporulation gene expression during fruiting body development. We also discuss orphan response regulators (ActB and FruA) that regulate a later phase of sporulation gene expression, which begins during the aggregation stage of fruiting body development. In addition, we summarize the research on a complex two-component system (Esp) that is important for the spatial regulation of sporulation.

Myxococcus xanthus is a rod-shaped deltaproteobacterium. The natural habitat of *M. xanthus* is topsoil, where it contributes to the "earthy smell" by producing the sesquiterpene geosmin (1). In its natural environment, *M. xanthus* is a microbial predator, hunting in swarming biofilms that collectively feed on other bacteria using hydrolytic enzymes (2–4). When prey bacteria are not available, *M. xanthus* cells form a second type of biofilm that contains a mat of rod-shaped cells, known as peripheral rods, and fruiting bodies containing thousands of metabolically dormant spores (Fig. 1) (5). The spores inside fruiting bodies have thick protective coats that provide resistance to environmental stresses and allow them to survive until nutrients for growth become available, an event which triggers spore germination and, eventually, the formation of swarms that engage in group feeding.

The process of spore formation in *M. xanthus*, and in related species known collectively as the myxobacteria, is fundamentally different from endospore formation in *Bacillus* sp., which is the best-characterized model of bacterial sporulation (6). For example, *M. xanthus* sporulation is a process by which one cell type (a rod-shaped metabolically active cell) differentiates into another cell type (a spherical spore that is dormant and stress resistant). In contrast, *Bacillus* sporulation is not a true cell differentiation: an asymmetric cell division event yields a relatively large mother cell and a smaller cell that eventually becomes a dormant spore. In addition, a *Bacillus* spore develops inside a mother cell, which protects it from the environment, whereas a developing *M. xanthus* spore is directly exposed to the environment. As a consequence, *M. xanthus* must have a mechanism for maintaining the integrity of the cell envelope as it is being reorganized during spore differentiation.

An interesting feature of *M. xanthus* sporulation is its spatial restriction to nascent fruiting bodies; the peripheral rods that surround these structures fail to differentiate into spores. Hence, sporulation in *M. xanthus* is under strict temporal and spatial control. Indeed, it was shown that expression of certain genetic loci that are important for sporulation is spatially localized to the aggregates of cells that develop into fruiting bodies (7, 8).

Abbreviations used in this article are as follows: TCS, two-component system; DHp domain, dimerization and histidine phosphorylation domain; CA domain, catalytic and ATP binding domain; EBP, enhancer binding protein; CRP, cyclic AMP receptor protein; and FNR, fumarate-nitrate reduction.

REGULATORS OF SPORULATION

A number of developmental signals that regulate the formation of spores inside nascent fruiting bodies have been identified, including the intracellular starvation signal (p)ppGpp (9–12), an extracellular cell density reporter known as A-signal (13–16), and a contact-stimulated cell-cell signal known as C-signal (17–22). For more detailed discussions of *M. xanthus* developmental signals, see reviews in references 23–26, and 27.

Signal transduction proteins, which are abundant in *M. xanthus* (28–30), have also been implicated in the regulation of sporulation genes (31, 32). Of these sporulation regulators, two-component signal transduction proteins are perhaps the best-characterized group.

The prototypical TCS contains a sensor histidine kinase protein and a response regulator protein (33). Histidine kinases typically contain a sensor input domain and a transmitter domain

Accepted manuscript posted online 14 September 2015

Citation Sarwar Z, Garza AG. 2016. Two-component signal transduction systems that regulate the temporal and spatial expression of *Myxococcus xanthus* sporulation genes. J Bacteriol 198:377–385. doi:10.1128/JB.00474-15.

Editor: W. Margolin

Address correspondence to Anthony G. Garza, agarza@syr.edu.

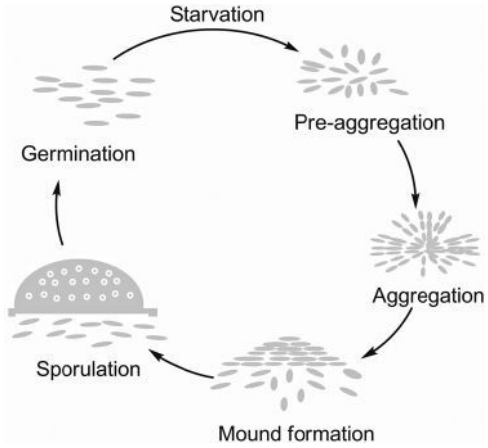


FIG 1 The *Myxococcus xanthus* life cycle. *M. xanthus* is a microbial predator that obtains nutrients by forming swarming biofilms and collectively feeding on other bacteria. When prey bacteria are not available to provide nutrients, cells undergo a multicellular developmental cycle that culminates with the formation of spore-filled fruiting bodies. The stages of development include preaggregation (1 to 5 h poststarvation), aggregation and mound formation (6 to 18 h poststarvation), and sporulation (24 to 120 h poststarvation). Sporulation occurs inside the dome-shaped mounds and is the process by which rod-shaped cells are converted into spherical, stress-resistant spores. The availability of nutrients, which presumably are provided by prey bacteria, triggers spore germination and, eventually, the formation of swarms that engage in group feeding.

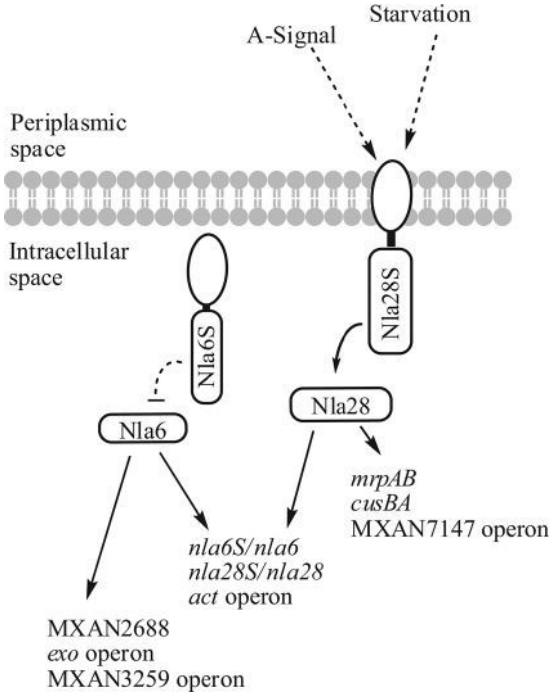


FIG 2 Diagram of the Nla6S/Nla6 and Nla28S/Nla28 two-component systems. Nla6S/Nla6 and Nla28S/Nla28 are early-acting two-component systems that directly regulate expression of sporulation genes. Nla6S and Nla28S are predicted to be cytoplasmic and membrane-bound histidine kinase sensors, respectively (28). The signal detected by Nla6S is unknown; however, it has been proposed that the cell density reporter known as A-signal or nutrient levels (starvation) may be the activating signal for Nla28S (67). It is believed that Nla6S modulates the activity of the Nla6 response regulator via dephosphorylation, whereas Nla28S modulates the activity of the Nla28 response regulator via phosphorylation (38, 67). Nla6 is a transcriptional activator that regulates genetic loci such as *exo*, MXAN2688, and MXAN3259 that are important for spore differentiation and spore stress resistance (59). Nla28 is a transcriptional activator that regulates genetic loci such as *cusBA* and MXAN7147, which are primarily involved in spore stress resistance (Li et al., unpublished). Nla6 and Nla28 also modulate expression of other transcriptional regulators that are important for sporulation: Nla6 modulates expression of *actB* and *nla28*, and Nla28 modulates expression of *actB*, *mrpAB*, and *nla6* (43) (Li et al., unpublished). In addition, Nla6 and Nla28 are involved in autoregulation (43).

(34). The sensor input domain is often a transmembrane domain for detecting extracellular signals and has high sequence variability. In contrast, the transmitter domain is a conserved cytoplasmic domain containing the DHp domain and the CA domain. When a signal is detected by the sensor domain of the histidine kinase, the CA domain transfers a phosphate from ATP to the conserved histidine residue within the DHp domain. The phosphoryl group is then transferred from the histidine kinase to a conserved aspartate residue in the receiver domain of the response regulator (35). Phosphorylation of the response regulator's receiver domain causes a conformational change in its effector domain, which is often a DNA binding domain that allows the response regulator to modulate transcription (36).

In addition to classical TCSs, histidine kinases and response regulators are used to create more-complex signal transduction systems (33). One common variant of the TCS is the phosphorelay system. In such a system, the phosphoryl group is passed from a hybrid histidine kinase, which contains a transmitter domain and a receiver domain, to a histidine phosphotransferase protein and then to a response regulator. TCSs are also frequently organized as branched systems in which one histidine kinase interacts with multiple response regulators or vice versa. In addition, many histidine kinases are bifunctional and can phosphorylate or dephosphorylate their cognate response regulators based on the stimuli (37).

In this review, we discuss prototypical TCSs that are important for an early, preaggregation phase of sporulation gene expression and orphan response regulators that are important for a later, aggregation phase of sporulation gene expression. We also discuss the Esp system, which is a complex TCS that helps ensure that sporulation is spatially localized in nascent fruiting bodies.

TWO-COMPONENT SYSTEM PROTEINS THAT REGULATE *M. XANTHUS* SPORULATION

Early regulators of sporulation. (i) Nla6S/Nla6. Nla6S (S for histidine kinase sensor) and Nla6 form an early-acting two-component system that begins functioning about 1 h after starvation initiates fruiting body development (Fig. 1 and 2). Nla6S is a cytoplasmic protein that has a remarkable property; it lacks many of the conserved sequence motifs of typical histidine kinases but still has all the *in vitro* properties of histidine kinases (38). In particular, an alignment of the C-terminal transmitter region of Nla6S with those of known histidine kinases uncovered a putative DHp domain; however, no CA domain was identified (28, 38–40). A closer look at the predicted secondary structure of the C-terminal transmitter region of Nla6S revealed characteristics of a functional CA domain. Moreover, biochemical analyses indicate that both the DHp and CA domains in the Nla6S transmitter region are functional (38). In subsequent phylogenetic analyses, potential orthologs of Nla6S were found, but only in the sequenced, fruiting

members of the *Cystobacterineae* suborder of the myxobacteria (38). Taken together, these data indicate that Nla6S is the prototype for a new family of histidine kinases thus far found only in fruiting *Cystobacterineae*.

Nla6 is a member of the EBP family of response regulators (41–43), which work with the σ^{54} protein to activate transcription at σ^{54} promoter elements. In particular, σ^{54} directs RNA polymerase to conserved sequences in the −24 and −12 regions of target promoters (44) and EBPs, which bind to tandem repeat sequences located upstream of the −24 and −12 regions (45–47), use the energy from ATP hydrolysis to help σ^{54}-RNA polymerase form an open promoter complex, and initiate transcription (48–50).

Since the *nla6S* histidine kinase and *nla6* response regulator genes are cotranscribed (28, 43) and two-component system partners are often regulated in such a manner, it has been assumed that Nla6S and Nla6 form a two-component pair. When this idea was tested experimentally, however, no *in vitro* transfer of a phosphoryl group from Nla6S to Nla6 was detected (38). Perhaps an additional component is required to facilitate the *in vitro* phosphotransfer from Nla6S to Nla6. Alternatively, the primary function of Nla6S may be to dephosphorylate Nla6; some proteins with similarity to histidine kinases, such as RedE in the Red signal transduction system of *M. xanthus* and CheA3 in the Crd/Che3 signal transduction system of *M. xanthus*, function as phosphatases (37, 51–53). This dephosphorylation activity, known as transmitter phosphatase activity, is mediated by a conserved D/EXXT/N motif found immediately adjacent to the phosphoaccepting histidine residue in the DHp domain (54, 55). Indeed, Nla6S has a DXXN motif immediately adjacent to the putative phosphate-accepting histidine residue in its DHp domain, suggesting that Nla6S could act as an Nla6 phosphatase. Thus, Nla6 may be phosphorylated by another histidine kinase or a small-molecule phosphate donor such as acetyl phosphate *in vivo* (56–58) and Nla6S may regulate Nla6 activity via dephosphorylation.

Attempts to inactive *nla6S* have thus far been unsuccessful. However, the function of Nla6S has been inferred from studies of Nla6. A mutation in *nla6* slightly delays but does not inhibit the formation of the tightly packed cell aggregates that become fruiting bodies. In contrast, its effect on the sporulation process is dramatic, reducing the number of viable, stress-resistant spores 500-fold compared to the level seen with the wild type (42). Recent work indicated that the *nla6* mutation affects the process by which rod-shaped cells inside nascent fruiting bodies differentiate into spherical spores and the spores' acquisition of stress resistance (59). These findings led to the proposal that the primary developmental function of the Nla6s/Nla6 TCS is to regulate production of stress-resistant spores.

To understand how the Nla6S/Nla6 TCS regulates sporulation, potential developmental targets of Nla6 were identified using its 10-bp tandem repeat binding site and the *M. xanthus* genome sequence (28, 43, 59). Nineteen operons containing 67 genes and 21 single genes were tagged as potential Nla6 targets using this strategy. Most of the 24 genes that have been experimentally confirmed to be Nla6 targets fall into two functional categories: (i) transcriptional regulators of sporulation such as *nla28S-nla28* and *actB* and (ii) genes that are important for spore differentiation and stress resistance, among which the *exo* operon is the best studied (8, 60–62). The *exo* operon contains nine genes (*exoA* to *exoI*), and most of those genes are known to be important for the export of spore coat polysaccharide, spore differentiation, and spore stress resistance (60, 61, 63). Other Nla6 targets that are important for spore differentiation and stress resistance include the MXAN3259 locus. The MXAN3259 gene encodes a putative member of the polysaccharide deacetylase family of enzymes, which are known to be involved in sporulation in bacteria such as *Bacillus subtilis* and *Streptococcus pneumoniae* (64, 65). The predicted functions of the unconfirmed targets of Nla6 include transcriptional regulation/signal transduction, cell wall/membrane biogenesis, and solute transport.

On the basis of recent studies, it has been suggested that the Nla6S/Nla6 TCS may be a general regulator of stress-associated genes (59, 66). Presumably, the confirmed developmental targets of the Nla6S/Nla6 TCS fall into this category, since they show Nla6-dependent activation 1 h after *M. xanthus* cells encounter starvation-induced stress (59). This early activation occurs well before Nla6S/Nla6 TCS targets are predicted to function; they are important for the sporulation, which starts after about 24 h of development (after the nascent fruiting body is constructed). It was suggested that the Nla6S/Nla6-mediated activation of early genes helps prepare cells for spore differentiation later in development (59).

(ii) Nla28S/Nla28. Nla28S (S for histidine kinase sensor) and Nla28 form a second two-component system that functions in the early stages (1 h poststarvation) of fruiting body development (Fig. 1 and 2). Nla28S and Nla28 were first tagged as potential two-component system partners based on DNA sequence and expression data indicating that *nla28S* and *nla28* genes are cotranscribed (28, 43). Nla28S is a transmembrane histidine kinase. *In vitro* studies showed that Nla28S is a functional histidine kinase with K_m and k_{cat} values comparable to those of other well-studied histidine kinases and that it specifically transfers a phosphoryl group to Nla28 (67), which is a member of the EBP family of response regulators (28, 43).

Two pieces of evidence suggest that Nla28S/Nla28 is an early-functioning two-component system. First, the *nla28S* and *nla28* genes are expressed shortly after starvation initiates development (59, 67). Second, a mutation in *nla28* starts to affect developmental gene expression patterns 1 h poststarvation (43). The early signal to which the Nla28S kinase responds has yet to be identified. However, one candidate is A-signal, which is an early-acting cell density reporter (13, 14, 16). In particular, expression of *nla28S* increases when exogenous A-signal is added to A-signal-deficient cells (67) and, since expression of the *nla28S-nla28* operon is autoregulated (43), it is possible that Nla28S is involved in the detection of A-signal.

Because *nla28S* is expressed at the onset of development, it is also possible that Nla28S monitors nutrient levels. Indeed, the developmental phenotypes of the *nla28S* mutant are different on media containing different levels of nutrients (67). While no significant defect in aggregation or sporulation efficiency was observed when the Δ*nla28S* strain developed on stringent-starvation agar containing no added nutrients, aggregation was delayed and the sporulation efficiency was reduced when the strain was placed on slow-starvation agar containing low levels of essential nutrients. Nutrient level-dependent developmental phenotypes were previously observed in an analysis of the *asgD* mutant (68). It was suggested that AsgD, a putative hybrid histidine kinase, is involved in detecting nutrient levels and in perceiving starvation. It was also suggested that nutrient levels must be relatively low for the *asgD* mutant cells to detect starvation; cells are unable to detect starva-

tion on agar containing low levels of nutrients, even though the nutrient levels are too low to sustain growth, but are capable of detecting starvation when no nutrients are added. Perhaps Nla28S has a nutrient-sensing function similar to that of AsgD, which would be consistent with developmental phenotypes of the *nla28S* mutant.

Mutations in the *nla28S* or *nla28* gene primarily affect the process of sporulation (42, 67), although it seems that the *nla28* mutation does cause a slight delay in the formation of tightly packed aggregates of cells (42). The sporulation phenotype of the *nla28* mutation has been characterized in detail, and it has little (if any) effect on the cell shape change associated with sporulation, but it causes the number of viable, stress-resistant spores to be reduced about 50-fold (42). This, together with the sporulation defect of the Δ*nla28S* strain on slow-starvation agar, led to the proposal that the Nla28S/Nla28 TCS regulates the acquisition of spore stress resistance properties. In contrast, its early-function counterpart, the Nla6S/Nla6 TCS, appears to regulate the shape change associated with sporulation and the spores' acquisition of stress resistance.

To date, 12 developmentally regulated operons containing 38 genes have been confirmed as direct Nla28 targets using the expression profiles of *nla28* mutant cells and *in vitro* promoter binding assays (43) (T. Li, D. Lemon, K. Murphy, and A. Garza, unpublished data). In addition, about 60 single genes or operons have been tagged as putative Nla28 targets using the 8-bp direct repeat binding site of Nla28 and the *M. xanthus* genome sequence (Li et al., unpublished).

Several of the confirmed Nla28 target operons contain genes for TCS proteins. These include the *nla6S-nla6* operon and the operon containing *actB*, which is an EBP-type response regulator that functions downstream of *nla6S-nla6* and *nla28S-nla28*. It also includes the operon containing *mrpA* and *mrpB*, which code for a putative histidine kinase and an EBP-type response regulator, respectively (69, 70). MrpB is thought to be a direct regulator of the *mrpC* gene. MrpC appears to be a member of the CRP/FNR family of transcription factors, and it coregulates several sporulation genes with the FruA response regulator-like protein (70–74). FruA and ActB are described in more detail below.

Most of the confirmed targets of Nla28 were previously uncharacterized. Insertions in each of these loci produced developmental phenotypes similar to that of an *nla28* mutation (Li et al., unpublished); they caused a slight delay in the formation of tightly packed cell aggregates and a strong defect in spore stress resistance. Interestingly, the putative products of many of the genes in these Nla28 target loci have similarity to proteins involved in stress resistance in other bacteria, which is consistent with the proposed spore stress resistance function of the Nla28S/Nla28 TCS.

Like the confirmed targets of the Nla6S/Nla6 TCS, the confirmed targets of the Nla28S/Nla28 TCS are activated early in development (1 h poststarvation) and yet they are predicted to function during the late, sporulation stage of development (Li et al., unpublished). Hence, the activation of early genes by both the Nla28S/Nla28 and Nla6S/Nla6 TCSs may help prepare cells for sporulation later in development.

LATE REGULATORS OF SPORULATION

(i) ActA and ActB. ActA and ActB are encoded by genes in the *act* operon. ActA is a putative response regulator that is predicted to have a GGDEF-type effector domain, which is found in diguanylate cyclase enzymes. ActB appears to be a member of the EBP family of response regulators (75). Developmental expression of the *act* operon is subjected to multiple levels of regulation, including direct regulation by the early-functioning Nla6S/Nla6 and Nla28S/Nla28 TCSs (43, 59) (Fig. 2), direct autoregulation via ActB (43), and direct or indirect regulation by the CsgA C-signaling protein and the FruA response regulator-like protein (76). The fact that the *act* operon is the hub for a large amount of regulatory inputs suggests that it has an important developmental function. Indeed, as discussed below, ActB may be a key regulatory switch for sporulation inside a nascent fruiting body.

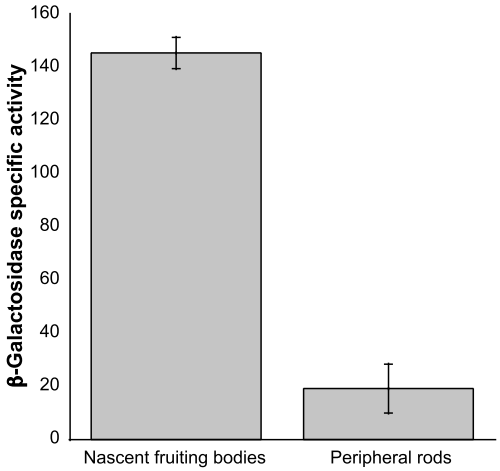


FIG 3 Expression of an *actB*::*lacZ* fusion in fruiting body cells and peripheral rods. Cells were harvested at 24 h of development, and the fruiting body cells and peripheral rods were separated by differential centrifugation as previously described (98). Mean β-galactosidase-specific activities from three independent replicates are shown (Garza, unpublished). Error bars are standard deviations of the means.

Mutational analyses indicate that *actA* and *actB* are required for sporulation inside nascent fruiting bodies but not for the formation of aggregates of cells (75). This finding and the fact that *actA* and *actB* are cotranscribed suggest that both proteins are required for sporulation and that they work in the same signal transduction pathway (75). Little else is known about the function of *actA*; however, *actB* function has been studied in some detail.

Gronewold and Kaiser showed that developmental gene expression patterns go awry in an *actB* deletion mutant only after aggregation is initiated (77), which is consistent with the idea that ActB is a sporulation regulator. Recently, we used differential centrifugation to examine *actB* expression in two developmental cell types: cells in nascent fruiting bodies, which eventually differentiate spores, and the peripheral rods, which fail to differentiate into spores (A. Garza, unpublished data). As shown in Fig. 3, expression of the *actB* gene is about 5-fold to 7-fold higher in cells isolated from nascent fruiting bodies than in the peripheral rods that surround the nascent fruiting bodies. This finding, together with the observed phenotypes of the *actB* mutant strain, suggest that ActB is a key regulatory switch that activates expression of sporulation genes inside nascent fruiting bodies and, in doing so, helps ensure that spore differentiation is spatially restricted to this structure. It is worth noting that ActB positively autoregulates the *act* operon and that such regulation is thought to be important for switch-like functions, which dramatically boost expression of key

transcription factors (78, 79). The histidine kinase sensor that activates ActB via phosphorylation and presumably triggers ActB-mediated expression of sporulation genes in nascent fruiting bodies has yet to be identified.

It was proposed that ActB triggers sporulation inside nascent fruiting bodies by directly regulating the *csgA* gene (75). This proposal was based on two pieces of data: an *actB* deletion reduces expression of the *csgA* gene and the CsgA protein, which is crucial for C-signaling (80, 81), and relatively high levels of C-signal are required to induce sporulation (82, 83). However, we failed to find a good match to the putative ActB binding site (43) or a good match to the σ^{54} promoter consensus sequence when we searched 1,000 bp of DNA upstream of the *csgA* gene. Thus, it seems unlikely that ActB directly regulates developmental expression of *csgA*. To date, the only direct ActB targets that have been identified are the *act* operon itself and the MXAN4899 gene (43), which encodes an EBP-type transcriptional regulator that is primarily involved in sporulation (84, 85). Clearly, additional developmental targets of ActB must be identified to better understand how it regulates sporulation inside developing fruiting bodies.

(ii) FruA. FruA has similarity to response regulators (86, 87); however, FruA lacks residues that are important for phosphorylation of other response regulators, a histidine kinase that phosphorylates FruA *in vitro* has yet to be identified, and several lines of evidence indicate that the *in vitro* DNA binding affinity of FruA is not altered by phosphorylation (73). Hence, it is unclear whether FruA is part of a TCS and whether phosphorylation is required for its *in vivo* activity.

FruA is part of the C-signaling network and is important for both aggregation and sporulation (86, 87). For a discussion of FruA's roles in aggregation and in the C-signaling network, see previous reviews in references 24, 88, 89, and 32; they will not be discussed in detail here. The results of several studies indicate that FruA plays a direct role in the regulation of sporulation. For example, FruA directly regulates the *exo* operon (90), which is important for the export of spore coat polysaccharide (61). As mentioned above, the *exo* operon is also a direct target of the Nla6S/Nla6 TCS (59). In addition, the *dev* operon and the *fmgBC* operon, which play important but undefined roles in sporulation, are coregulated by FruA and CRP-like transcription factor MrpC (72, 91).

REGULATORS OF THE TIMING OF SPORULATION

The Esp system. The Esp TCS is required for the temporal and spatial regulation of sporulation (31, 92, 93). Esp is a well-characterized complex TCS and an excellent example of the plasticity of signal transduction networks in *M. xanthus*; the Esp TCS contains two hybrid histidine kinases and two serine/threonine protein kinases (Fig. 4). The Esp components of this TCS are encoded by the *espAB* genes, which are in one operon, and the *espC* gene, which is located at a separate locus. The *espA* and *espC* genes encode hybrid histidine kinases (cytosolic and membrane bound, respectively), and the *espB* gene encodes a putative oligopeptide transport membrane protein. Cells carrying an *espA* or *espC* deletion aggregate and sporulate earlier than the wild-type cells, whereas cells carrying an *espB* deletion aggregate and sporulate many hours after wild-type cells (92, 93). Furthermore, in contrast to wild-type cells, which sporulate in nascent fruiting bodies, sporulation of the Δ*espA* mutant and the Δ*espC* mutant can occur outside these structures (92, 93).

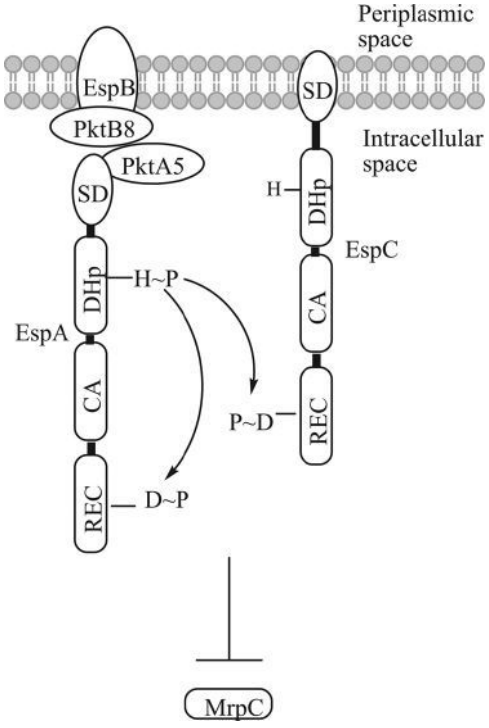


FIG 4 Diagram of the Esp two-component pathway. Esp is a complex two-component pathway that regulates the timing of sporulation (92, 93, 95). This two-component pathway consists of the following proteins: the cytoplasmic EspA and membrane-bound EspC hybrid histidine kinases, which contain a histidine kinase transmitter domain and a response regulator receiver domain, the serine/threonine protein kinases called PktA5 and PktB8, and the putative EspB oligopeptide transport membrane protein. Several lines of evidence indicate that EspA and EspC form a signaling unit; EspA performs autophosphorylation via its transmitter domain and then transfers a phosphoryl group to its receiver domain and the receiver domain of EspC (94). It has been proposed that EspB, PktA5, and PktB8 work together to modulate the activity of EspA (92, 95). It is believed that the EspAC signaling unit directly or indirectly inhibits the accumulation of MrpC, which is a positive regulator of sporulation, until the proper time in fruiting body development (97). SD, sensor domain; DHp, dimerization and histidine phosphorylation domain; CA, catalytic and ATP binding domain; REC, receiver domain.

The EspA and EspC proteins are hybrid histidine kinases that contain a histidine kinase transmitter domain and a response regulator receiver domain. Recent *in vitro* studies showed that the transmitter domain of EspA transfers a phosphoryl group to its own receiver domain (94). Furthermore, the transmitter domain of EspA, but not the transmitter domain of EspC, is capable of transferring a phosphoryl group to the receiver domain of EspC (94). This result and the similarity of the early sporulation phenotypes of the *espA* and *espC* mutants (92, 93) suggest that EspA and EspC form a signaling unit that inhibits sporulation until the proper time in fruiting body development.

In addition to the EspA and EspC hybrid histidine kinases, the Esp TCS contains serine/threonine protein kinases called PktA5 and PktB8. PktA5 and PktB8 were originally identified as potential components in the Esp TCS based on the chromosomal location of their corresponding genes: *pktA5* and *pktB8* are located immediately upstream and downstream of the *espAB* genes, respectively (95). The results of epistasis experiments, coupled with the similarity of the *pktA5* mutant, *pktB8* mutant, and *espB* mutant developmental phenotypes, led to the proposal that EspB, PktA5, and

PktB8 work together to modulate the activity of the EspAC signaling unit (92, 95).

Two pieces of evidence support the idea that PktA5 and PktB8 interact with the EspAC signaling unit via EspA. First, *in vitro* studies indicate that PktA5 autophosphorylates on a threonine residue(s) (95) and that the sensor region of EspA contains a putative forkhead-associated domain, which is a phosphopeptide recognition domain with specificity toward phosphothreonine (96). Second, using lysates from developing cells and immunoprecipitation, interactions between EspA and both of the Pkt proteins were identified (95). Interactions between EspB, which is predicted to be membrane bound, and EspAC have not been identified; however, it has been suggested that EspB interacts with EspA indirectly via a Pkt protein and through this interaction relays environmental information (92, 95).

How does the Esp TCS influence the timing of sporulation? Higgs et al. (97) showed that MrpC protein levels, but not *mrpC* mRNA levels, increase earlier in development in the Δ*espA* mutant than in a wild-type strain. In addition, it was shown that a deletion of *espA* and *espC* affects the turnover of MrpC (94). Since MrpC is a crucial, positive regulator of sporulation, this finding suggests that the Esp TCS directly or indirectly controls the timing of MrpC accumulation and, as a consequence, the timing of sporulation in developing cells. Presumably, the Esp TCS inhibits the accumulation of MrpC and the onset of sporulation until the appropriate time in fruiting body development.

SUMMARY AND FUTURE WORK

When starved for nutrients, *M. xanthus* produces a biofilm that contains a mat of peripheral rods and aerial structures called fruiting bodies, which house thousands of dormant and stress-resistant spherical spores. Because *M. xanthus* spores represent a differentiated cell type and this cellular differentiation occurs only in nascent fruiting bodies, many genes and multiple levels of regulation are required for sporulation.

In this review, we have described two phases of induction of *M. xanthus* sporulation genes: an early phase that begins prior to the onset of aggregation and a later phase that begins after aggregation commences. We have discussed prototypical TCSs (Nla6S/Nla6 and Nla28S/Nla28) that are involved in the early phase of sporulation gene expression and orphan response regulators (ActB and FruA) that are involved in the later phase of sporulation gene expression. We have also discussed the Esp system, which is a complex TCS that regulates the timing of sporulation.

Many of the promoter/gene targets (output) of these TCS and TCS components have been identified, whereas little is known about the extracellular or intracellular signals (inputs) that modulate their activities. In fact, the only TCS for which a candidate activating signal (nutrient levels or A-signal) has been identified is Nla28S/Nla28. In the case of the orphan response regulator ActB, the histidine kinase that detects the input signal is also unknown. In the case of the response regulator-like protein FruA, it is unclear whether it has a signal-detecting histidine kinase partner.

There are a number of interesting and unanswered questions about the function of the sporulation regulators discussed here. Why do the Nla6S/Nla6 and Nla28S/Nla28 TCSs activate expression of their sporulation gene targets well before cells start to construct the structures that will eventually house the spores? It has been suggested that these TCSs help prepare cells for sporulation early in the developmental process (59). Perhaps these TCSs activate genes that make cells more stress resistant, allowing them to cope with the prolonged period (about 12 to 24 h) of starvation that precedes the formation of nascent fruiting bodies. Since only a fraction of the cells that enter development become spores, it would be interesting to know whether the Nla6S/Nla6 and Nla28S/Nla28 TCS targets are expressed only in this subpopulation of developing cells.

The Esp TCS seems to function as a negative regulatory switch, ensuring that sporulation is inhibited until nascent fruiting bodies have been built (94). Does ActB (and, presumably, its histidine kinase partner) serve as a positive regulatory switch that activates expression of sporulation genes inside nascent fruiting bodies, as suggested? Are the apparent opposing activities of the Esp TCS and ActB coordinately regulated? Perhaps the answers to these questions will help address the long-standing problem of how sporulation in *M. xanthus* is restricted to the nascent fruiting body. Of course, there are other regulators such as the contact-stimulated cell-cell signal known as C-signal that have been implicated in the spatially localized sporulation inside nascent fruiting bodies (7). The challenge for the future will be to determine the connections between the different spatial regulators of sporulation in *M. xanthus* and to connect these regulatory circuits to those functioning early in development.

ACKNOWLEDGMENT

Much of the work on ActB, Nla6, and Nla28 was supported by National Science Foundation grant IOS-1354779 to Anthony G. Garza.

FUNDING INFORMATION

National Science Foundation (NSF) provided funding to Anthony G. Garza under grant number IOS-1354779.

REFERENCES

1. **Dickschat JS, Wenzel SC, Bode HB, Muller R, Schulz S.** 2004. Biosynthesis of volatiles by the myxobacterium *Myxococcus xanthus*. Chembiochem **5:**778–787. http://dx.doi.org/10.1002/cbic.200300813.
2. **Berleman JE, Kirby JR.** 2009. Deciphering the hunting strategy of a bacterial wolfpack. FEMS Microbiol Rev **33:**942–957. http://dx.doi.org/10.1111/j.1574-6976.2009.00185.x.
3. **Rosenberg E, Keller KH, Dworkin M.** 1977. Cell density-dependent growth of *Myxococcus xanthus* on casein. J Bacteriol **129:**770–777.
4. **Varon M, Cohen S, Rosenberg E.** 1984. Autocides produced by *Myxococcus xanthus*. J Bacteriol **160:**1146–1150.
5. **O'Connor KA, Zusman DR.** 1991. Development in *Myxococcus xanthus* involves differentiation into two cell types, peripheral rods and spores. J Bacteriol **173:**3318–3333.
6. **Higgins D, Dworkin J.** 2012. Recent progress in *Bacillus subtilis* sporulation. FEMS Microbiol Rev **36:**131–148. http://dx.doi.org/10.1111/j.1574-6976.2011.00310.x.
7. **Julien B, Kaiser AD, Garza A.** 2000. Spatial control of cell differentiation in *Myxococcus xanthus*. Proc Natl Acad Sci U S A **97:**9098–9103. http://dx.doi.org/10.1073/pnas.97.16.9098.
8. **Müller FD, Treuner-Lange A, Heider J, Huntley SM, Higgs PI.** 2010. Global transcriptome analysis of spore formation in *Myxococcus xanthus* reveals a locus necessary for cell differentiation. BMC Genomics **11:**264. http://dx.doi.org/10.1186/1471-2164-11-264.
9. **Harris BZ, Kaiser D, Singer M.** 1998. The guanosine nucleotide (p)ppGpp initiates development and A-factor production *in Myxococcus xanthus*. Genes Dev **12:**1022–1035. http://dx.doi.org/10.1101/gad.12.7.1022.
10. **Manoil C, Kaiser D.** 1980. Accumulation of guanosine tetraphosphate and guanosine pentaphosphate in *Myxococcus xanthus* during starvation and myxospore formation. J Bacteriol **141:**297–304.
11. **Manoil C, Kaiser D.** 1980. Guanosine pentaphosphate and guanosine tetraphosphate accumulation and induction of *Myxococcus xanthus* fruiting body development. J Bacteriol **141:**305–315.

12. **Singer M, Kaiser D.** 1995. Ectopic production of guanosine penta- and tetraphosphate can initiate early developmental gene expression in *Myxococcus xanthus*. Genes Dev **9:**1633–1644. http://dx.doi.org/10.1101/gad.9.13.1633.
13. **Kuspa A, Plamann L, Kaiser D.** 1992. Identification of heat-stable A-factor from *Myxococcus xanthus*. J Bacteriol **174:**3319–3326.
14. **Kuspa A, Plamann L, Kaiser D.** 1992. A-signalling and the cell density requirement for *Myxococcus xanthus* development. J Bacteriol **174:**7360–7369.
15. **LaRossa R, Kuner J, Hagen D, Manoil C, Kaiser D.** 1983. Developmental cell interactions of *Myxococcus xanthus*: analysis of mutants. J Bacteriol **153:**1394–1404.
16. **Plamann L, Kuspa A, Kaiser D.** 1992. Proteins that rescue A-signal-defective mutants of *Myxococcus xanthus*. J Bacteriol **174:**3311–3318.
17. **Hagen TJ, Shimkets LJ.** 1990. Nucleotide sequence and transcriptional products of the *csg* locus of *Myxococcus xanthus*. J Bacteriol **172:**15–23.
18. **Jelsbak L, Sogaard-Andersen L.** 2003. Cell behavior and cell-cell communication during fruiting body morphogenesis in *Myxococcus xanthus*. J Microbiol Methods **55:**829–839. http://dx.doi.org/10.1016/j.mimet.2003.08.007.
19. **Kim SK, Kaiser D.** 1990. C-factor: a cell-cell signaling protein required for fruiting body morphogenesis of *M. xanthus*. Cell **61:**19–26. http://dx.doi.org/10.1016/0092-8674(90)90211-V.
20. **Kim SK, Kaiser D.** 1990. Cell motility is required for the transmission of C-factor, an intercellular signal that coordinates fruiting body morphogenesis of *Myxococcus xanthus*. Genes Dev **4:**896–904. http://dx.doi.org/10.1101/gad.4.6.896.
21. **Kim SK, Kaiser D.** 1990. Purification and properties of *Myxococcus xanthus* C-factor, an intercellular signaling protein. Proc Natl Acad Sci U S A **87:**3635–3639. http://dx.doi.org/10.1073/pnas.87.10.3635.
22. **Søgaard-Andersen L, Slack FJ, Kimsey H, Kaiser D.** 1996. Intercellular C-signaling in *Myxococcus xanthus* involves a branched signal transduction pathway. Genes Dev **10:**740–754. http://dx.doi.org/10.1101/gad.10.6.740.
23. **Kaiser D.** 1998. How and why myxobacteria talk to each other. Curr Opin Microbiol **1:**663–668. http://dx.doi.org/10.1016/S1369-5274(98)80112-4.
24. **Kaiser D.** 2004. Signaling in myxobacteria. Annu Rev Microbiol **58:**75–98. http://dx.doi.org/10.1146/annurev.micro.58.030603.123620.
25. **Kaplan HB.** 1991. Cell-cell interactions that direct fruiting body development in *Myxococcus xanthus*. Curr Opin Genet Dev **1:**363–369. http://dx.doi.org/10.1016/S0959-437X(05)80301-6.
26. **Shimkets LJ.** 1999. Intercellular signaling during fruiting-body development of *Myxococcus xanthus*. Annu Rev Microbiol **53:**525–549. http://dx.doi.org/10.1146/annurev.micro.53.1.525.
27. **Søgaard-Andersen L.** 2004. Cell polarity, intercellular signalling and morphogenetic cell movements in *Myxococcus xanthus*. Curr Opin Microbiol **7:**587–593. http://dx.doi.org/10.1016/j.mib.2004.10.004.
28. **Goldman BS, Nierman WC, Kaiser D, Slater SC, Durkin AS, Eisen JA, Ronning CM, Barbazuk WB, Blanchard M, Field C, Halling C, Hinkle G, Iartchuk O, Kim HS, Mackenzie C, Madupu R, Miller N, Shvartsbeyn A, Sullivan SA, Vaudin M, Wiegand R, Kaplan HB.** 2006. Evolution of sensory complexity recorded in a myxobacterial genome. Proc Natl Acad Sci U S A **103:**15200–15205. http://dx.doi.org/10.1073/pnas.0607335103.
29. **Shi X, Wegener-Feldbrugge S, Huntley S, Hamann N, Hedderich R, Sogaard-Andersen L.** 2008. Bioinformatics and experimental analysis of proteins of two-component systems in *Myxococcus xanthus*. J Bacteriol **190:**613–624. http://dx.doi.org/10.1128/JB.01502-07.
30. **Whitworth DE, Millard A, Hodgson DA, Hawkins PF.** 2008. Protein-protein interactions between two-component system transmitter and receiver domains of *Myxococcus xanthus*. Proteomics **8:**1839–1842. http://dx.doi.org/10.1002/pmic.200700544.
31. **Munoz-Dorado J, Higgs PI, Elias-Arnanz M.** 2014. Abundance and complexity of signalling mechanisms in myxobacteria, p 127–149. *In* Yang Z, Higgs PI (ed), Myxobacteria: genomics, cellular and molecular biology. Caister Academic Press, Norwich, United Kingdom.
32. **Rajagopalan R, Sarwar Z, Garza A, Kroos L.** 2014. Developmental gene regulation, p 105–126. *In* Yang Z, Higgs PI (ed), Myxobacteria: genomics, cellular and molecular biology. Caister Academic Press, Norwich, United Kingdom.
33. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. Annu Rev Biochem **69:**183–215. http://dx.doi.org/10.1146/annurev.biochem.69.1.183.
34. **Krell T, Lacal J, Busch A, Silva-Jiménez H, Guazzaroni M-E, Ramos JL.** 2010. Bacterial sensor kinases: diversity in the recognition of environmental signals. Annu Rev Microbiol **64:**539–559. http://dx.doi.org/10.1146/annurev.micro.112408.134054.
35. **Gao R, Mack TR, Stock AM.** 2007. Bacterial response regulators: versatile regulatory strategies from common domains. Trends Biochem Sci **32:**225–234. http://dx.doi.org/10.1016/j.tibs.2007.03.002.
36. **Galperin MY.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. J Bacteriol **188:**4169–4182. http://dx.doi.org/10.1128/JB.01887-05.
37. **Huynh TN, Stewart V.** 2011. Negative control in two-component signal transduction by transmitter phosphatase activity. Mol Microbiol **82:**275–286. http://dx.doi.org/10.1111/j.1365-2958.2011.07829.x.
38. **Sarwar Z, Garza AG.** 2012. The Nla6S protein of *Myxococcus xanthus* is the prototype for a new family of bacterial histidine kinases. FEMS Microbiol Lett **335:**86–94. http://dx.doi.org/10.1111/j.1574-6968.2012.02639.x.
39. **Grebe TW, Stock JB.** 1999. The histidine protein kinase superfamily. Adv Microb Physiol **41:**139–227. http://dx.doi.org/10.1016/S0065-2911(08)60167-8.
40. **Karniol B, Vierstra RD.** 2004. The HWE histidine kinases, a new family of bacterial two-component sensor kinases with potentially diverse roles in environmental signaling. J Bacteriol **186:**445–453. http://dx.doi.org/10.1128/JB.186.2.445-453.2004.
41. **Bush M, Dixon R.** 2012. The role of bacterial enhancer binding proteins as specialized activators of σ^{54}-dependent transcription. Microbiol Mol Biol Rev **76:**497–529. http://dx.doi.org/10.1128/MMBR.00006-12.
42. **Caberoy NB, Welch RD, Jakobsen JS, Slater SC, Garza AG.** 2003. Global mutational analysis of NtrC-like activators in *Myxococcus xanthus*: identifying activator mutants defective for motility and fruiting body development. J Bacteriol **185:**6083–6094. http://dx.doi.org/10.1128/JB.185.20.6083-6094.2003.
43. **Giglio KM, Caberoy N, Suen G, Kaiser D, Garza AG.** 2011. A cascade of coregulating enhancer binding proteins initiates and propagates a multicellular developmental program. Proc Natl Acad Sci U S A **108:**E431–E439. http://dx.doi.org/10.1073/pnas.1105876108.
44. **Barrios H, Valderrama B, Morett E.** 1999. Compilation and analysis of σ^{54}-dependent promoter sequences. Nucleic Acids Res **27:**4305–4313. http://dx.doi.org/10.1093/nar/27.22.4305.
45. **Minchin SD, Austin S, Dixon RA.** 1988. The role of activator binding sites in transcriptional control of the divergently transcribed *nifF* and *nifLA* promoters from *Klebsiella pneumoniae*. Mol Microbiol **2:**433–442. http://dx.doi.org/10.1111/j.1365-2958.1988.tb00049.x.
46. **Morett E, Buck M.** 1988. NifA-dependent *in vivo* protection demonstrates that the upstream activator sequence of nif promoters is a protein binding site. Proc Natl Acad Sci U S A **85:**9401–9405. http://dx.doi.org/10.1073/pnas.85.24.9401.
47. **Reitzer LJ, Magasanik B.** 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell **45:**785–792. http://dx.doi.org/10.1016/0092-8674(86)90553-2.
48. **Popham DL, Szeto D, Keener J, Kustu S.** 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science **243:**629–635. http://dx.doi.org/10.1126/science.2563595.
49. **Sasse-Dwight S, Gralla JD.** 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ^{54}. Cell **62:**945–954. http://dx.doi.org/10.1016/0092-8674(90)90269-K.
50. **Wedel A, Kustu S.** 1995. The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. Genes Dev **9:**2042–2052. http://dx.doi.org/10.1101/gad.9.16.2042.
51. **Jagadeesan S, Mann P, Schink CW, Higgs PI.** 2009. A novel "four-component" two-component signal transduction mechanism regulates developmental progression in *Myxococcus xanthus*. J Biol Chem **284:**21435–21445. http://dx.doi.org/10.1074/jbc.M109.033415.
52. **Weiss V, Kramer G, Dunnebier T, Flotho A.** 2002. Mechanism of regulation of the bifunctional histidine kinase NtrB in *Escherichia coli*. J Mol Microbiol Biotechnol **4:**229–233.
53. **Willett JW, Kirby JR.** 2011. CrdS and CrdA comprise a two-component system that is cooperatively regulated by the Che3 chemosensory system in *Myxococcus xanthus*. mBio **2:**e00110-11. http://dx.doi.org/10.1128/mBio.00110-11.
54. **Huynh TN, Noriega CE, Stewart V.** 2010. Conserved mechanism for sensor phosphatase control of two-component signaling revealed in the nitrate sensor NarX. Proc Natl Acad Sci U S A **107:**21140–21145. http://dx.doi.org/10.1073/pnas.1013081107.

55. **Willett JW, Kirby JR.** 2012. Genetic and biochemical dissection of a HisKA domain identifies residues required exclusively for kinase and phosphatase activities. PLoS Genet **8:**e1003084. http://dx.doi.org/10.1371/journal.pgen.1003084.
56. **Hiratsu K, Nakata A, Shinagawa H, Makino K.** 1995. Autophosphorylation and activation of transcriptional activator PhoB of *Escherichia coli* by acetyl phosphate *in vitro*. Gene **161:**7–10. http://dx.doi.org/10.1016/0378-1119(95)00259-9.
57. **Lukat GS, Mccleary WR, Stock AM, Stock JB.** 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phosphodonors. Proc Natl Acad Sci U S A **89:**718–722. http://dx.doi.org/10.1073/pnas.89.2.718.
58. **McCleary WR, Stock JB.** 1994. Acetyl phosphate and the activation of two-component response regulators. J Biol Chem **269:**31567–31572.
59. **Giglio KM, Zhu C, Klunder C, Kummer S, Garza AG.** 2015. The enhancer binding protein Nla6 regulates developmental genes that are important for *Myxococcus xanthus* sporulation. J Bacteriol **197:**1276–1287. http://dx.doi.org/10.1128/JB.02408-14.
60. **Holkenbrink C, Hoiczyk E, Kahnt J, Higgs PI.** 2014. Synthesis and assembly of a novel glycan layer in *Myxococcus xanthus* spores. J Biol Chem **289:**32364–32378. http://dx.doi.org/10.1074/jbc.M114.595504.
61. **Müller FD, Schink CW, Hoiczyk E, Cserti E, Higgs PI.** 2012. Spore formation in *Myxococcus xanthus* is tied to cytoskeleton functions and polysaccharide spore coat deposition. Mol Microbiol **83:**486–505. http://dx.doi.org/10.1111/j.1365-2958.2011.07944.x.
62. **Wartel M, Ducret A, Thutupalli S, Czerwinski F, Le Gall AV, Mauriello EM, Bergam P, Brun YV, Shaevitz J, Mignot T.** 2013. A versatile class of cell surface directional motors gives rise to gliding motility and sporulation in *Myxococcus xanthus*. PLoS Biol **11:**e1001728. http://dx.doi.org/10.1371/journal.pbio.1001728.
63. **Licking E, Gorski L, Kaiser D.** 2000. A common step for changing cell shape in fruiting body and starvation-independent sporulation of *Myxococcus xanthus*. J Bacteriol **182:**3553–3558. http://dx.doi.org/10.1128/JB.182.12.3553-3558.2000.
64. **Fukushima T, Yamamoto H, Atrih A, Foster SJ, Sekiguchi J.** 2002. A polysaccharide deacetylase gene (*pdaA*) is required for germination and for production of muramic delta-lactam residues in the spore cortex of *Bacillus subtilis*. J Bacteriol **184:**6007–6015. http://dx.doi.org/10.1128/JB.184.21.6007-6015.2002.
65. **Vollmer W, Tomasz A.** 2000. The *pgdA* gene encodes a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. J Biol Chem **275:**20496–20501. http://dx.doi.org/10.1074/jbc.M910189199.
66. **Pan HW, Liu H, Liu T, Li CY, Li ZF, Cai K, Zhang CY, Zhang Y, Hu W, Wu ZH, Li YZ.** 2009. Seawater-regulated genes for two-component systems and outer membrane proteins in *Myxococcus*. J Bacteriol **191:**2102–2111. http://dx.doi.org/10.1128/JB.01556-08.
67. **Sarwar Z, Garza AG.** 2012. The Nla28S/Nla28 two-component signal transduction system regulates sporulation in *Myxococcus xanthus*. J Bacteriol **194:**4698–4708. http://dx.doi.org/10.1128/JB.00225-12.
68. **Cho K, Zusman DR.** 1999. AsgD, a new two-component regulator required for A-signalling and nutrient sensing during early development of *Myxococcus xanthus*. Mol Microbiol **34:**268–281. http://dx.doi.org/10.1046/j.1365-2958.1999.01594.x.
69. **Sun H, Shi W.** 2001. Analyses of *mrp* genes during *Myxococcus xanthus* development. J Bacteriol **183:**6733–6739. http://dx.doi.org/10.1128/JB.183.23.6733-6739.2001.
70. **Sun H, Shi W.** 2001. Genetic studies of *mrp*, a locus essential for cellular aggregation and sporulation of *Myxococcus xanthus*. J Bacteriol **183:**4786–4795. http://dx.doi.org/10.1128/JB.183.16.4786-4795.2001.
71. **Boysen A, Ellehauge E, Julien B, Sogaard-Andersen L.** 2002. The DevT protein stimulates synthesis of FruA, a signal transduction protein required for fruiting body morphogenesis in *Myxococcus xanthus*. J Bacteriol **184:**1540–1546. http://dx.doi.org/10.1128/JB.184.6.1540-1546.2002.
72. **Campbell A, Viswanathan P, Barrett T, Son B, Saha S, Kroos L.** 2015. Combinatorial regulation of the *dev* operon by MrpC2 and FruA during *Myxococcus xanthus* development. J Bacteriol **197:**240–251. http://dx.doi.org/10.1128/JB.02310-14.
73. **Mittal S, Kroos L.** 2009. A combination of unusual transcription factors binds cooperatively to control *Myxococcus xanthus* developmental gene expression. Proc Natl Acad Sci U S A **106:**1965–1970. http://dx.doi.org/10.1073/pnas.0808516106.
74. **Viswanathan P, Murphy K, Julien B, Garza AG, Kroos L.** 2007. Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. J Bacteriol **189:**3738–3750. http://dx.doi.org/10.1128/JB.00187-07.
75. **Gronewold TM, Kaiser D.** 2001. The *act* operon controls the level and time of C-signal production for *Myxococcus xanthus* development. Mol Microbiol **40:**744–756. http://dx.doi.org/10.1046/j.1365-2958.2001.02428.x.
76. **Gronewold TM, Kaiser D.** 2007. Mutations of the act promoter in *Myxococcus xanthus*. J Bacteriol **189:**1836–1844. http://dx.doi.org/10.1128/JB.01618-06.
77. **Gronewold TM, Kaiser D.** 2002. *act* operon control of developmental gene expression in *Myxococcus xanthus*. J Bacteriol **184:**1172–1179. http://dx.doi.org/10.1128/jb.184.4.1172-1179.2002.
78. **Dubnau D, Losick R.** 2006. Bistability in bacteria. Mol Microbiol **61:**564–572. http://dx.doi.org/10.1111/j.1365-2958.2006.05249.x.
79. **Smits WK, Kuipers OP, Veening JW.** 2006. Phenotypic variation in bacteria: the role of feedback regulation. Nat Rev Microbiol **4:**259–271. http://dx.doi.org/10.1038/nrmicro1381.
80. **Lobedanz S, Sogaard-Andersen L.** 2003. Identification of the C-signal, a contact-dependent morphogen coordinating multiple developmental responses in *Myxococcus xanthus*. Genes Dev **17:**2151–2161. http://dx.doi.org/10.1101/gad.274203.
81. **Rolbetzki A, Ammon M, Jakovljevic V, Konovalova A, Sogaard-Andersen L.** 2008. Regulated secretion of a protease activates intercellular signaling during fruiting body formation in *M. xanthus*. Dev Cell **15:**627–634. http://dx.doi.org/10.1016/j.devcel.2008.08.002.
82. **Kim SK, Kaiser D.** 1991. C-factor has distinct aggregation and sporulation thresholds during *Myxococcus* development. J Bacteriol **173:**1722–1728.
83. **Li S, Lee BU, Shimkets LJ.** 1992. *csgA* expression entrains *Myxococcus* xanthus development. Genes Dev **6:**401–410. http://dx.doi.org/10.1101/gad.6.3.401.
84. **Jelsbak L, Givskov M, Kaiser D.** 2005. Enhancer-binding proteins with a forkhead-associated domain and the σ^{54} regulon in Myxococcus *xanthus* fruiting body development. Proc Natl Acad Sci U S A **102:**3010–3015. http://dx.doi.org/10.1073/pnas.0409371102.
85. **Volz C, Kegler C, Muller R.** 2012. Enhancer binding proteins act as hetero-oligomers and link secondary metabolite production to myxococcal development, motility, and predation. Chem Biol **19:**1447–1459. http://dx.doi.org/10.1016/j.chembiol.2012.09.010.
86. **Ellehauge E, Norregaard-Madsen M, Sogaard-Andersen L.** 1998. The FruA signal transduction protein provides a checkpoint for the temporal co-ordination of intercellular signals in *Myxococcus xanthus* development. Mol Microbiol **30:**807–817. http://dx.doi.org/10.1046/j.1365-2958.1998.01113.x.
87. **Ogawa M, Fujitani S, Mao X, Inouye S, Komano T.** 1996. FruA, a putative transcription factor essential for the development of *Myxococcus xanthus*. Mol Microbiol **22:**757–767. http://dx.doi.org/10.1046/j.1365-2958.1996.d01-1725.x.
88. **Inouye S, Nariya H.** 2008. Dual regulation with Ser/Thr kinase cascade and a His/Asp TCS in *Myxococcus xanthus*. Adv Exp Med Biol **631:**111–121. http://dx.doi.org/10.1007/978-0-387-78885-2_7.
89. **Kaiser D, Robinson M, Kroos L.** 2010. Myxobacteria, polarity, and multicellular morphogenesis. Cold Spring Harb Perspect Biol **2:**a000380.
90. **Ueki T, Inouye S.** 2005. Identification of a gene involved in polysaccharide export as a transcription target of FruA, an essential factor for *Myxococcus xanthus* development. J Biol Chem **280:**32279–32284. http://dx.doi.org/10.1074/jbc.M507191200.
91. **Mittal S, Kroos L.** 2009. Combinatorial regulation by a novel arrangement of FruA and MrpC2 transcription factors during *Myxococcus xanthus* development. J Bacteriol **191:**2753–2763. http://dx.doi.org/10.1128/JB.01818-08.
92. **Cho K, Zusman DR.** 1999. Sporulation timing in *Myxococcus xanthus* is controlled by the *espAB* locus. Mol Microbiol **34:**714–725. http://dx.doi.org/10.1046/j.1365-2958.1999.01633.x.
93. **Lee B, Higgs PI, Zusman DR, Cho K.** 2005. EspC is involved in controlling the timing of development in *Myxococcus xanthus*. J Bacteriol **187:**5029–5031. http://dx.doi.org/10.1128/JB.187.14.5029-5031.2005.
94. **Schramm A, Lee B, Higgs PI.** 2012. Intra- and interprotein phosphorylation between two-hybrid histidine kinases controls *Myxococcus xanthus* developmental progression. J Biol Chem **287:**25060–25072. http://dx.doi.org/10.1074/jbc.M112.387241.
95. **Stein EA, Cho K, Higgs PI, Zusman DR.** 2006. Two Ser/Thr protein kinases essential for efficient aggregation and spore morphogenesis in

Myxococcus xanthus. Mol Microbiol **60:**1414–1431. http://dx.doi.org/10.1111/j.1365-2958.2006.05195.x.
96. **Durocher D, Henckel J, Fersht AR, Jackson SP.** 1999. The FHA domain is a modular phosphopeptide recognition motif. Mol Cell **4:**387–394. http://dx.doi.org/10.1016/S1097-2765(00)80340-8.
97. **Higgs PI, Jagadeesan S, Mann P, Zusman DR.** 2008. EspA, an orphan hybrid histidine protein kinase, regulates the timing of expression of key developmental proteins of *Myxococcus xanthus*. J Bacteriol **190:**4416–4426. http://dx.doi.org/10.1128/JB.00265-08.
98. **O'Connor KA, Zusman DR.** 1991. Analysis of *Myxococcus xanthus* cell types by two-dimensional polyacrylamide gel electrophoresis. J Bacteriol **173:**3334–3341.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

MINIREVIEW

Yersinia Type III Secretion System Master Regulator LcrF

Leah Schwiesow,[a] Hanh Lam,[b] Petra Dersch,[c] Victoria Auerbuch[b]

Department of Molecular, Cell, and Developmental Biology[a] and Department of Microbiology and Environmental Toxicology,[b] University of California, Santa Cruz, Santa Cruz, California, USA; Department of Molecular Infection Biology, Helmholtz Center for Infection Research, Braunschweig, Lower Saxony, Germany[c]

Many Gram-negative pathogens express a type III secretion (T3SS) system to enable growth and survival within a host. The three human-pathogenic *Yersinia* species, *Y. pestis*, *Y. pseudotuberculosis*, and *Y. enterocolitica*, encode the Ysc T3SS, whose expression is controlled by an AraC-like master regulator called LcrF. In this review, we discuss LcrF structure and function as well as the environmental cues and pathways known to regulate LcrF expression. Similarities and differences in binding motifs and modes of action between LcrF and the *Pseudomonas aeruginosa* homolog ExsA are summarized. In addition, we present a new bioinformatics analysis that identifies putative LcrF binding sites within *Yersinia* target gene promoters.

There are three *Yersinia* species pathogenic to humans. *Y. pestis* is the causative agent of bubonic and pneumonic plague and is transmitted through a flea vector or through airborne transmission from one mammalian host to another (1). In contrast, the enteropathogenic *Yersinia* species *Y. enterocolitica* and *Y. pseudotuberculosis* grow in the environment but can be transmitted to mammalian hosts through ingestion of contaminated food or water (1). *Y. enterocolitica* and *Y. pseudotuberculosis* cause typically self-limiting mesenteric lymphadenitis or gastroenteritis in otherwise healthy individuals but can cause a serious blood-borne infection in people with iron overload disorders such as hereditary hemochromatosis (2). In addition, sequelae following enteropathogenic *Yersinia* infection, such as erythema nodosum and reactive arthritis, have also been reported (3–5).

Human-pathogenic *Yersinia* species share a virulence plasmid, called pCD1 in *Y. pestis* and pYV in enteropathogenic yersiniae, encoding the Ysc type III secretion system (T3SS) essential for causing disease (6). These 70-kb plasmids carry dozens of T3SS structural genes and encode five or six T3SS effector proteins called Yops and their dedicated chaperones, as well as genes encoding proteins involved in regulating expression and function of the T3SS. One of these regulatory proteins, LcrF, serves as the *Yersinia* T3SS master regulator, controlling transcription of a large number of plasmid-borne genes. Several environmental cues influence expression of LcrF itself, possibly enabling *Yersinia* to control T3SS expression during transitions from one niche to another. Recent reviews have highlighted important advances in our understanding of T3SS structure and modulation of the innate immune response by T3SS effector proteins (7). In this review, we focus on factors controlling LcrF expression, the target genes LcrF regulates, and how LcrF activity influences *Yersinia* pathogenesis.

LcrF HISTORY, STRUCTURE, AND FUNCTION

It has long been appreciated that human-pathogenic *Yersinia* carrying T3SS genes requires millimolar concentrations of calcium to grow at 37°C, and this phenomenon was termed the low-calcium response, or Lcr (8, 9). The absence of calcium, in combination with a shift to 37°C, triggers secretion of T3SS effector proteins, mimicking the effect of host cell contact. *Yersinia* undergoes growth arrest during active type III secretion, explaining why calcium is required for growth at 37°C. Mechanistic explanations for how calcium ions regulate Yop secretion and why type III secretion is associated with cessation of growth *in vitro* remain unclear. However, a number of T3SS genes were originally named for the low-calcium response, as mutations in these genes were shown to alter the Lcr phenotype (10, 11).

lcrF was first identified in *Y. pestis* by Goguen and colleagues as a gene required for thermal induction of several pCD1 genes (12). Using a similar approach, VirF, the *Y. enterocolitica* LcrF homolog, was discovered (13). Finally, while sequence analysis of *Y. pestis*, *Y. enterocolitica*, and *Y. pseudotuberculosis* virulence plasmids revealed evidence of several rearrangements, the low-calcium response at 37°C was found to be highly conserved in all three species, suggesting the presence of LcrF in *Y. pseudotuberculosis* as well (14). In this review, we refer to the regulator as LcrF unless specifically referring to the *Y. enterocolitica* VirF homolog.

LcrF is a 30-kDa AraC-like protein that shares homology with AraC in its carboxy-terminal DNA binding region (15). AraC is well known for its role in *Escherichia coli* as a DNA binding transcriptional regulator (reviewed in reference 16). Early on, *Y. pestis* LcrF and *Y. enterocolitica* VirF were shown by the use of gel shift assays to bind directly to sequences in the *yopE* and *yopH* promoters (17). The amino-terminal domains for AraC-like proteins have been shown to be involved in self-association. Additionally, the amino-terminal domain of several AraC-like proteins binds cofactors that influence the ability of the protein to regulate transcription (16). For example, binding of the arabinose cofactor to *E. coli* AraC induces a conformational change, allowing AraC to activate transcription (16). It is thought that LcrF exists as a dimer in solution through the self-association of its amino-terminal domain; however, unlike that of *E. coli* AraC, the amino-terminal domain of LcrF has not been shown to bind additional cofactors.

As in all other AraC family transcriptional regulators, the carboxy-terminal DNA binding region of LcrF contains two helix-turn-helix (HTH) domains (18). Like all HTH domains, the recognition helix binds specific DNA residues within the major

Accepted manuscript posted online 7 December 2015

Citation Schwiesow L, Lam H, Dersch P, Auerbuch V. 2016. *Yersinia* type III secretion system master regulator LcrF. J Bacteriol 198:604–614. doi:10.1128/JB.00686-15.

Editor: W. Margolin

Address correspondence to Petra Dersch, petra.dersch@helmholtz-hzi.de, or Victoria Auerbuch, vastone@ucsc.edu.

groove (19). DNA binding sites of LcrF and its homologs have been experimentally investigated in a number of studies (15–17, 20, 21). The first study, using DNase protection assays performed on promoters of genes activated by LcrF, suggested a common DNA binding motif, TTTTaGYcTgTat (capital letters represent more highly conserved residues, and Y stands for C or T) (17). Wattiau and Cornelis identified this "half-site" upstream of several T3SS genes. However, these proposed LcrF binding sites were highly variable in terms of distance from the transcriptional start site, directionality, and distance of the half-sites from each other (17). In a more recent analysis, King et al. observed that the carboxy-terminal domain of LcrF is nearly identical to the DNA binding region of the homologous AraC-like master regulator ExsA of the *Pseudomonas aeruginosa* T3SS (18). The authors further showed that *Y. pestis* LcrF binds and activates ExsA-dependent promoters in *P. aeruginosa*. Similarly, ExsA was able to induce expression of T3SS genes in *Y. pestis* in the absence of LcrF (18). LcrF and ExsA were shown to interact with a common nucleotide sequence motif (AaAAAnwnMygrCynnnmYTGyaAk), which is also recognized by activators of T3SS genes from *Photorhabdus luminescens*, *Aeromonas hydrophilus*, and *Vibrio parahaemolyticus* (W stands for A or T, M for A or C, Y for C or T, R for A or G, and K for G or T [with uppercase letters representing more highly conserved residues]) (18, 21, 22).

To reconcile these two dissimilarly presented LcrF consensus binding sites, we attempted to align the promoter regions of genes known to be controlled by LcrF. First, all previously identified LcrF binding site sequences from the *virA*, *virB*, *virC*, *yopE*, *lcrG*, and *yopH* promoters, as well as known ExsA binding sites from *P. aeruginosa*, were input into MEME motif discovery to identify a consensus motif (23). The resulting motif was subsequently used to search the virulence plasmids of the three human-pathogenic *Yersinia* species to identify all putative LcrF binding sites using FIMO (Find Individual Motif Occurrences), where motifs were called with a *P* value threshold of <0.0001 (24). Alignment of LcrF binding sites of selected genes (Fig. 1A) shows a 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motif located 20 to 21 bp upstream of the predicted −10 TATA box for most genes. This consensus motif is similar to that of ExsA described above. Sequences containing these motifs 20 to 21 bp upstream of the −10 TATA box overlap well the regions found experimentally to bind LcrF or ExsA (Fig. 1A, solid underlines) (15–18, 20, 21). In the case of the *yopH* and *sycE* genes, a second 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motif was found further upstream and coincided with regions previously shown to be weakly bound by LcrF (Fig. 1A, dashed underlines) (17).

A closer look at the *yopE* promoter revealed surprising features that required clarification. A putative LcrF binding site within the *yopE* promoter was not located 20 to 21 bp upstream of a −10 TATA box (Fig. 1B), as are the majority of putative LcrF sites, but overlapped the annotated translational start site of *yopE* (Fig. 1A). The −10 region of *yopE*, however, may be located 30 bp upstream of the translational start site, which is 12 bp upstream of the transcriptional start site (black arrow in Fig. 1B) (25). Interestingly, Wattiau and Cornelis identified two different regions further upstream from *Y. enterocolitica yopE* that were protected by VirF during DNase I footprinting (Fig. 1B in blue boxes) (17). These regions contain sequences that are the reverse complement of the 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motif. Because *sycE* and *yopE* are located in close proximity but are transcribed in opposite directions, it is possible that the two protected regions upstream of *yopE*, one strongly bound by VirF (underlined by a solid line in Fig. 1B) and the other weakly bound (underlined by dashed lines in Fig. 1B) (17), are in fact LcrF binding motifs on the reverse DNA strand, belonging to the *sycE* promoter. This LcrF binding site phenomenon of two adjacent, coregulated, but divergent genes can also be seen in the case of *virA* and *virB* (blue solid-line box in Fig. 1A). It is worth noting that some AraC-like activators require binding sites downstream of the −10 region (26, 27). For example, in the *E. coli* AraC-like activator Rns, it was found that the inverted Rns motifs upstream of target gene −10 regions and the motifs downstream of −10 regions function in synergy (26, 27). Therefore, it is possible that both the putative LcrF binding sites far upstream and those downstream of the *yopE* transcriptional start site are required for activation of the *yopE* promoter. Interestingly, *yopE* was reported to be transcribed strongly under T3SS-inducing conditions, while *sycE* was minimally transcribed (28). Because of this departure from the known characteristics of other LcrF binding sites, whether the putative LcrF binding sites within the *yopE-sycE* promoter regions are functional and, if so, how they function remain to be determined.

Putative LcrF binding sites appear in a number of locations in the virulence plasmids of *Yersinia*. As shown in Fig. 1A, a 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motif was found within the *sycE*, *yopO*, *yadA*, *yopK*, *sycH*, *yscH*, and *yscC* promoters. The transcriptional dependence of *yadA* and *yopK* (homologous to *yopQ* in *Y. enterocolitica*) on LcrF and VirF was previously demonstrated in transcriptional reporter assays in *Y. pseudotuberculosis* (29) and *Y. enterocolitica* (20), respectively. It is worth mentioning that we could not find LcrF binding sites upstream of two secreted effectors, *yopJ* and *yopM*, unless the threshold for motif calling was much lower (data not shown). Note also that *yscC* and *yopK* (from *Y. enterocolitica* and *Y. pestis*) do not seem to have a consensus TATA box following the TGANA motif, although these two genes are expressed under T3SS-inducing conditions (28). In contrast, the two LcrF binding sites upstream of *virA* and *virB* (Fig. 1A) have a TATA box at a proper distance but exhibit a weak dependence on LcrF (20).

Our analysis identified three putative LcrF binding sites within *virC* (*yscA* to *yscL*), upstream of *yscA*, *yscC*, and *yscH* (Fig. 1A). Transcriptional regulation of this region seems to be complex. Michiels et al. proposed the presence of a single operon from *yscA* to *yscL* in *Y. enterocolitica* (30). However, Haddix and Straley demonstrated the presence of at least two operons within the *Y. pestis virC* region, one starting upstream of *yscA* and the other starting upstream of *yscF* (31). The exact operon structure of the *virC* locus, as well as the LcrF binding sites important for *virC* gene expression, remains to be clarified.

Recent findings have shown that ExsA/LcrF-dependent promoters are σ^{70} promoters (32). The conserved TGANA sequence within the broader LcrF binding motif was previously mistaken for a −35 box and is thought to bind one of the two monomers of LcrF or ExsA dimers (20). Notably, the distance between the −10 box and the TGANA motif of LcrF-and-ExsA-dependent genes is 21 to 22 nucleotides (nt), while the spacing seen in typical σ^{70}-dependent promoters is 17 nucleotides between the −10 and −35 boxes (32). Decreasing the distance between the −10 box and the TGANA motif from 21 or 22 nucleotides to 17 nucleotides in the ExsA-dependent *exoT* and *exsD* genes abolishes transcriptional activation of these genes, suggesting that the 21-to-22-nucleotide

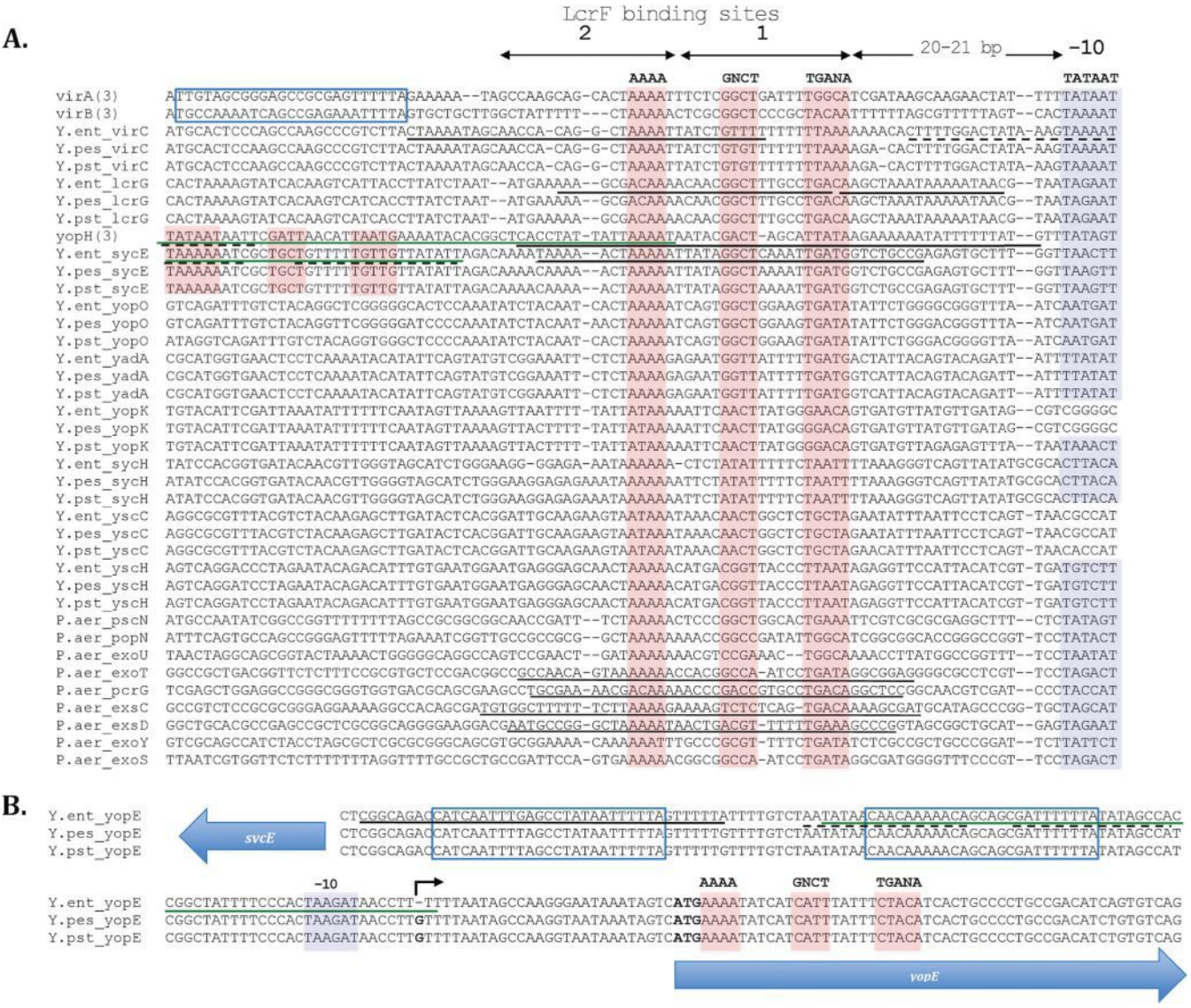


FIG 1 Alignment of verified and putative LcrF and ExsA binding sites within target gene promoters. (A) Promoter regions of *Yersinia* and *Pseudomonas* genes controlled by LcrF and ExsA, respectively, were aligned using SeaView (81). The sequences are from *Y. enterocolitica* 8081 (NC 008791), *Y. pestis* CO92 (NC 003131), *Y. pseudotuberculosis* IP 32953 (NC 006153), *P. aeruginosa* (NC 002516), and *P. aeruginosa* UCBPP PA14 (NC 008463). Predicted −10 regions are highlighted in blue. Identified 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ consensus sites containing three conserved regions are highlighted in red, and the conserved nucleotides at each position of the motif are denoted in bold above the alignment, with uppercase letters denoting highly conserved residues and lowercase letters denoting more-degenerate residues. LcrF binding sites 1 and 2 are indicated by arrows. The sequences of the *virA*, *virB*, and *yopH* promoters from all three *Yersinia* species are identical, and thus a single sequence is shown. *virA* and *virB* are proximal but are divergently encoded. Thus, we propose that what appears to be an inverted 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motif upstream of the *virA* and *virB* promoters (blue solid-line box) actually belongs to the divergent *virB* and *virA* promoters, respectively. *yopH* and *sycE* have two tandem 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motifs (both highlighted in red), which overlap the protected regions identified by Wattiau and Cornelis (17). (B) Sequences upstream of *Yersinia yopE* were aligned using SeaView. Two regions containing inverted 5′-AAAA-N_6-GNCT-N_5-TGANA-3′ motifs which may belong to *sycE* are outlined in blue. The putative translational *yopE* start site is denoted in bold. (A and B) YtxR binding sites upstream of *yopH* (A) and *yopE* (B) are underlined in green (38). Regions experimentally found to be strongly bound by LcrF/VirF or ExsA are underlined with black solid lines, and those found to be weakly bound are underlined with black dashed lines (17). Identified putative LcrF binding sites are highlighted in red and the conserved nucleotides denoted in bold above the alignment, with uppercase letters denoting highly conserved residues and lowercase letters denoting more-degenerate residues.

spacing ensures that activation of these promoters does not occur without ExsA binding (32).

Despite the resemblance in the consensus binding sequences, the oligomeric states of LcrF and ExsA during DNA binding as well as their binding properties are distinct, leading to differences in the DNA binding specificity and kinetics of transcriptional activation. ExsA is predominantly monomeric in solution, and two molecules are sequentially recruited to target promoters and generate two higher-order DNA-protein complexes to activate T3SS genes, whereas LcrF is dimeric and its presence results in the formation of only one large higher-order DNA-protein complex (18, 33). In addition, LcrF-induced promoter bending is more pronounced and may account for its overall (2.5-fold to 20-fold) higher activator activity relative to that of ExsA (18, 22). Lower

basal activity of ExsA could be compensated by its positive autoregulatory feedback loop, which leads to a rapid increase of ExsA expression under inducing conditions (34). While mutations in critical nucleotides within the major groove disrupted both LcrF and ExsA DNA binding, LcrF, but not ExsA, was able to tolerate certain mutations within the consensus binding site, presumably because LcrF binds DNA as a preformed dimer and not in the sequential, ordered manner seen with ExsA monomers (18). Indeed, an LcrF mutant unable to dimerize was shown to bind to each half-site but, similarly to ExsA, was more sensitive to mutations within the binding site (18).

REGULATION OF LcrF TARGET GENES

To date, LcrF/VirF is the only characterized transcriptional activator of T3SS genes in pathogenic *Yersinia*, and all known LcrF target genes are carried on pCD1/pYV. Several groups have shown transcriptional activation or direct binding of LcrF to the promoter regions of the *yopE*, *yopH*, *yadA*, and *ylpA* genes as well as of the *yopBD-lcrGVH* and *virC* operons, which encode T3SS effector, regulatory, and structural proteins as well as the YadA adhesin and the YlpA lipoprotein (17, 20, 29, 30, 35–37).

Several reports have suggested that regulation of T3SS genes can be mediated by proteins antagonizing LcrF (38, 39). Darwin and colleagues identified YtxR as a global transcriptional regulator that is conserved in all human-pathogenic *Yersinia* species. Overexpression of YtxR rendered *Yersinia* defective in secretion of Yops into the bacterial culture supernatant (38). Using a DNase footprinting approach, the authors showed that YtxR protected specific regions in the *yopE-sycE* and *yopH* promoters that overlapped known LcrF binding sites (denoted by green lines in Fig. 1A), suggesting that YtxR competes with LcrF for binding to T3SS gene promoters (38). While YtxR may introduce an additional layer of regulation to T3SS gene expression, the environmental conditions under which YtxR is expressed have yet to be elucidated (38).

Recently, Li et al. proposed another model in which the pYV-encoded regulatory protein LcrQ inhibits LcrF activity (39). *Y. pseudotuberculosis* LcrQ was first discovered as a gene required for calcium dependence at 37°C and was shown to be secreted by the T3SS, relieving repression of Yop expression when the concentration of LcrQ in the bacterial cytoplasm decreased as a result of active secretion (40, 41). The *Y. enterocolitica* LcrQ orthologs YscM1 and YscM2 were subsequently suggested to control Yop expression in combination with pYV-encoded factors other than VirF (42). Indeed, the T3SS YopD translocon protein and its LcrH/SycD chaperone are thought to cooperate with LcrQ or YscM1/YscM2 to negatively control T3SS gene expression (43–47). LcrQ has no obvious DNA binding domain and does not bind Yop promoter regions, as has been shown for LcrF and YtxR (39). In fact, it has been suggested that *Y. enterocolitica* YscM1 and YscM2, in cooperation with the YopD-LcrH complex, bind to the 5′ untranslated regions of Yop mRNAs to inhibit translation (46). However, Li et al. proposed that *Y. pseudotuberculosis* LcrQ inhibits LcrF activity until the T3SS is assembled and LcrQ is secreted out of the cell. The authors showed that overexpression of LcrF mimicked the secretion profiles of a Δ*lcrQ* mutant. This suggested that the LcrF/LcrQ ratio may be important for activating T3SS gene transcription, but they could not detect a direct interaction of LcrF and LcrQ (39). Moreover, previous data indicated that *Y. enterocolitica* YscM1 was not capable of inhibiting VirF activity on the *yopH* promoter in the absence of other pYV-encoded factors (42). Therefore, how LcrQ and YscM1/YscM2 repress T3SS gene expression in pathogenic *Yersinia* remains to be clarified.

TRANSCRIPTIONAL CONTROL OF LcrF

Thermoregulation of *lcrF* transcription and *ymoA*. There are several lines of evidence showing that transcription of *lcrF* is regulated by temperature. The first observations were made with *Y. enterocolitica*, in which induction of a *virF-cat* fusion, and a substantial increase in the abundance of the *virF* transcript, was detected upon temperature upshift (15, 48). A comprehensive expression analysis revealed that the *lcrF* gene of *Y. pseudotuberculosis* is transcribed from a σ^{70}-dependent promoter located upstream of the *yscW* gene (named *virG* in *Y. enterocolitica*), which in turn is located 124 bp upstream of the *lcrF* coding sequence (Fig. 2). Temperature-dependent *yscW-lcrF* transcription from this promoter is controlled by the nucleoid-associated YmoA protein, which shows homology to the Hha protein of *E. coli* (49). YmoA (for "*Y*ersinia modulator") was identified in a search for chromosomal insertion mutants of *Y. enterocolitica* transcribing *virF* and, hence, VirF-dependent *yop* and *yadA* genes at low temperatures (48). Mutations in *ymoA* led to increased *virF* and *yop* expression at 25°C and decreased expression at 37°C compared to that in the control strain, although *ymoA* mutants still induced expression of *virF* and *yop* upon temperature shift (48). Elevated expression of a *yscW-lacZ* fusion in a *ymoA*-deficient *Y. pseudotuberculosis* strain further suggested that YmoA represses *yscW-lcrF* transcription from a promoter located 264 nt upstream of the start codon of *yscW* (49) (Fig. 2). As YmoA dependency was lost when sequences downstream of the *yscW* transcriptional start site were deleted but was maintained when the regulatory region upstream of the *yscW* promoter was removed, it is assumed that YmoA influences *lcrF* expression via sequences located downstream of the *yscW* promoter (49). However, even high concentrations of purified YmoA homodimers were unable to interact specifically with the *yscW* regulatory region, indicating that an additional factor contributes to YmoA-mediated repression of *yscW-lcrF* transcription at moderate temperatures (15°C to 30°C) (49).

The small size of YmoA, its unusual high number of charged amino acid residues, and its influence on the fragility of the chromosomal DNA suggested that YmoA is a histone-like protein involved in chromosome structure similarly to histone-like nucleoid-structuring (H-NS) protein and controls *lcrF* expression through temperature-induced changes in DNA topology (48, 50, 51). Domains of members of the Hha/YmoA protein family have a striking similarity to the oligomerization domain of the H-NS nucleoid-structuring protein and its paralogs and were shown to interact specifically with different enterobacterial H-NS proteins (52, 53). Band shift analyses performed with *yscW* promoter fragments demonstrated that YmoA copurified with H-NS was able to interact specifically with sequences located downstream of the *yscW* promoter (Fig. 2) (49). Therefore, heterocomplex formation of YmoA with H-NS seems responsible for the thermoregulation of the *yscW-lcrF* operon. Interestingly, H-NS alone is also able to interact with *yscW* promoter fragments (49). In this context, it would be important to know how H-NS homodimers and H-NS/YmoA heterodimers differ in their abilities to interact with the *yscW* promoter region and how this influences expression of the *yscW-lcrF* operon in response to temperature. Several independent attempts to construct an *hns*-deficient mutant in *Yersinia* to

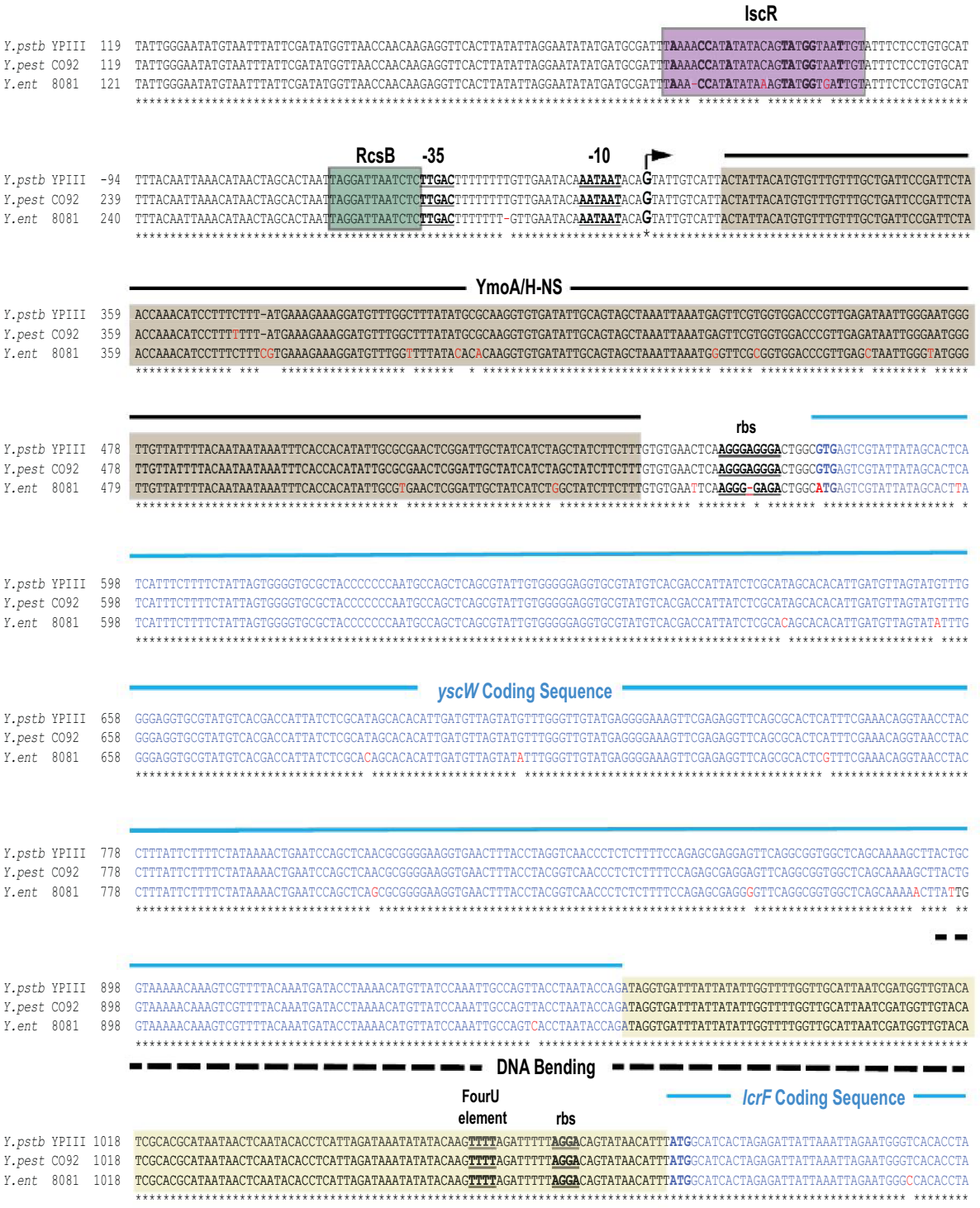


FIG 2 Regulatory elements encoded within the *yscW-lcrF* sequence. Nucleotide sequences of the *yscW-lcrF* promoter regions for *Y. pestis* CO92, *Y. enterocolitica* 8081, and *Y. pseudotuberculosis* YPIII were aligned using ClustalW2. Nucleotides whose sequences are not identical are marked in red, while the conserved nucleotides are indicated by asterisks. The identified binding sequence for IscR is marked in purple, and critical residues required for IscR binding within this motif are in bold (28, 64). The binding site for RcsB is marked in green (71). The region experimentally determined to be involved in YmoA/H-NS binding, marked in brown, is downstream of the transcriptional start site (49). The sequence encoding the RNA thermometer is marked in yellow and is encoded within the intergenic region between *yscW* and *lcrF* (49). The DNA bending region identified within the *yscW-lcrF* operon is denoted with a dotted line and is within the intergenic region between *yscW* and *lcrF* (59). The −10 and −35 boxes, the transcriptional start site, the ribosome binding sites (rbs) upstream of the *yscW* and *lcrF* coding sequence, and the fourU element in the intergenic region between the *yscW* and *lcrF* coding regions are marked in bold (49). All coding sequences are marked in blue.

address this issue failed, indicating that H-NS is essential for the biological fitness of members of this genus (54, 55).

It is also quite reasonable that the DNA binding abilities of H-NS/H-NS and H-NS/YmoA complexes are differentially influenced by temperature-induced topological changes of the promoter and/or intrinsic conformational alterations of the regulatory proteins (56, 57). In fact, it has been reported that the DNA topology of the virulence plasmid undergoes a conformational change upon upshift to 37°C, leading to significant derepression of *virF-lcrF* expression (15, 30, 58, 59). Rohde et al. observed that *Y. enterocolitica* mutants resistant to the DNA gyrase inhibitor novobiocin constitutively expressed Yops, similarly to a Δ*ymoA* mutant (59). Using a two-dimensional (2-D) gel-based assay, the authors identified several DNA-intrinsic bends in the pYV plasmid. Interestingly, the presence of a DNA bend within the intergenic region between the *yscW* and *virF* genes was identified (Fig. 2). This bend was shown to melt at 37°C, suggesting that this intrinsic bend could potentially inhibit transcription of *virF* at noninducing temperatures (59).

Lastly, YmoA of *Y. pestis* and *Y. pseudotuberculosis* is subject to proteolysis by the Lon and ClpP proteases at 37°C but not at common environmental temperatures between 15°C and 30°C (49, 57). As a result, YmoA-mediated repression of the *yscW-lcrF* operon is rapidly eliminated at 37°C to induce the T3SS, but the repression effect remains present at all temperatures in a *clpP-lon* deletion mutant (57). Nevertheless, LcrF synthesis is still significantly enhanced in a *ymoA*-deficient strain at host body temperature compared to 25°C, indicating the importance of thermally induced DNA bending, the RNA thermometer (see below), and, possibly, H-NS homodimers.

Cross-regulation with flagellar system. Many *Yersinia* flagellar genes, including the alternative flagellum-specific σ^{28} sigma factor encoded by *fliA*, are strictly controlled by temperature. In contrast to T3SS genes, they are upregulated only at moderate temperatures and repressed at body temperature (60). *Yersinia* likely utilizes flagellar motility in the environment at temperatures under 37°C and does not require the T3SS prior to transmission into a mammalian host. This suggests an inverse regulation of flagellar and T3SS genes, and some evidence exists that σ^{28}/FliA is crucial for this process (61). Through microarray analysis comparing the *Y. enterocolitica* wild-type strain and a Δ*fliA* mutant, several pYV-borne genes such as *virF* were found to be upregulated in the Δ*fliA* mutant at 25°C (61). Furthermore, a putative binding site for FliA was identified in the *virF* promoter (61), suggesting that FliA binds to the *virF* promoter to repress transcription under temperature conditions that induce flagellar expression and assembly (61). However, we could not identify the putative FliA site described by Horne and Prüss (61) within the *virF* upstream region either by scanning for the exact FliA binding site sequence suggested by the authors or by using known FliA binding sites to make a motif model to search against the *Y. enterocolitica* genome (unpublished observations). While a discernible FliA motif may not be present in the *virF* promoter region in the current *Y. enterocolitica* genome assembly, the discrepancy might be due to differences between the current genome assembly and the one used in Horne and Prüss. Interestingly, *Y. pseudotuberculosis* bacteria lacking the RNA chaperone Hfq are hypermotile but defective in type III secretion (62). Thus, further studies are needed to identify the mechanism(s) involved in maintaining this inverse relationship between flagellar motility and T3SS expression.

IscR. It has been recently suggested that LcrF might be affected by environmental signals other than temperature. Through a forward genetic screen for modulators of the *Y. pseudotuberculosis* T3SS, the iron-sulfur cluster coordinating transcription regulator IscR was identified as important for type III secretion and *Yersinia* virulence (28). IscR is a global transcriptional regulator that has been extensively characterized in *E. coli* (63). The ability of IscR to modulate transcription of target genes depends on coordination of a [2Fe-2S] cluster, and IscR is an active transcription factor in both the apo-IscR and holo-IscR forms (63–65). This is due to the ability of IscR to recognize two separate DNA binding motifs: type I motifs are bound by holo-IscR, while type II motifs are recognized by both apo-IscR and holo-IscR (63, 64, 66). In *Y. pseudotuberculosis*, it was shown that IscR binds to a type II motif within the *lcrF* promoter (Fig. 2) (28), suggesting that IscR controls transcription of *Yersinia* type III secretion directly. Furthermore, this motif within the *lcrF* promoter is 100% conserved between *Y. pseudotuberculosis and Y. pestis* species and contains all nine residues found to be critical for IscR binding in *E. coli* (64). While the IscR binding site is not 100% conserved in *Y. enterocolitica*, the nine residues that were found to be critical for IscR binding are conserved (Fig. 2, marked in bold). This indicates a possible mechanism of T3SS gene regulation that is conserved among the three pathogens. It has been suggested that oxidative stress and oxygen limitation (as a result of Fe-S cluster damage) as well as iron availability influence the apo-IscR/holo-IscR ratio and that these environmental signals may affect IscR expression and activity (67–70). Therefore, while it has not yet been demonstrated, oxidative stress, oxygen limitation, or iron availability may influence expression of LcrF and the T3SS.

RcsB and CpxR in response to extracytoplasmic stress. A recent study reported that T3SS/*yop* expression in *Yersinia* is regulated by the Rcs phosphorelay system—a complex signaling pathway used by members of the *Enterobacteriaceae* to adapt their cellular physiology, biofilm and capsule formation, and motility in response to perturbations in external or surface-associated processes, e.g., overproduction of envelope components, osmotic shock, or desiccation (71). Overexpression of the wild type or a constitutive active variant of the response regulator RcsB enhanced mRNA levels of LcrF as well as Yop protein expression and secretion, suggesting that RcsB influences T3SS/Yop through LcrF (71). This was confirmed by the fact that activated/phosphorylated RcsB has the capacity to bind directly to a conserved RcsB box just upstream of the −35 promoter element of the *yscW-lcrF* operon (Fig. 2). RcsB binding most likely enhances RNA polymerase binding and/or function (71).

In addition to RcsB, it has been established that a second response regulator, CpxR, of another prominent phosphorelay system that responds to extracytoplasmic stress conditions targets the *yscW-lcrF* promoter region. However, in contrast to RcsB, CpxR represses transcription of the operon, suggesting that the two regulatory components either are part of a joint regulatory cascade or are separately induced during different infection stages or in different niches (71, 72).

Effects on *lcrF* revealed by global expression analyses. To gain insight into genes expressed by *Yersinia* during septicemia, global transcription patterns of bacteria grown in human plasma were compared with those of bacteria grown in Luria-Bertani broth, a standard laboratory medium (73, 74). *Y. pestis virG* and *yscW* and *Y. pseudotuberculosis lcrF* were specifically upregulated

in human plasma at 37°C, indicating that induction of the *yscW-lcrF* operon is important for virulence during the septicemic phase of the infection. A strong upregulation of *lcrF* gene expression was also observed at various time points after nasal infection of mice with *Y. pestis* in the lungs, spleen, and liver, supporting previous studies indicating that LcrF-controlled T3SS/*yop* genes are crucial in resisting immune and inflammatory defensive responses during the development of pneumonic plague (75).

A transcriptomic study designed to identify genes under the control of the recently recognized YbeY endonuclease demonstrated that deletion of the *ybeY* gene led to an upregulation of *lcrF* and the T3SS/*yop* genes in *Y. enterocolitica* serotype O:3 under conditions in which these genes are usually repressed (i.e., at 22°C) (76). This derepression was not caused by lower YmoA levels, since the amount of *ymoA* transcript was increased in the *ybeY* mutant. Instead, the authors suggested that *lcrF* upregulation in the *ybeY* mutant was due to increased copy numbers of the virulence plasmid or to altered regulation of global regulators (e.g., Hfq, nucleoid-structuring proteins) affecting the noncoding RNA network and/or DNA supercoiling (76).

Transcriptional profiling further revealed that bacterial membrane permeability may affect expression of LcrF and the T3SS/Yops. Microarray analysis showed that LcrF was upregulated by ~2-fold in a Δ*rovA Y. pestis* mutant grown under T3SS-inducing conditions (77). This result seems surprising, as RovA is a transcriptional regulator that is upregulated at 25°C and best known for inducing expression of invasin, an important virulence factor in enteropathogenic *Yersinia* but absent from *Y. pestis*. Through gel shift analysis, the possibility that RovA binds to the *lcrF* promoter to modulate transcription was ruled out. However, small electron-dense particles were found surrounding the Δ*rovA* mutant membrane by the use of transmission electron microscopy and the membrane permeability of the mutant was decreased compared to that of the wild type (77). The authors suggested that membrane construction is altered in the absence of RovA and that this might impact assembly and regulation of the T3SS. However, the connection between membrane integrity and LcrF has not been further explored.

TRANSLATIONAL CONTROL OF *lcrF* IN RESPONSE TO TEMPERATURE

Several early studies on temperature sensing in *Yersinia* reported that levels of protein encoded by LcrF-regulated genes such as *yopE* still change in response to temperature even under conditions in which *lcrF* transcription remains constant (37). Moreover, forced transcription of *lcrF* at low temperatures did not cause induction of LcrF-dependent genes (20). This implied that posttranscriptional mechanisms modulate LcrF synthesis and/or its specific activity in response to temperature. Hoe et al. (78) reported that *Y. pestis* LcrF is controlled at the translational level in response to temperature. This thermal control was maintained when *lcrF*, containing only 208 bp of the upstream 5′ untranslated region, was transcribed by the T7 polymerase in *E. coli* (78). Based on these data, a simple model of thermal regulation of *lcrF* translation was proposed in which the presence of a short predicted thermolabile stem-loop structure, or RNA thermometer, sequesters the *lcrF* ribosome binding site (rbs) sequence and blocks translation initiation at moderate temperatures. The decreased stability or melting of this structure at higher temperatures liberates the rbs sequence and allows formation of a productive mRNA complex and efficient translation (78).

A comprehensive secondary-structure prediction of the 124-bp intergenic region of *yscW* and *lcrF* performed using algorithms such as mfold and RNAfold suggested that this untranslated region of the bicistronic operon folds into two hairpin structures with a free energy of −19.67 kcal/mol (Fig. 2) (49). The second temperature-sensitive stem-loop includes a stretch of four uridines base-paired with the AGGA sequence of the rbs (49). This short motif with its simple design is referred to as a fourU element. It was first discovered in the *agsA* (aggregation suppression A) heat shock gene in *Salmonella enterica* serovar Typhimurium and bears resemblance to other potential fourU elements identified in the 5′ untranslated region of the *groES* and *dnaJ* heat shock genes of *Staphylococcus aureus* and *Brucella melitensis* (79). The presence of small loops and several noncanonical base pairs coupled with a network of weak hydrogen bonds, which facilitate liberation of the *lcrF* rbs, argued for a thermolabile RNA structure prone to melting within a physiological temperature range (49).

Enzymatic structural probing experiments using RNases T1 and V1 confirmed *in silico* predictions and demonstrated temperature-induced partial, but not complete, opening of the second hairpin loop (49). Existence of a thermosensing RNA element (RNA thermometer) was further confirmed by (i) *lcrF-lacZ* translational fusions, which were thermally induced when the fusion was transcribed from a temperature-independent P_{BAD} promoter, (ii) toe printing assays demonstrating that binding of ribosomes to the *lcrF* translational start site is restricted to 37°C and does not occur at 25°C, and (iii) base substitutions within the second hairpin of the thermosensing RNA element designed to stabilize or destabilize the second stem-loop (49). These stabilizing point mutations led to a "closed" conformation of the RNA element, resulting in full repression of LcrF synthesis at 37°C. In contrast, the destabilizing mutations allowed an opening of RNA structure ("open" conformation) and enhanced LcrF production at moderate and higher temperatures. Posttranscriptional control in an RNA thermometer-like fashion was further confirmed with a deletion of a stretch of nucleotides implicated in the formation of hairpin II, which completely abolished the thermosensing function of the intergenic RNA element and provoked constitutive increased synthesis of LcrF in a temperature range from 25°C to 37°C. In contrast, an RNA thermometer variant, which consisted only of the second hairpin, was more open than the full RNA thermometer and was characterized by higher LcrF levels. Nevertheless, this shortened version was still temperature inducible, indicating that the first stem-loop is not essential for thermosensing but seems to promote proper folding and/or supports the stability of the second hairpin (49).

In order to test the physiological relevance of the *lcrF* RNA thermometer and its role in *Yersinia* virulence, the pathogenicities of a closed RNA thermometer variant and an open variant of *Y. pseudotuberculosis* were compared with that of the isogenic wild-type strain in an oral mouse infection model. Despite the fact that all mice were challenged with a normally lethal dose of *Y. pseudotuberculosis*, all animals infected with *Yersinia* mutants encoding the closed RNA thermometer variant survived and showed no visible signs of infection, similarly to an *lcrF*-deficient strain, and displayed decreased colonization of the Peyer's patches, mesenteric lymph nodes, liver, and spleen (49). Intriguingly, overexpression of the T3SS/*yop* virulence program, as displayed by the open variant, was not beneficial and did not cause greater host mortality. In contrast, colonization of some host tissues was slightly re-

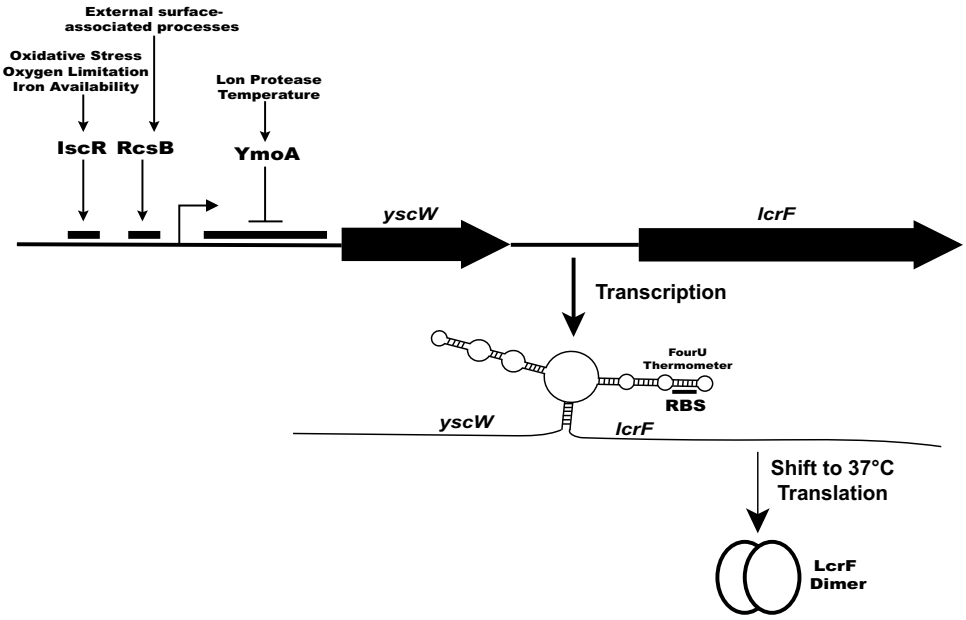


FIG 3 Multiple environmental signals control *lcrF* expression and, subsequently, T3SS expression through several distinct transcriptional and translational regulatory mechanisms. The data summarized in this review suggest that the YmoA, RcsB, and IscR regulators control transcription of *lcrF* in response to temperature, extracytoplasmic stress, iron bioavailability, oxygen tension, and reactive oxygen species. In addition, the RNA thermometer found upstream of *lcrF* allows LcrF translation only at the mammalian host body temperature, 37°C. As *Yersinia* transits from the environment or the flea vector to the mammalian host and then from localized to disseminated sites of infection, changes in temperature, iron availability, and stresses such as ROS may direct the regulatory network controlling LcrF, optimizing T3SS deployment and virulence.

duced, and the average time to death remained unchanged or was even increased by several days, most likely due to biological fitness impediments of the pathogen or increased inflammation in the host (49). This clearly illustrated that the *lcrF* RNA thermometer, all examples of which are 100% identical in all human-pathogenic *Yersinia* species, is a decisive posttranscriptional control element evolved to produce just the right amount of LcrF to promote the most ideal infection efficiency.

Apparently, both the YmoA- and RNA thermometer-mediated thermosensing mechanisms achieve a very rapid and efficient response. However, recent reports indicate that this control strategy seems to be complemented by additional regulatory modules adjusting LcrF synthesis according to host cell contact and T3SS-mediated effector translocation. In fact, the YopD translocator protein was recently found to bind to the 5′ untranslated sequences of multiple T3SS/*yop* genes, including *lcrF*, which facilitates their degradation (47). The molecular mechanism of YopD-mediated repression of LcrF synthesis is still unknown, but it is possible that YopD binding to the *lcrF* transcript in the absence of host cells (i) promotes a more closed conformation of the RNA thermometer and/or (ii) accelerates the degradation of the *lcrF* transcript as a consequence of the blockage of ribosome binding and translation. Alternatively, it is possible that YopD controls expression of additional factors influencing *lcrF* transcript stability.

CONCLUSIONS

Yersinia bacteria and other T3SS-expressing pathogens have complex regulatory networks in place to control expression of T3SS genes. While the *Yersinia* T3SS master regulator LcrF was identified almost three decades ago, recent work has greatly expanded our understanding of how expression of LcrF is regulated and what environmental signals might contribute to its regulation. YmoA, RcsB, and IscR all enhance transcription of *lcrF* (Fig. 3). As *Yersinia* experiences changes in temperature during the transition from the environment or the flea vector to the mammalian host, as well as various stresses, including iron limitation and reactive oxygen species (ROS) production during infection, sensing these environmental cues to control LcrF and T3SS expression may enable *Yersinia* to optimize T3SS deployment and virulence. While temperature has indeed been shown to contribute to *lcrF* transcription, future work focusing on expression of LcrF and T3SS genes in different host niches will enable a more complete understanding of how this T3SS master regulator facilitates optimization of T3SS expression to promote virulence.

FUNDING INFORMATION

German Research Foundation provided funding to Petra Dersch under grant number DE616/4. German Center for Infection Research provided funding to Petra Dersch under grant number DZIF-TTU 06.801. HHS | National Institutes of Health (NIH) provided funding to Victoria Auerbuch under grant number R21AI099747. HHS | National Institutes of Health (NIH) provided funding to Leah Schwiesow under grant number T32GM008646.

We also acknowledge the Vietnam Education Foundation for support (to H.L.).

REFERENCES

1. **Putzker M, Sauer H, Sobe D.** 2001. Plague and other human infections caused by *Yersinia* species. Clin Lab **47:**453–466.
2. **Khan FA, Fisher MA, Khakoo RA.** 2007. Association of hemochromatosis with infectious diseases: expanding spectrum. Int J Infect Dis **11:**482–487. http://dx.doi.org/10.1016/j.ijid.2007.04.007.
3. **Bottone EJ.** 1999. *Yersinia enterocolitica*: overview and epidemiologic correlates. Microbes Infect **1:**323–333. http://dx.doi.org/10.1016/S1286-4579(99)80028-8.
4. **Smego RA, Frean J, Koornhof HJ.** 1999. Yersiniosis I: microbiological

and clinicoepidemiological aspects of plague and non-plague *Yersinia* infections. Eur J Clin Microbiol Infect Dis **18:**1–15. http://dx.doi.org/10.1007/s100960050219.
5. **Koornhof HJ, Smego RA, Jr, Nicol M.** 1999. Yersiniosis. II: the pathogenesis of *Yersinia* infections. Eur J Clin Microbiol Infect Dis **18:**87–112. http://dx.doi.org/10.1007/s100960050237.
6. **Cornelis GR, Boland A, Boyd AP, Geuijen C, Iriarte M, Neyt C, Sory MP, Stainier I.** 1998. The virulence plasmid of *Yersinia*, an antihost genome. Microbiol Mol Biol Rev **62:**1315–1352.
7. **Bliska JB, Wang X, Viboud GI, Brodsky IE.** 2013. Modulation of innate immune responses by *Yersinia* type III secretion system translocators and effectors. Cell Microbiol **15:**1622–1631.
8. **Kupferberg LL, Higuchi K.** 1958. Role of calcium ions in the stimulation of growth of virulent strains of *Pasteurella pestis*. J Bacteriol **76:**120–121.
9. **Higuchi K, Kupferberg LL, Smith JL.** 1959. Studies on the nutrition and physiology of *Pasteurella pestis*. *III*. Effects of calcium ions on the growth of virulent and avirulent strains of Pasteurella pestis. J Bacteriol **77:**317–321.
10. **Yother J, Goguen JD.** 1985. Isolation and characterization of Ca2+-blind mutants of Yersinia pestis. J Bacteriol **164:**704–711.
11. **Perry RD, Harmon PA, Bowmer WS, Straley SC.** 1986. A low-Ca^{2+} response operon encodes the V antigen of *Yersinia pestis*. Infect Immun **54:**428–434.
12. **Yother J, Chamness TW, Goguen JD.** 1986. Temperature-controlled plasmid regulon associated with low calcium response in *Yersinia pestis*. J Bacteriol **165:**443–447.
13. **Cornelis G, Vanootegem JC, Sluiters C.** 1987. Transcription of the *yop* regulon from *Y. enterocolitica* requires *trans* acting pYV and chromosomal genes. Microb Pathog **2:**367–379. http://dx.doi.org/10.1016/0882-4010(87)90078-7.
14. **Bölin I, Forsberg A, Norlander L, Skurnik M, Wolf-Watz H.** 1988. Identification and mapping of the temperature-inducible, plasmid-encoded proteins of *Yersinia* spp. Infect Immun **56:**343–348.
15. **Cornelis G, Sluiters C, de Rouvroit CL, Michiels T.** 1989. Homology between virF, the transcriptional activator of the *Yersinia* virulence regulon, and AraC, the *Escherichia coli* arabinose operon regulator. J Bacteriol **171:**254–262.
16. **Schleif R.** 2010. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. FEMS Microbiol Rev **34:**779–796. http://dx.doi.org/10.1111/j.1574-6976.2010.00226.x.
17. **Wattiau P, Cornelis GR.** 1994. Identification of DNA sequences recognized by VirF, the transcriptional activator of the *Yersinia yop* regulon. J Bacteriol **176:**3878–3884.
18. **King JM, Schesser Bartra S, Plano G, Yahr TL.** 2013. ExsA and LcrF recognize similar consensus binding sites, but differences in their oligomeric state influence interactions with promoter DNA. J Bacteriol **195:**5639–5650. http://dx.doi.org/10.1128/JB.00990-13.
19. **Rhee S, Martin RG, Rosner JL, Davies DR.** 1998. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. Proc Natl Acad Sci U S A **95:**10413–10418. http://dx.doi.org/10.1073/pnas.95.18.10413.
20. **Lambert de Rouvroit C, Sluiters C, Cornelis GR.** 1992. Role of the transcriptional activator, VirF, and temperature in the expression of the pYV plasmid genes of *Yersinia enterocolitica*. Mol Microbiol **6:**395–409. http://dx.doi.org/10.1111/j.1365-2958.1992.tb01483.x.
21. **Diaz MR, King JM, Yahr TL.** 2011. Intrinsic and extrinsic regulation of type III secretion gene expression in *Pseudomonas aeruginosa*. Front Microbiol **2:**89.
22. **Brutinel ED, Vakulskas CA, Brady KM, Yahr TL.** 2008. Characterization of ExsA and of ExsA-dependent promoters required for expression of the *Pseudomonas aeruginosa* type III secretion system. Mol Microbiol **68:**657–671. http://dx.doi.org/10.1111/j.1365-2958.2008.06179.x.
23. **Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS.** 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res **37:**W202–W208. http://dx.doi.org/10.1093/nar/gkp335.
24. **Grant CE, Bailey TL, Noble WS.** 2011. FIMO: scanning for occurrences of a given motif. Bioinformatics **27:**1017–1018. http://dx.doi.org/10.1093/bioinformatics/btr064.
25. **Nuss AM, Heroven AK, Waldmann B, Reinkensmeier J, Jarek M, Beckstette M, Dersch P.** 2015. Transcriptomic profiling of *Yersinia pseudotuberculosis* reveals reprogramming of the Crp regulon by temperature and uncovers Crp as a master regulator of small RNAs. PLoS Genet **11:**e1005087. http://dx.doi.org/10.1371/journal.pgen.1005087.
26. **Munson GP, Scott JR.** 2000. Rns, a virulence regulator within the AraC family, requires binding sites upstream and downstream of its own promoter to function as an activator. Mol Microbiol **36:**1391–1402.
27. **Munson GP, Holcomb LG, Scott JR.** 2001. Novel group of virulence activators within the AraC family that are not restricted to upstream binding sites. Infect Immun **69:**186–193. http://dx.doi.org/10.1128/IAI.69.1.186-193.2001.
28. **Miller HK, Kwuan L, Schwiesow L, Bernick DL, Mettert E, Ramirez HA, Ragle JM, Chan PP, Kiley PJ, Lowe TM, Auerbuch V.** 2014. IscR is essential for *Yersinia pseudotuberculosis* type III secretion and virulence. PLoS Pathog **10:**e1004194. http://dx.doi.org/10.1371/journal.ppat.1004194.
29. **China B, Michiels T, Cornelis GR.** 1990. The pYV plasmid of *Yersinia* encodes a lipoprotein, YlpA, related to TraT. Mol Microbiol **4:**1585–1593. http://dx.doi.org/10.1111/j.1365-2958.1990.tb02070.x.
30. **Michiels T, Vanooteghem JC, Lambert de Rouvroit C, China B, Gustin A, Boudry P, Cornelis GR.** 1991. Analysis of *virC*, an operon involved in the secretion of Yop proteins by *Yersinia enterocolitica*. J Bacteriol **173:**4994–5009.
31. **Haddix PL, Straley SC.** 1992. Structure and regulation of the *Yersinia pestis yscBCDEF* operon. J Bacteriol **174:**4820–4828.
32. **Vakulskas CA, Brady KM, Yahr TL.** 2009. Mechanism of transcriptional activation by *Pseudomonas aeruginosa* ExsA. J Bacteriol **191:**6654–6664. http://dx.doi.org/10.1128/JB.00902-09.
33. **Brutinel ED, Vakulskas CA, Yahr TL.** 2009. Functional domains of ExsA, the transcriptional activator of the *Pseudomonas aeruginosa* type III secretion system. J Bacteriol **191:**3811–3821. http://dx.doi.org/10.1128/JB.00002-09.
34. **Yahr TL, Frank DW.** 1994. Transcriptional organization of the trans-regulatory locus which controls exoenzyme S synthesis in *Pseudomonas aeruginosa*. J Bacteriol **176:**3832–3838.
35. **Michiels T, Cornelis G.** 1988. Nucleotide sequence and transcription analysis of yop51 from *Yersinia enterocolitica* W22703. Microb Pathog **5:**449–459. http://dx.doi.org/10.1016/0882-4010(88)90006-X.
36. **Skurnik M, Toivanen P.** 1992. LcrF is the temperature-regulated activator of the *yadA* gene of *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. J Bacteriol **174:**2047–2051.
37. **Hoe NP, Minion FC, Goguen JD.** 1992. Temperature sensing in *Yersinia pestis*: regulation of *yopE* transcription by *lcrF*. J Bacteriol **174:**4275–4286.
38. **Axler-DiPerte GL, Hinchliffe SJ, Wren BW, Darwin AJ.** 2009. YtxR acts as an overriding transcriptional off switch for the *Yersinia enterocolitica* Ysc-Yop type 3 secretion system. J Bacteriol **191:**514–524. http://dx.doi.org/10.1128/JB.01305-08.
39. **Li L, Yan H, Feng L, Li Y, Lu P, Hu Y, Chen S.** 2014. LcrQ blocks the role of LcrF in regulating the Ysc-Yop type III secretion genes in *Yersinia pseudotuberculosis*. PLoS One **9:**e92243. http://dx.doi.org/10.1371/journal.pone.0092243.
40. **Pettersson J, Nordfelth R, Dubinina E, Bergman T, Gustafsson M, Magnusson KE, Wolf-Watz H.** 1996. Modulation of virulence factor expression by pathogen target cell contact. Science **273:**1231–1233. http://dx.doi.org/10.1126/science.273.5279.1231.
41. **Rimpiläinen M, Forsberg A, Wolf-Watz H.** 1992. A novel protein, LcrQ, involved in the low-calcium response of Yersinia pseudotuberculosis shows extensive homology to YopH. J Bacteriol **174:**3355–3363.
42. **Stainier I, Iriarte M, Cornelis GR.** 1997. YscM1 and YscM2, two *Yersinia enterocolitica* proteins causing downregulation of *yop* transcription. Mol Microbiol **26:**833–843. http://dx.doi.org/10.1046/j.1365-2958.1997.6281995.x.
43. **Kopaskie KS, Ligtenberg KG, Schneewind O.** 2013. Translational regulation of *Yersinia enterocolitica* mRNA encoding a type III secretion substrate. J Biol Chem **288:**35478–35488. http://dx.doi.org/10.1074/jbc.M113.504811.
44. **Williams AW, Straley SC.** 1998. YopD of *Yersinia pestis* plays a role in negative regulation of the low-calcium response in addition to its role in translocation of Yops. J Bacteriol **180:**350–358.
45. **Anderson DM, Ramamurthi KS, Tam C, Schneewind O.** 2002. YopD and LcrH regulate expression of *Yersinia enterocolitica* YopQ by a post-transcriptional mechanism and bind to yopQ RNA. J Bacteriol **184:**1287–1295. http://dx.doi.org/10.1128/JB.184.5.1287-1295.2002.
46. **Cambronne ED, Schneewind O.** 2002. *Yersinia enterocolitica* type III secretion: *yscM1* and *yscM2* regulate *yop* gene expression by a posttran-

scriptional mechanism that targets the 5′ untranslated region of *yop* mRNA. J Bacteriol **184:**5880–5893. http://dx.doi.org/10.1128/JB.184.21.5880-5893.2002.
47. **Chen Y, Anderson DM.** 2011. Expression hierarchy in the *Yersinia* type III secretion system established through YopD recognition of RNA. Mol Microbiol **80:**966–980. http://dx.doi.org/10.1111/j.1365-2958.2011.07623.x.
48. **Cornelis G, Sluiters C, Delor I, Geib D, Kaniga K, Lambert de Rouvroit C, Sory MP, Vanooteghem JC, Michiels T.** 1991. *ymoA*, a *Yersinia enterocolitica* chromosomal gene modulating the expression of virulence functions. Mol Microbiol **5:**1023–1034. http://dx.doi.org/10.1111/j.1365-2958.1991.tb01875.x.
49. **Böhme K, Steinmann R, Kortmann J, Seekircher S, Heroven AK, Berger E, Pisano F, Thiermann T, Wolf-Watz H, Narberhaus F, Dersch P.** 2012. Concerted actions of a thermo-labile regulator and a unique intergenic RNA thermosensor control *Yersinia* virulence. PLoS Pathog **8:**e1002518. http://dx.doi.org/10.1371/journal.ppat.1002518.
50. **Cornelis GR.** 1993. Role of the transcription activator *virF* and the histone-like protein YmoA in the thermoregulation of virulence functions in yersiniae. Zentralbl Bakteriol **278:**149–164. http://dx.doi.org/10.1016/S0934-8840(11)80833-9.
51. **Mikulskis AV, Cornelis GR.** 1994. A new class of proteins regulating gene expression in enterobacteria. Mol Microbiol **11:**77–86. http://dx.doi.org/10.1111/j.1365-2958.1994.tb00291.x.
52. **Madrid C, Nieto JM, Juarez A.** 2002. Role of the Hha/YmoA family of proteins in the thermoregulation of the expression of virulence factors. Int J Med Microbiol **291:**425–432.
53. **Nieto J, Madrid C, Miquelay E, Parra JL, Rodriguez S, Juarez A.** 2002. Evidence for direct protein-protein interaction between members of the enterobacterial Hha/YmoA and H-NS families of proteins. J Bacteriol **184:**629–635. http://dx.doi.org/10.1128/JB.184.3.629-635.2002.
54. **Heroven A, Nagel G, Tran HJ, Parr S, Dersch P.** 2004. RovA is autoregulated and antagonizes H-NS-mediated silencing of invasin and *rovA* expression in *Yersinia pseudotuberculosis*. Mol Microbiol **53:**871–888. http://dx.doi.org/10.1111/j.1365-2958.2004.04162.x.
55. **Ellison DW, Miller VL.** 2006. H-NS represses inv transcription in *Yersinia enterocolitica* through competition with RovA and interaction with YmoA. J Bacteriol **188:**5101–5112. http://dx.doi.org/10.1128/JB.00862-05.
56. **Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JC, Ladbury JE.** 2005. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem J **391:**203–213. http://dx.doi.org/10.1042/BJ20050453.
57. **Jackson M, Silva-Herzog E, Plano GV.** 2004. The ATP-dependent ClpXP and Lon proteases regulate expression of the *Yersinia pestis* type III secretion system via regulated proteolysis of YmoA, a small histone-like protein. sMol Microbiol **54:**1364–1378. http://dx.doi.org/10.1111/j.1365-2958.2004.04353.x.
58. **Rohde JR, Fox JM, Minnich SA.** 1994. Thermoregulation in *Yersinia enterocolitica* is coincident with changes in DNA supercoiling. Mol Microbiol **12:**187–199. http://dx.doi.org/10.1111/j.1365-2958.1994.tb01008.x.
59. **Rohde JR, Luan XS, Rohde H, Fox JM, Minnich SA.** 1999. The *Yersinia enterocolitica* pYV virulence plasmid contains multiple intrinsic DNA bends which melt at 37 degrees C. J Bacteriol **181:**4198–4204.
60. **Kapatral V, Olson JW, Pepe JC, Miller VL, Minnich SA.** 1996. Temperature-dependent regulation of *Yersinia enterocolitica* Class III flagellar genes. Mol Microbiol **19:**1061–1071. http://dx.doi.org/10.1046/j.1365-2958.1996.452978.x.
61. **Horne SM, Prüss BM.** 2006. Global gene regulation in *Yersinia enterocolitica*: effect of FliA on the expression levels of flagellar and plasmid-encoded virulence genes. Arch Microbiol **185:**115–126. http://dx.doi.org/10.1007/s00203-005-0077-1.
62. **Schiano CA, Bellows LE, Lathem WW.** 2010. The small RNA chaperone Hfq is required for the virulence of *Yersinia pseudotuberculosis*. Infect Immun **78:**2034–2044. http://dx.doi.org/10.1128/IAI.01046-09.
63. **Schwartz CJ, Giel JL, Patschkowski T, Luther C, Ruzicka FJ, Beinert H, Kiley PJ.** 2001. IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. Proc Natl Acad Sci U S A **98:**14895–14900. http://dx.doi.org/10.1073/pnas.251550898.
64. **Nesbit AD, Giel JL, Rose JC, Kiley PJ.** 2009. Sequence-specific binding to a subset of IscR-regulated promoters does not require IscR Fe-S cluster ligation. J Mol Biol **387:**28–41. http://dx.doi.org/10.1016/j.jmb.2009.01.055.
65. **Fleischhacker AS, Stubna A, Hsueh KL, Guo Y, Teter SJ, Rose JC, Brunold TC, Markley JL, Munck E, Kiley PJ.** 2012. Characterization of the [2Fe-2S] cluster of Escherichia coli transcription factor IscR. Biochemistry **51:**4453–4462. http://dx.doi.org/10.1021/bi3003204.
66. **Rajagopalan S, Teter SJ, Zwart PH, Brennan RG, Phillips KJ, Kiley PJ.** 2013. Studies of IscR reveal a unique mechanism for metal-dependent regulation of DNA binding specificity. Nat Struct Mol Biol **20:**740–747. http://dx.doi.org/10.1038/nsmb.2568.
67. **Mettert EL, Kiley PJ.** 2014. Coordinate regulation of the Suf and Isc Fe-S cluster biogenesis pathways by IscR is essential for viability of *Escherichia coli*. J Bacteriol **196:**4315–4323. http://dx.doi.org/10.1128/JB.01975-14.
68. **Giel JL, Rodionov D, Liu M, Blattner FR, Kiley PJ.** 2006. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O2-regulated genes in *Escherichia coli*. Mol Microbiol **60:**1058–1075. http://dx.doi.org/10.1111/j.1365-2958.2006.05160.x.
69. **Yeo WS, Lee JH, Lee KC, Roe JH.** 2006. IscR acts as an activator in response to oxidative stress for the suf operon encoding Fe-S assembly proteins. Mol Microbiol **61:**206–218. http://dx.doi.org/10.1111/j.1365-2958.2006.05220.x.
70. **Wu Y, Outten FW.** 2009. IscR controls iron-dependent biofilm formation in *Escherichia coli* by regulating type I fimbria expression. J Bacteriol **191:**1248–1257. http://dx.doi.org/10.1128/JB.01086-08.
71. **Li Y, Hu Y, Francis MS, Chen S.** 2015. RcsB positively regulates the *Yersinia* Ysc-Yop type III secretion system by activating expression of the master transcriptional regulator LcrF. Environ Microbiol **17:**1219–1233. http://dx.doi.org/10.1111/1462-2920.12556.
72. **Liu J, Thanikkal EJ, Obi IR, Francis MS.** 2012. Elevated CpxR∼P levels repress the Ysc-Yop type III secretion system of *Yersinia pseudotuberculosis*. Res Microbiol **163:**518–530. http://dx.doi.org/10.1016/j.resmic.2012.07.010.
73. **Rosso ML, Chauvaux S, Dessein R, Laurans C, Frangeul L, Lacroix C, Schiavo A, Dillies MA, Foulon J, Coppee JY, Medigue C, Carniel E, Simonet M, Marceau M.** 2008. Growth of *Yersinia pseudotuberculosis* in human plasma: impacts on virulence and metabolic gene expression. BMC Microbiol **8:**211. http://dx.doi.org/10.1186/1471-2180-8-211.
74. **Chauvaux S, Rosso ML, Frangeul L, Lacroix C, Labarre L, Schiavo A, Marceau M, Dillies MA, Foulon J, Coppee JY, Medigue C, Simonet M, Carniel E.** 2007. Transcriptome analysis of *Yersinia pestis* in human plasma: an approach for discovering bacterial genes involved in septicaemic plague. Microbiology **153:**3112–3124. http://dx.doi.org/10.1099/mic.0.2007/006213-0.
75. **Liu H, Wang H, Qiu J, Wang X, Guo Z, Qiu Y, Zhou D, Han Y, Du Z, Li C, Song Y, Yang R.** 2009. Transcriptional profiling of a mice plague model: insights into interaction between *Yersinia pestis* and its host. J Basic Microbiol **49:**92–99. http://dx.doi.org/10.1002/jobm.200800027.
76. **Leskinen K, Varjosalo M, Skurnik M.** 2015. Absence of YbeY RNase compromises the growth and enhances the virulence plasmid gene expression of *Yersinia enterocolitica* O:3. Microbiology **161:**285–299. http://dx.doi.org/10.1099/mic.0.083097-0.
77. **Yang F, Ke Y, Tan Y, Bi Y, Shi Q, Yang H, Qiu J, Wang X, Guo Z, Ling H, Yang R, Du Z.** 2010. Cell membrane is impaired, accompanied by enhanced type III secretion system expression in *Yersinia pestis* deficient in RovA regulator. PLoS One **5:**e12840. http://dx.doi.org/10.1371/journal.pone.0012840.
78. **Hoe NP, Goguen JD.** 1993. Temperature sensing in *Yersinia pestis*: translation of the LcrF activator protein is thermally regulated. J Bacteriol **175:**7901–7909.
79. **Waldminghaus T, Heidrich N, Brantl S, Narberhaus F.** 2007. FourU: a novel type of RNA thermometer in *Salmonella*. Mol Microbiol **65:**413–424. http://dx.doi.org/10.1111/j.1365-2958.2007.05794.x.
80. **Venecia K, Young GM.** 2005. Environmental regulation and virulence attributes of the Ysa type III secretion system of *Yersinia enterocolitica* biovar 1B. Infect Immun **73:**5961–5977. http://dx.doi.org/10.1128/IAI.73.9.5961-5977.2005.
81. **Gouy M, Guindon S, Gascuel O.** 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol **27:**221–224. http://dx.doi.org/10.1093/molbev/msp259.

Victoria Auerbuch received her diploma from Cornell University in 1997 and her Ph.D. from the University of California, Berkeley (U.C. Berkeley), in 2002, where she worked with Daniel Portnoy. She was a postdoctoral researcher at U.C. Berkeley from 2003 to 2004 and then with Ralph Isberg at Tufts University School of Medicine, where she began working on the immune response of mammalian cells to the *Yersinia* type III secretion system (T3SS). In 2009, she joined the faculty in the Department of Microbiology and Environmental Toxicology at U.C. Santa Cruz, where she is now an associate professor. Her current main research interests include IscR control of virulence factor utilization in *Yersinia* and development of T3SS inhibitors.

Petra Dersch performed the work for her Ph.D. thesis at University Konstanz and the Max Planck Institute for Terrestrial Microbiology in Marburg on the nucleoid-associated H-NS protein of *E. coli*. She was a postdoctoral fellow with Ralph Isberg at the Tufts Medical School/Howard Hughes Institute in Boston, MA, where she became interested in the function and regulation of crucial virulence-related factors pertaining to enteropathogenic *Yersinia* species. She continued her work as independent group leader at the Free University in Berlin and as a junior research group leader at the Robert Koch Institute in Berlin. She joined the faculty of microbiology at the Technical University Braunschweig in 2005 and since 2008 has been head of the Department of Molecular Infection Biology at the Helmholtz Center for Infection Research in Braunschweig. Her current main research interest is the global regulation of *Yersinia* virulence factors by posttranscriptional control mechanisms, with an emphasis on sensory and regulatory RNAs.

Hanh Lam completed her Ph.D. in plant pathology and plant-microbe biology at Cornell University. Her thesis research was on the plant pathogen *Pseudomonas syringae* pv. DC3000, focusing on regulation of that organism's type III secretion system. At Cornell, Hanh developed interests in and was trained in bacterial gene regulation, genomics, and computational biology. In 2014, she joined Dr. Auerbuch's research group at the University of California, Santa Cruz, as a postdoctoral scientist. Dr. Hanh's main research focus is on developing methods for high-throughput screening to identify antimicrobial compounds targeting the T3SS. The resulting compounds could be used for therapeutics and as biochemical tools to study T3SS regulation.

Leah Schwiesow received her B.S. in chemistry from Gonzaga University in Spokane, WA, in 2009. While an undergraduate, she became interested in molecular biology and worked to characterize the role of calcium in the regulation of photosynthesis in *Arabidopsis thaliana*. After graduating, she took a position at Dow Agrosciences, where she worked on a team optimizing maize transformation protocols. She started the molecular, cellular, and developmental biology Ph.D. program at the University of California, Santa Cruz, in 2011 and joined Dr. Auerbuch's research group, where she discovered her love for microbes. Her current research is focused on how IscR in *Yersinia pseudotuberculosis* responds to changing environmental signals within the host to optimize expression of factors important for virulence.

JB Journals.ASM.org

Resolving Nonstop Translation Complexes Is a Matter of Life or Death

Kenneth C. Keiler, Heather A. Feaga

The Pennsylvania State University, Department of Biochemistry & Molecular Biology, University Park, Pennsylvania, USA

Problems during gene expression can result in a ribosome that has translated to the 3′ end of an mRNA without terminating at a stop codon, forming a nonstop translation complex. The nonstop translation complex contains a ribosome with the mRNA and peptidyl-tRNA engaged, but because there is no codon in the A site, the ribosome cannot elongate or terminate the nascent chain. Recent work has illuminated the importance of resolving these nonstop complexes in bacteria. Transfer-messenger RNA (tmRNA)-SmpB specifically recognizes and resolves nonstop translation complexes in a reaction known as *trans*-translation. *trans*-Translation releases the ribosome and promotes degradation of the incomplete nascent polypeptide and problematic mRNA. tmRNA and SmpB have been found in all bacteria and are essential in some species. However, other bacteria can live without *trans*-translation because they have one of the alternative release factors, ArfA or ArfB. ArfA recruits RF2 to nonstop translation complexes to promote hydrolysis of the peptidyl-tRNAs. ArfB recognizes nonstop translation complexes in a manner similar to tmRNA-SmpB recognition and directly hydrolyzes the peptidyl-tRNAs to release the stalled ribosomes. Genetic studies indicate that most or all species require at least one mechanism to resolve nonstop translation complexes. Consistent with such a requirement, small molecules that inhibit resolution of nonstop translation complexes have broad-spectrum antibacterial activity. These results suggest that resolving nonstop translation complexes is a matter of life or death for bacteria.

Bacteria perform transcription and translation in the same cellular compartment because they do not have nuclei. One advantage to this arrangement is that bacteria can rapidly respond to environmental challenges by producing new proteins. The time between transcription of a gene and the availability of the corresponding protein is minimized because the mRNA does not have to be processed or exported, and translation of an mRNA can initiate before transcription is complete. However, using a single compartment for transcription and translation has serious consequences for protein quality control because there are limited opportunities for mRNA proofreading. Mechanisms used by eukaryotes to ensure that the mRNA is intact are generally absent in bacteria. For example, in eukaryotes, 3′ polyadenylation is used as a signal that the mRNA transcript is complete. This signal is read at several steps, including nuclear export and translation initiation, which requires interaction between poly(A)-binding proteins and translation initiation factors (1, 2). In contrast, the bacterial ribosome does not require any information from the 3′ end of the mRNA to initiate translation, so there is no assurance that the mRNA is complete or intact (3). mRNAs can be truncated by many events, including premature termination of transcription, nuclease activity, and physical damage. As a consequence, bacterial ribosomes frequently translate mRNAs that do not have a stop codon ("nonstop" mRNAs). When a ribosome reaches the 3′ end of a nonstop mRNA, it is trapped in a nonstop translation complex. In this complex, the mRNA and peptidyl-tRNA in the P site prevent dissociation of the ribosome, but the complex cannot elongate or terminate because there is no codon in the A site. A nonstop complex can also be formed when a ribosome stalls during translation and the mRNA is cleaved in the A site (4–6). Estimates from *Escherichia coli* suggest that 2% to 4% of translation reactions end in a nonstop translation complex (7). At that rate, an average ribosome is involved in ~5 nonstop translation complexes per cell division cycle. Clearly, the protein synthesis capacity of the cell would be severely compromised if these complexes could not be quickly resolved. To cope with the prevalence of nonstop translation complexes, bacteria have a remarkable mechanism known as *trans*-translation, which can release the ribosome and target the nonstop mRNA and nascent polypeptide for rapid degradation.

RESOLUTION OF NONSTOP TRANSLATION COMPLEXES BY tmRNA-SmpB

trans-Translation is performed by a ribonucleoprotein complex consisting of transfer-messenger RNA (tmRNA), a specialized RNA molecule, and SmpB, a small protein. tmRNA has elements of both a tRNA and an mRNA. The 5′ and 3′ ends of tmRNA form a structure resembling the acceptor arm and TΨC arm of alanyl-tRNA (8, 9). The remainder of tmRNA includes several pseudoknots and a specialized reading frame that is decoded during *trans*-translation (8, 10–12). SmpB binds tightly with tmRNA and completes the tRNA-like structure by mimicking the anticodon stem (13–15). The acceptor arm of tmRNA is charged with alanine by alanyl-tRNA synthetase and bound by EF-Tu in the same manner as $tRNA^{Ala}$ (8, 16, 17). During *trans*-translation, tmRNA-SmpB specifically recognizes a nonstop translation complex and is accommodated in the ribosomal A site (Fig. 1) (18–21). The nascent polypeptide is transferred to the alanine charged to tmRNA, and SmpB-tmRNA is translocated to the P site. During translocation, a large swivel of the 30S head of the ribosome allows the reading frame of tmRNA to enter the mRNA channel (22). The first codon of the tmRNA reading frame is aligned in the A site, and translation resumes using the tmRNA reading frame as a mes-

Published ahead of print 4 April 2014

Address correspondence to Kenneth C. Keiler, kkeiler@psu.edu.

doi:10.1128/JB.01490-14

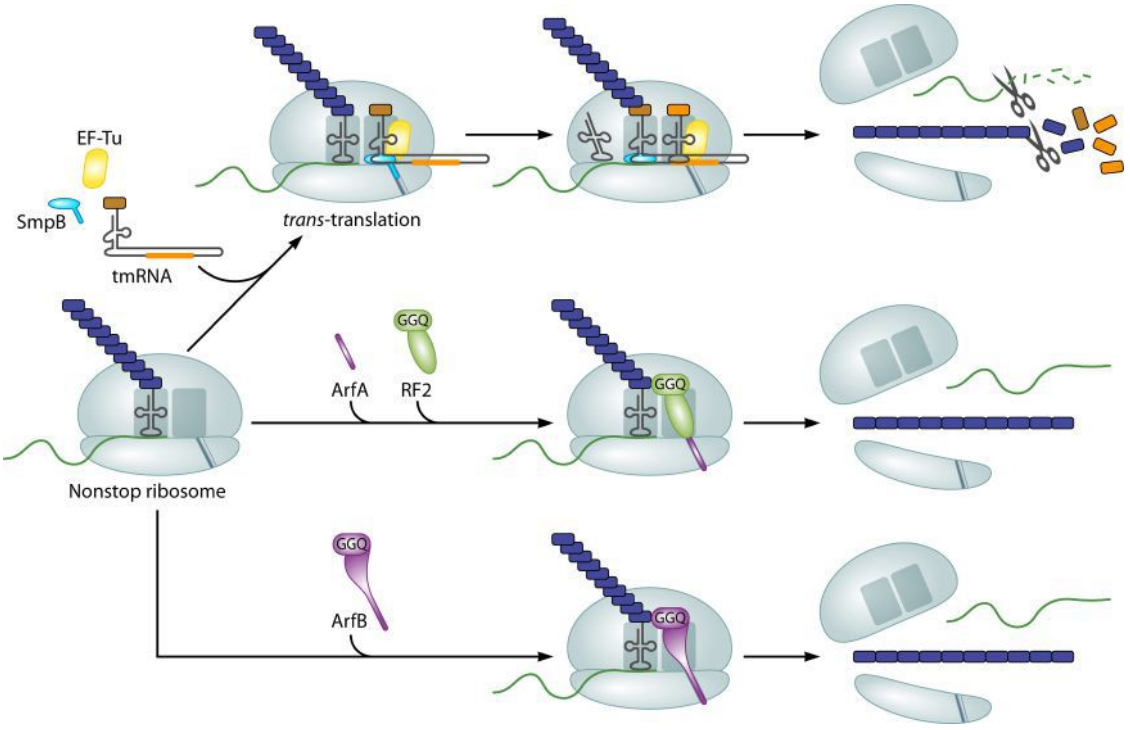


FIG 1 Mechanisms for resolving nonstop translation complexes. During *trans*-translation (top), tmRNA-SmpB recognizes nonstop translation complexes by binding in the empty mRNA channel and uses a reading frame within tmRNA to mediate the release of the ribosome and target the nascent polypeptide for proteolysis. The problematic mRNA is also degraded. Some bacteria have backup systems that use either ArfA or ArfB to recognize nonstop translation complexes. ArfA recruits RF2, which uses its GGQ motif to hydrolyze the peptidyl-tRNA in the ribosome. It is not known how ArfA recognizes nonstop translation complexes, but it might bind in the empty mRNA channel in a manner similar to that of SmpB and ArfB binding. ArfB contains a GGQ motif and directly hydrolyzes the peptidyl-tRNA on the ribosome. ArfA and ArfB release the ribosome but do not target the nascent polypeptide for degradation. See the text for details.

sage. Correct alignment of tmRNA in the mRNA channel requires sequence-specific contacts between tmRNA and SmpB (23). Translation of the tmRNA reading frame terminates at a stop codon, releasing the ribosome and a protein that includes the tmRNA-encoded peptide tag at the C terminus (24). The peptide tag is recognized by multiple proteases in the cell, ensuring that the protein is rapidly degraded (24–28). The nonstop mRNA is also targeted for degradation during *trans*-translation (29–31). Thus, the overall effect of the reaction is to remove the problematic mRNA and the incomplete protein and to release the ribosome (Fig. 1).

A crystal structure from Neubauer et al. captures an early step of *trans*-translation and shows how tmRNA-SmpB recognizes nonstop translation complexes (32) (Fig. 2). In the structure, the

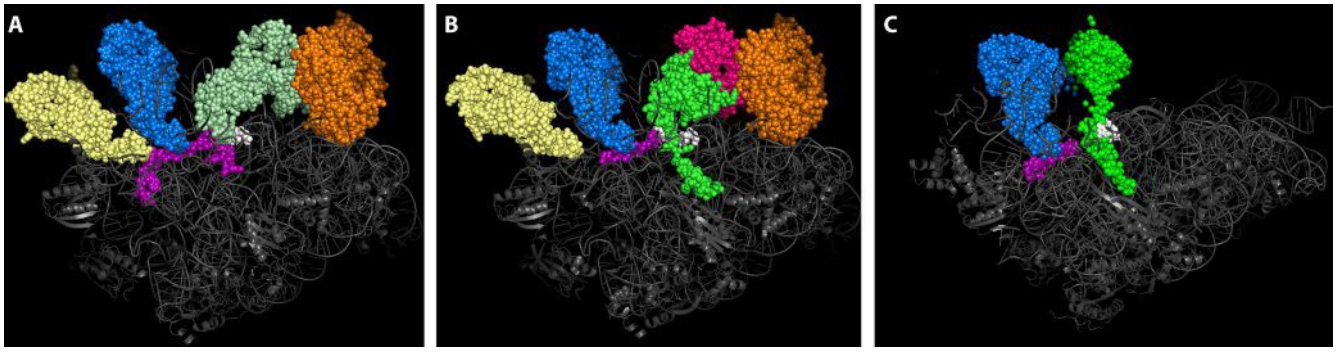


FIG 2 Recognition of nonstop translation complexes. Structure models of an elongation complex (A) with an intact mRNA compared to recognition of nonstop translation complexes by tmRNA-SmpB (B) and ArfB (C) are shown. The 30S ribosomal subunits are shown in gray, with decoding nucleotides G530, A1492, and A1493 in white. (A) An elongation complex trapped by kirromycin from PDB 2WRQ, with mRNA (purple), E-site tRNA (yellow), P-site tRNA (blue), and A-site tRNA (green) bound with EF-Tu (orange). (B) *trans*-Translation complex trapped by kirromycin from PDB 4ABR. The tRNA-like domain of tmRNA (pink) bound with EF-Tu (orange) is in an orientation similar to that of the acceptor stem of the tRNA shown in panel A. SmpB (green) occupies the codon-anticodon region and extends into the empty mRNA channel. (C) In nonstop translation complexes recognized by ArfB (from PDB 4DH9), ArfB (green) extends into the empty mRNA channel, with the catalytic GGQ domain near the peptidyl-tRNA in the P site (blue).

tRNA-like domain of tmRNA, bound with SmpB and EF-Tu, is trapped in the A site of a nonstop translation complex during accommodation using the drug kirromycin. Overall, the structure resembles an elongation complex with tmRNA-SmpB in place of the acylated tRNA. The acceptor arm of tmRNA is in the same orientation as the acceptor arm of the acylated tRNA, and SmpB takes the place of the anticodon stem. However, SmpB also makes contacts in the decoding center and empty mRNA channel that appear to mimic the missing mRNA. The 16S rRNA residues A1492, A1493, and G530, which interact with the mRNA in an elongation complex, directly contact SmpB in the nonstop complex. In addition, the C terminus of SmpB forms a helix that extends into the empty mRNA channel between the decoding center and the leading edge of the ribosome. Chemical footprinting and mutational studies support the hypothesis of the presence of these interactions during *trans*-translation (33, 34). This crystal structure suggests that tmRNA-SmpB could not be accommodated in elongating ribosomes because the mRNA would obstruct SmpB interactions with the 16S rRNA (Fig. 2). Consistent with this model, competition experiments show that tmRNA-SmpB does not interfere with translation elongation or termination *in vivo* (35).

Whereas the crystal structure suggests that the mRNA channel downstream of the A site must be empty for tmRNA-SmpB to bind, kinetic data indicate that the mRNA channel does not always have to be empty for *trans*-translation to occur. The rate of *trans*-translation *in vitro* was measured using ribosomes stalled on mRNAs of different lengths (36). When the ribosomes were stalled with the mRNA channel completely occupied (with >15 nucleotides downstream of the P site), the reaction was extremely slow, consistent with the mRNA blocking tmRNA-SmpB. However, the reaction was rapid when the ribosomes were stalled with 0 to 6 nucleotides of mRNA downstream of the P site and was inhibited only partially with 9 to 12 nucleotides downstream of the P site. These results imply that mRNA in the A site, and even several codons downstream of the A site, does not interfere with *trans*-translation. The substrates used for the kinetic measurements were generated by omitting a tRNA from the reaction, so they probably do not occur frequently *in vivo*. However, the issue of whether tmRNA-SmpB can act on ribosomes with mRNA extending past the A site has important implications for the mechanism of *trans*-translation. It is possible that the interactions between SmpB and 16S rRNA observed in the crystal structure represent the lowest energy conformation, but these interactions are not required for tmRNA-SmpB to initiate *trans*-translation. Alternatively, when a ribosome stalls on an mRNA that does not completely fill the mRNA channel, it might undergo a structural change that allows SmpB access to the 16S rRNA. For example, the 3′ end of the mRNA might loop out of the A site, or the ribosome could slide to the 3′ end of the mRNA, leaving the A site empty. Such rearrangements could be facilitated by communication between the mRNA channel and the decoding center of the ribosome. Further biochemical experiments are required to determine whether *trans*-translation always requires an empty mRNA channel.

SUBSTRATES FOR *trans*-TRANSLATION

Some of the known substrates for *trans*-translation are consistent with nonstop translation complexes generated by mRNA damage, but others suggest nonrandom or intentional mRNA cleavage to target translation reactions to *trans*-translation. Truncation of mRNA by premature termination of transcription, damage to the mRNA, or 3′-5′ exonucleolytic mRNA turnover would be expected to be largely random and should produce nonstop translation complexes at a variety of positions along many mRNAs. Two proteomic-analysis-scale studies identified proteins tagged by *trans*-translation in *Caulobacter crescentus* and *Francisella tularensis*. Both studies found that many proteins are tagged and that tagging occurs at locations throughout the protein sequence, as would be expected for activity on damaged mRNAs (37, 38).

On the other hand, investigation of *E. coli* proteins that are tagged with high frequency indicates that there are some sequences prone to generation of nonstop translation complexes (39). For example, in some substrates, tagging occurs with high frequency after runs of rare codons or highly inefficient translation termination sequences (40–42). The mRNA is initially complete in these cases, but ribosome stalling during translation elongation or termination exposes the downstream mRNA to exonucleases, which chew back the mRNA to the leading edge of the ribosome to generate substrates for *trans*-translation (4, 43–45). Exonuclease activity by RNase II can promote cleavage of the mRNA in the A site through an unknown mechanism, but RNase II and the corresponding A-site cleavage are not essential for *trans*-translation on known substrates (46, 47). Redundant nuclease activities may ensure that translation complexes stalled for an extended time are targeted for resolution by *trans*-translation.

In addition to ribosome stalling, errors during translation can lead to *trans*-translation. Suppressor tRNAs and drugs that promote miscoding increase the number of proteins tagged by *trans*-translation, demonstrating that readthrough of the stop codon and frameshifting can result in nonstop translation complexes when there is not an in-frame stop codon downstream (48, 49). The examples described above all result in nonproductive translation complexes, which could sequester ribosomes and limit new protein synthesis. The main purpose of *trans*-translation on these substrates is likely to be release of the ribosomes to maintain protein synthesis capacity.

There is also evidence that *trans*-translation is used to ensure the quality of the protein pool. *trans*-Translation increases on large proteins when *dnaK* is deleted, suggesting that misfolding of the nascent polypeptide might trigger mRNA cleavage to target the nascent polypeptide for proteolysis (50). It is now clear that interactions of the nascent chain in the peptide exit tunnel and communication between ribosome-associated chaperones and the catalytic center of the ribosome can affect the rate of translation (51, 52). Terminally misfolded nascent proteins might be targeted to *trans*-translation to ensure that they are rapidly degraded. It is not yet known whether there is a dedicated pathway for generating nonstop complexes that is triggered by misfolding or whether misfolding slows elongation enough to expose the mRNA to nonspecific exonuclease activity.

Finally, *trans*-translation is used intentionally as part of several regulatory circuits. RNase toxin components of toxin-antitoxin systems such as RelE and MazF cut most mRNAs in the cell, generating a large number of nonstop mRNAs and nonstop translation complexes (53, 54). Toxin activity is used to induce stasis, allowing the cell to conserve resources during severe stress (53, 54). Toxins are also activated in a small percentage of cells under optimal growth conditions to generate persister cells that can survive sudden stresses (55). *E. coli* mutants lacking *trans*-translation

activity are defective in recovery from toxin-induced stasis, indicating that resolution of the nonstop translation complexes resulting from toxin activity is important for resuming growth after severe nutritional stress or persistence (56, 57). Individual proteins are also targeted for *trans*-translation through truncation of the cognate mRNAs. Nuclease cleavage sites or transcriptional terminators 5′ of the stop codon have been found in some *arfA* and *kinA* genes (58–60). Translation of these genes results in proteins that are rapidly degraded unless *trans*-translation is impaired, making the encoded protein activity dependent on the state of *trans*-translation. The *arfA* example is described in more detail below. *trans*-Translation is used by LacI in *E. coli* to prevent excess protein accumulation (61). At high concentrations, LacI binds within the 3′ end of its own gene. LacI binding to this site blocks transcription elongation and generates a nonstop mRNA, thereby targeting all newly expressed LacI for proteolysis. The use of *trans*-translation in regulatory circuits may be important for individual species or behaviors, but the evolutionary conservation of *trans*-translation is almost certainly due to the ability to maintain the protein synthesis capacity of the cell.

PHYSIOLOGY OF AND ALTERNATIVES TO *trans*-TRANSLATION

Genes encoding tmRNA (*ssrA*) and SmpB (*smpB*) have been identified in all sequenced bacterial species, including those with severely reduced genomes (62). This conservation suggests that *trans*-translation confers a selective advantage in all environments that support bacterial life. In fact, tmRNA and SmpB have been shown to be essential in several species, including *Neisseria gonorrhoeae*, *Shigella flexneri*, *Helicobacter pylori*, and *Mycobacterium tuberculosis* (63–66). Saturating genome-wide mutagenesis experiments suggest that tmRNA and SmpB are also required for viability in *Haemophilus influenzae*, *Mycoplasma genitalium*, and *Staphylococcus aureus* (67–69). In other bacteria, tmRNA can be deleted with widely varying consequences. In some species, phenotypes of mutants lacking *trans*-translation activity are severe, including defects in virulence (*Salmonella enterica*, *Yersinia pestis*, *Francisella tularensis*, and *Streptococcus pneumoniae*), symbiosis (*Bradyrhizobium japonicum*), and cell cycle control (*C. crescentus*) (38, 70–75). However, *E. coli* and *Bacillus subtilis* mutants that lack *trans*-translation have relatively mild phenotypes, such as increased antibiotic susceptibility and stress response defects (48, 76–78). Recent discoveries have shown that most or all species that do not require *trans*-translation have backup systems that resolve nonstop translation complexes when *trans*-translation activity is not available.

ArfA

On the basis of the evolutionary conservation of *trans*-translation and the differences in phenotypes between *E. coli* and species in which tmRNA is essential, Chadani and coworkers performed a screen for genes that are essential in strains deleted for *ssrA* (79). They identified a single gene, *arfA*, and showed that the ArfA protein can promote hydrolysis of peptidyl-tRNA on nonstop translation complexes in an *in vitro* translation reaction. Release of the ribosomes by ArfA requires RF2, suggesting that ArfA recognizes the empty mRNA channel and recruits RF2 to hydrolyze the peptidyl-tRNA (Fig. 1) (80). However, it is not yet clear how ArfA recognizes nonstop translation complexes.

ArfA is a true backup system for *trans*-translation in that it is active only when *trans*-translation activity is not available. The *arfA* mRNA in *E. coli* includes a cleavage site for RNase III before the stop codon and is efficiently cut by RNase III to produce a nonstop mRNA (58). Translation of *arfA* when *trans*-translation is active results in a tagged ArfA protein that is rapidly degraded. When *ssrA* is deleted, stable and active ArfA protein is produced. Presumably, regulation by *trans*-translation allows ArfA to release nonstop complexes only under physiological conditions where *trans*-translation is inactive or saturated. Most *arfA* genes from other species encode the RNase III cleavage site, but some use a transcriptional terminator before the stop codon to produce a nonstop mRNA (60). Thus, regulation of ArfA by *trans*-translation is conserved even though the mechanism for producing the nonstop mRNA is not.

Genetic experiments with *arfA* suggest that release of ribosomes from nonstop translation complexes is essential in *E. coli* and related species. In *E. coli*, deletion of *arfA* and *ssrA* is synthetically lethal (79). In contrast, *ssrA* is essential in *S. flexneri*, which does not have *arfA*, but *ssrA* can be deleted in *S. flexneri* cells that are engineered to express *E. coli arfA* (64). *arfA* genes have been identified in only a subset of beta- and gammaproteobacteria and a few other species (60). However, the small size of *arfA* makes bioinformatic identification in distantly related bacteria difficult. The presence of *arfA* does not ensure that *trans*-translation is dispensable. *N. gonorrhoeae* has an *arfA* gene, and yet *trans*-translation is essential. The *N. gonorrhoeae arfA* is active when expressed in *E. coli* (60), so either *arfA* is not expressed in *N. gonorrhoeae* or its activity is not sufficient to support viability in the absence of *trans*-translation.

ArfB

A second alternative system, ArfB, was discovered in a multicopy suppressor screen for genes that allowed *E. coli* to survive without tmRNA or ArfA (Fig. 1) (81). Peptidyl-tRNA hydrolase (Pth) activity had been predicted for ArfB on the basis of the presence of a GGQ motif common to release factors and peptidyl-tRNA hydrolases (82). In fact, purified ArfB specifically hydrolyzes peptidyl-tRNA in nonstop translation complexes *in vitro* (81, 83). Structural studies show that ArfB recognizes nonstop complexes in a manner similar to that of SmpB-tmRNA: a C-terminal helix of ArfB extends into the empty mRNA channel, and residues in this helix make contacts with 16S rRNA that are important for activity (Fig. 2) (84, 85). The physiological role of ArfB in *E. coli* is not clear. The chromosomal copy of a*rfB* will not support growth of *E. coli* in the absence of tmRNA and ArfA, and *ssrA* is essential in *S. flexneri* even though *arfB* is present (64, 79). Either ArfB is reserved for special conditions in these species or the availability of ArfA has made ArfB redundant and control of its expression has been lost. In contrast, ArfB in *Caulobacter crescentus* is functional in its chromosomal context and allows cells to survive without *trans*-translation. The *C. crescentus arfB* gene was identified in transposon sequencing (Tn-Seq) experiments as a gene that is essential in cells lacking *ssrA* but not in wild-type cells (H. A. Feaga and K. C. Keiler, unpublished data). ArfB homologs are widely distributed throughout bacterial species. No regulation of ArfB by *trans*-translation has been identified, so, unlike ArfA, ArfB may provide a constitutive, low level of resolution activity that becomes significant only when *trans*-translation is saturated or inactivated.

Mitochondria also have an ArfB homolog, which is named

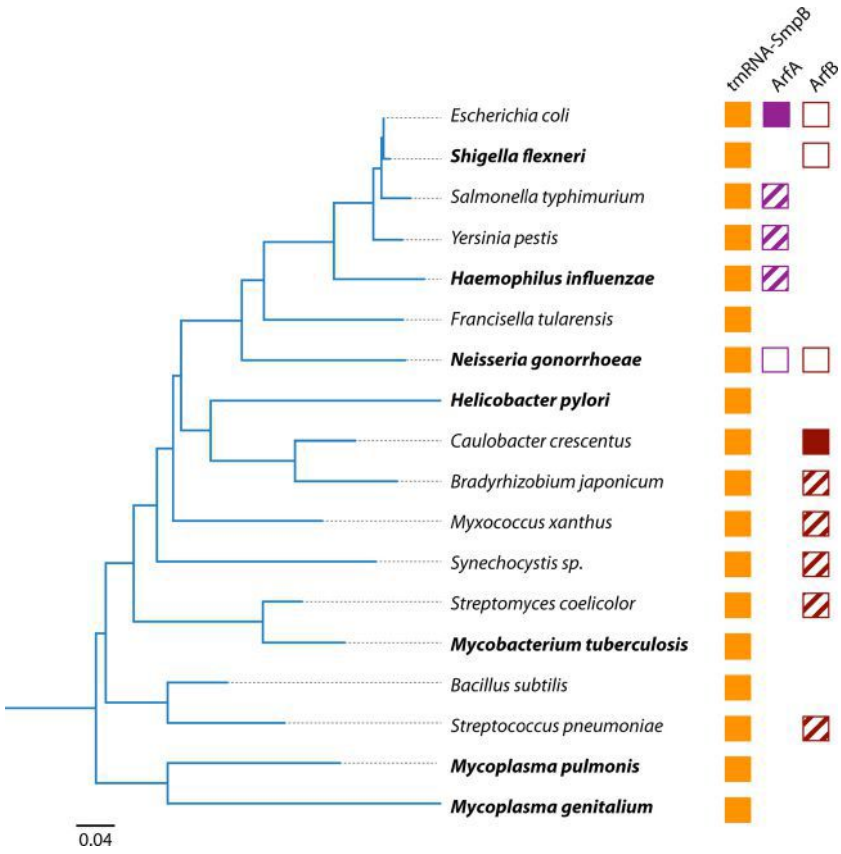


FIG 3 Phylogenetic distribution of *trans*-translation, ArfA, and ArfB. Species in which the phenotype of deleting *ssrA* or *smpB* is known are shown on a phylogenetic tree based on 16S rRNA sequences. Bold names indicate species in which *ssrA* or *smpB* is essential. The presence of genes encoding tmRNA-SmpB, ArfA, and ArfB is shown. For ArfA and ArfB, a filled box indicates that the system is sufficient to maintain viability in the absence of tmRNA-SmpB, an empty box indicates that the system is not sufficient to maintain viability in the absence of tmRNA-SmpB, and a hashed box indicates that it is not yet known whether the system is sufficient to maintain viability. *Salmonella typhimurium*, *Salmonella enterica* serovar Typhimurium.

ICT1 (85, 87). ICT1 hydrolyzes peptidyl-tRNA on the ribosome, and this activity is essential for human cells (87). ArfB and ICT1 both contain an N-terminal GGQ motif and a C-terminal $R(X_3)K(X_6)K(X_2)R$ motif that are required for peptidyl-tRNA hydrolase activity (85). As in bacteria, transcription and translation are performed in a single compartment in mitochondria, so ICT1 may serve to release nonstop complexes and maintain protein synthesis capacity in these organelles. tmRNA has been identified in organelles of some primitive eukaryotes but is not retained in metazoans (62, 88). It appears that most eukaryotic mitochondria kept ArfB and dispensed with *trans*-translation, whereas all bacteria retained *trans*-translation.

The discoveries of ArfA and ArfB have important implications for understanding the role of *trans*-translation and the consequences of nonstop translation complexes. With the exception of *B. subtilis* and *F. tularensis*, all species in which *ssrA* or *smpB* has been deleted encode either ArfA or ArfB (Fig. 3). Moreover, in all cases that have been tested, the ArfA or ArfB backup system becomes essential when *ssrA* is deleted. Therefore, at least one mechanism to resolve nonstop complexes may be required for viability in most or all bacteria. Investigation of unknown alternative resolution mechanisms in *B. subtilis* and *F. tularensis* would test how universal this requirement is. Some nonstop translation complexes may be resolved by "drop-off," dissociation of the peptidyl-tRNA from the ribosome followed by hydrolysis of the free peptidyl-tRNA by peptidyl-tRNA hydrolase (Pth). Drop-off occurs with some nascent chains of two to five amino acids, but longer chains have not been shown to dissociate without prior peptidyl-tRNA hydrolysis within the ribosome (89, 90). Interactions between the nascent polypeptide and the exit channel may prevent drop-off in most cases. The discoveries of ArfA and ArfB make it clear that drop-off alone cannot support viability for most species in the absence of *trans*-translation.

Why is it that all bacteria use *trans*-translation to resolve nonstop complexes, and some use only *trans*-translation, but none use only ArfA or ArfB? ArfA and ArfB do not completely mimic *trans*-translation, because they do not directly target the nascent polypeptide for proteolysis. Presumably, incomplete proteins released by ArfA or ArfB activity must be recognized and degraded by other proteolytic pathways in the cell. The fate of the mRNA during ArfA and ArfB activity is not yet known. It is likely that *trans*-translation is the preferred pathway because it promotes degradation of the incomplete proteins and damaged mRNAs from nonstop complexes in addition to releasing the stalled ribosomes.

TARGETING *trans*-TRANSLATION FOR ANTIBIOTICS

The *trans*-translation pathway is an attractive target for development of new antibiotics because it is required for viability or vir-

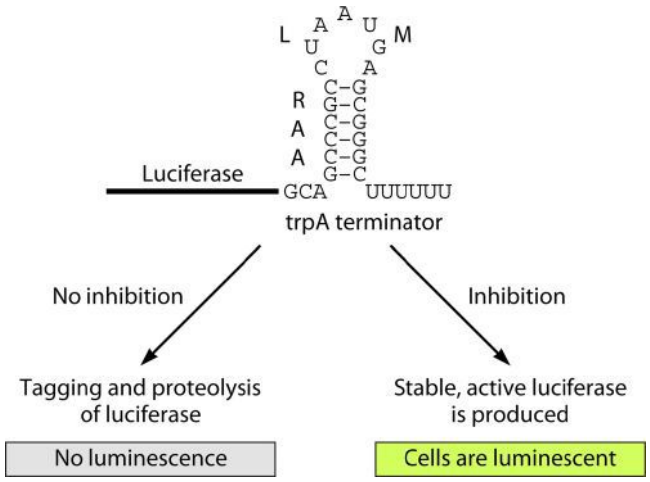


FIG 4 High-throughput screening assay to identify *trans*-translation inhibitors. The reporter contains a gene encoding luciferase with a strong transcriptional terminator inserted before the stop codon, such that transcription results in a nonstop mRNA. *E. coli* cells containing the reporter were screened in high-throughput format to identify compounds that inhibit *trans*-translation. When no inhibitor is present, translation of the nonstop mRNA results in *trans*-translation followed by proteolysis of luciferase, and cells produce no luminescence. Conversely, active luciferase is produced when a *trans*-translation inhibitor is present, resulting in luminescence.

ulence in many pathogenic strains and is not found in metazoans. Therefore, compounds that specifically inhibit *trans*-translation and not translation are likely to be effective for treating infections and yet have low toxicity for host cells. Compounds that inhibit *trans*-translation should kill *M. tuberculosis*, *N. gonorrhoeae*, *S. flexneri*, *H. influenzae*, *S. aureus*, and other species in which *trans*-translation is essential and could also prevent infection by *S. enterica*, *Y. pestis*, *F. tularensis*, *S. pneumoniae*, and other species that require *trans*-translation for virulence. Compounds that inhibit ArfA and ArfB in addition to the effect of *trans*-translation may have antibacterial activity against all species.

Several cell-based assays for *trans*-translation activity have been described, and they all have the same basic construction (91). A strong transcriptional terminator is inserted before the stop codon of a reporter gene, such as the *luc* gene encoding luciferase (Fig. 4). Because the reporter protein is made from a nonstop mRNA, the protein is tagged and degraded if there is no inhibitor present. In the presence of an inhibitor, active reporter protein is produced. In principle, such assays could be used for screening any compound library for inhibitors.

The results of one high-throughput screening (HTS) investigation of inhibitors of *trans*-translation have been reported (36, 91). Several small molecules identified by HTS inhibit *trans*-translation but not translation *in vitro*. Growth inhibition assays with these compounds showed that they have broad-spectrum antibacterial activity (91). One compound, KKL-35, has a MIC of <2 μg/ml against pathogenic strains of *F. tularensis*, *Y. pestis*, *B. anthracis*, *Burkholderia mallei*, and *S. aureus* (K. Keiler, unpublished data). For KKL-35, growth inhibition of *E. coli* was antagonized by low concentrations of puromycin, a drug that can release nonstop translation complexes by hydrolyzing peptidyl-tRNA on the ribosome (91). Likewise, growth inhibition of *S. flexneri* was antagonized by overexpression of *E. coli* ArfA. These results suggest that KKL-35 inhibits growth by preventing release of nonstop translation complexes. Although many additional tests are required to determine if KKL-35 can be developed into a new antibiotic, it is clear that *trans*-translation and alternate pathways to resolve nonstop translation complexes are druggable. These pathways should be considered a prime target for further antibiotic development.

ACKNOWLEDGMENT

K.C.K. and H.A.F. are supported by National Institutes of Health grant GM68720 (to K.C.K.).

REFERENCES

1. **Kahvejian A, Svitkin YV, Sukarieh R, M'Boutchou MN, Sonenberg N.** 2005. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. Genes Dev. **19:**104–113. http://dx.doi.org/10.1101/gad.1262905.
2. **Doma MK, Parker R.** 2007. RNA quality control in eukaryotes. Cell **131:**660–668. http://dx.doi.org/10.1016/j.cell.2007.10.041.
3. **Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU.** 2005. Initiation of protein synthesis in bacteria. Microbiol. Mol. Biol. Rev. **69:**101–123. http://dx.doi.org/10.1128/MMBR.69.1.101-123.2005.
4. **Hayes CS, Sauer RT.** 2003. Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. Mol. Cell **12:**903–911. http://dx.doi.org/10.1016/S1097-2765(03)00385-X.
5. **Sunohara T, Jojima K, Yamamoto Y, Inada T, Aiba H.** 2004. Nascent-peptide-mediated ribosome stalling at a stop codon induces mRNA cleavage resulting in nonstop mRNA that is recognized by tmRNA. RNA **10:**378–386. http://dx.doi.org/10.1261/rna.5169404.
6. **Sunohara T, Jojima K, Tagami H, Inada T, Aiba H.** 2004. Ribosome stalling during translation elongation induces cleavage of mRNA being translated in *Escherichia coli*. J. Biol. Chem. **279:**15368–15375. http://dx.doi.org/10.1074/jbc.M312805200.
7. **Ito K, Chadani Y, Nakamori K, Chiba S, Akiyama Y, Abo T.** 2011. Nascentome analysis uncovers futile protein synthesis in *Escherichia coli*. PLoS One **6:**e28413. http://dx.doi.org/10.1371/journal.pone.0028413.
8. **Komine Y, Kitabatake M, Yokogawa T, Nishikawa K, Inokuchi H.** 1994. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **91:**9223–9227. http://dx.doi.org/10.1073/pnas.91.20.9223.
9. **Ushida C, Himeno H, Watanabe T, Muto A.** 1994. tRNA-like structures in 10Sa RNAs of *Mycoplasma capricolum* and *Bacillus subtilis*. Nucleic Acids Res. **22:**3392–3396. http://dx.doi.org/10.1093/nar/22.16.3392.
10. **Felden B, Himeno H, Muto A, McCutcheon JP, Atkins JF, Gesteland RF.** 1997. Probing the structure of the Escherichia coli 10Sa RNA (tmRNA). RNA **3:**89–103.
11. **Tu GF, Reid GE, Zhang JG, Moritz RL, Simpson RJ.** 1995. C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. J. Biol. Chem. **270:**9322–9326. http://dx.doi.org/10.1074/jbc.270.16.9322.
12. **Williams KP, Bartel DP.** 1996. Phylogenetic analysis of tmRNA secondary structure. RNA **2:**1306–1310.
13. **Karzai AW, Susskind MM, Sauer RT.** 1999. SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA). EMBO J. **18:**3793–3799. http://dx.doi.org/10.1093/emboj/18.13.3793.
14. **Bessho Y, Shibata R, Sekine S-I, Murayama K, Higashijima K, Hori-Takemoto C, Shirouzu M, Kuramitsu S, Yokoyama S.** 2007. Structural basis for functional mimicry of long-variable-arm tRNA by transfer-messenger RNA. Proc. Natl. Acad. Sci. U. S. A. **104:**8293–8298. http://dx.doi.org/10.1073/pnas.0700402104.
15. **Shimizu Y, Ueda T.** 2006. SmpB triggers GTP hydrolysis of elongation factor Tu on ribosomes by compensating for the lack of codon-anticodon interaction during trans-translation initiation. J. Biol. Chem. **281:**15987–15996. http://dx.doi.org/10.1074/jbc.M512165200.
16. **Barends S, Wower J, Kraal B.** 2000. Kinetic parameters for tmRNA binding to alanyl-tRNA synthetase and elongation factor Tu from Escherichia coli. Biochemistry **39:**2652–2658. http://dx.doi.org/10.1021/bi992439d.
17. **Rudinger-Thirion J, Giegé R, Felden B.** 1999. Aminoacylated tmRNA from *Escherichia coli* interacts with prokaryotic elongation factor Tu. RNA **5:**989–992. http://dx.doi.org/10.1017/S135583829999101X.
18. **Ivanova N, Pavlov MY, Bouakaz E, Ehrenberg M, Holmberg Schiavone**

L. 2005. Mapping the interaction of SmpB with ribosomes by footprinting of ribosomal RNA. Nucleic Acids Res. **33:**3529–3539. http://dx.doi.org/10.1093/nar/gki666.
19. **Kurita D, Sasaki R, Muto A, Himeno H.** 2007. Interaction of SmpB with ribosome from directed hydroxyl radical probing. Nucleic Acids Res. **35:**7248–7255. http://dx.doi.org/10.1093/nar/gkm677.
20. **Sundermeier TR, Karzai AW.** 2007. Functional SmpB-ribosome interactions require tmRNA. J. Biol. Chem. **282:**34779–34786. http://dx.doi.org/10.1074/jbc.M707256200.
21. **Valle M, Gillet R, Kaur S, Henne A, Ramakrishnan V, Frank J.** 2003. Visualizing tmRNA entry into a stalled ribosome. Science **300:**127–130. http://dx.doi.org/10.1126/science.1081798.
22. **Ramrath DJF, Yamamoto H, Rother K, Wittek D, Pech M, Mielke T, Loerke J, Scheerer P, Ivanov P, Teraoka Y, Shpanchenko O, Nierhaus KH, Spahn CMT.** 2012. The complex of tmRNA-SmpB and EF-G on translocating ribosomes. Nature **485:**526–529. http://dx.doi.org/10.1038/nature11006.
23. **Camenares D, Dulebohn DP, Svetlanov A, Karzai AW.** 2013. Active and accurate trans-translation requires distinct determinants in the C-terminal tail of SmpB protein and the mRNA-like domain of transfer messenger RNA (tmRNA). J. Biol. Chem. **288:**30527–30542. http://dx.doi.org/10.1074/jbc.M113.503896.
24. **Keiler KC, Waller PR, Sauer RT.** 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science **271:**990–993. http://dx.doi.org/10.1126/science.271.5251.990.
25. **Flynn JM, Levchenko I, Seidel M, Wickner SH, Sauer RT, Baker TA.** 2001. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proc. Natl. Acad. Sci. U. S. A. **98:**10584–10589. http://dx.doi.org/10.1073/pnas.191375298.
26. **Gottesman S, Roche E, Zhou Y, Sauer RT.** 1998. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. Genes Dev. **12:**1338–1347. http://dx.doi.org/10.1101/gad.12.9.1338.
27. **Herman C, Thévenet D, Bouloc P, Walker GC, D'Ari R.** 1998. Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). Genes Dev. **12:**1348–1355. http://dx.doi.org/10.1101/gad.12.9.1348.
28. **Choy JS, Aung LL, Karzai AW.** 2007. Lon protease degrades transfer-messenger RNA-tagged proteins. J. Bacteriol. **189:**6564–6571. http://dx.doi.org/10.1128/JB.00860-07.
29. **Richards J, Mehta P, Karzai AW.** 2006. RNase R degrades non-stop mRNAs selectively in an SmpB-tmRNA-dependent manner. Mol. Microbiol. **62:**1700–1712. http://dx.doi.org/10.1111/j.1365-2958.2006.05472.x.
30. **Mehta P, Richards J, Karzai AW.** 2006. tmRNA determinants required for facilitating nonstop mRNA decay. RNA **12:**2187–2198. http://dx.doi.org/10.1261/rna.247706.
31. **Yamamoto Y, Sunohara T, Jojima K, Inada T, Aiba H.** 2003. SsrA-mediated trans-translation plays a role in mRNA quality control by facilitating degradation of truncated mRNAs. RNA **9:**408–418. http://dx.doi.org/10.1261/rna.2174803.
32. **Neubauer C, Gillet R, Kelley AC, Ramakrishnan V.** 2012. Decoding in the absence of a codon by tmRNA and SmpB in the ribosome. Science **335:**1366–1369. http://dx.doi.org/10.1126/science.1217039.
33. **Kurita D, Muto A, Himeno H.** 2010. Role of the C-terminal tail of SmpB in the early stage of trans-translation. RNA **16:**980–990. http://dx.doi.org/10.1261/rna.1916610.
34. **Sundermeier TR, Dulebohn DP, Cho HJ, Karzai AW.** 2005. A previously uncharacterized role for small protein B (SmpB) in transfer messenger RNA-mediated trans-translation. Proc. Natl. Acad. Sci. U. S. A. **102:**2316–2321. http://dx.doi.org/10.1073/pnas.0409694102.
35. **Moore SD, Sauer RT.** 2005. Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*. Mol. Microbiol. **58:**456–466. http://dx.doi.org/10.1111/j.1365-2958.2005.04832.x.
36. **Ivanova N, Pavlov MY, Felden B, Ehrenberg M.** 2004. Ribosome rescue by tmRNA requires truncated mRNAs. J. Mol. Biol. **338:**33–41. http://dx.doi.org/10.1016/j.jmb.2004.02.043.
37. **Hong S-J, Lessner FH, Mahen EM, Keiler KC.** 2007. Proteomic identification of tmRNA substrates. Proc. Natl. Acad. Sci. U. S. A. **104:**17128–17133. http://dx.doi.org/10.1073/pnas.0707671104.
38. **Svetlanov A, Puri N, Mena P, Koller A, Karzai AW.** 2012. *Francisella tularensis* tmRNA system mutants are vulnerable to stress, avirulent in mice, and provide effective immune protection. Mol. Microbiol. **85:**122–141. http://dx.doi.org/10.1111/j.1365-2958.2012.08093.x.
39. **Roche ED, Sauer RT.** 2001. Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons. J. Biol. Chem. **276:**28509–28515. http://dx.doi.org/10.1074/jbc.M103864200.
40. **Roche ED, Sauer RT.** 1999. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. EMBO J. **18:**4579–4589. http://dx.doi.org/10.1093/emboj/18.16.4579.
41. **Hayes CS, Bose B, Sauer RT.** 2002. Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. J. Biol. Chem. **277:**33825–33832. http://dx.doi.org/10.1074/jbc.M205405200.
42. **Hayes CS, Bose B, Sauer RT.** 2002. Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **99:**3440–3445. http://dx.doi.org/10.1073/pnas.052707199.
43. **Li X, Yokota T, Ito K, Nakamura Y, Aiba H.** 2007. Reduced action of polypeptide release factors induces mRNA cleavage and tmRNA tagging at stop codons in *Escherichia coli*. Mol. Microbiol. **63:**116–126. http://dx.doi.org/10.1111/j.1365-2958.2006.05498.x.
44. **Li X, Hirano R, Tagami H, Aiba H.** 2006. Protein tagging at rare codons is caused by tmRNA action at the 3′ end of nonstop mRNA generated in response to ribosome stalling. RNA **12:**248–255. http://dx.doi.org/10.1261/rna.2212606.
45. **Garza-Sánchez F, Gin JG, Hayes CS.** 2008. Amino acid starvation and colicin D treatment induce A-site mRNA cleavage in *Escherichia coli*. J. Mol. Biol. **378:**505–519. http://dx.doi.org/10.1016/j.jmb.2008.02.065.
46. **Garza-Sánchez F, Shoji S, Fredrick K, Hayes CS.** 2009. RNase II is important for A-site mRNA cleavage during ribosome pausing. Mol. Microbiol. **73:**882–897. http://dx.doi.org/10.1111/j.1365-2958.2009.06813.x.
47. **Janssen BD, Garza-Sánchez F, Hayes CS.** 2013. A-site mRNA cleavage is not required for tmRNA-mediated ssrA-peptide tagging. PLoS One **8:**e81319. http://dx.doi.org/10.1371/journal.pone.0081319.
48. **Abo T, Ueda K, Sunohara T, Ogawa K, Aiba H.** 2002. SsrA-mediated protein tagging in the presence of miscoding drugs and its physiological role in *Escherichia coli*. Genes Cells **7:**629–638. http://dx.doi.org/10.1046/j.1365-2443.2002.00549.x.
49. **Ueda K, Yamamoto Y, Ogawa K, Abo T, Inokuchi H, Aiba H.** 2002. Bacterial SsrA system plays a role in coping with unwanted translational readthrough caused by suppressor tRNAs. Genes Cells **7:**509–519. http://dx.doi.org/10.1046/j.1365-2443.2002.00537.x.
50. **Ruhe ZC, Hayes CS.** 2010. The N-terminus of GalE induces tmRNA activity in *Escherichia coli*. PLoS One **5:**e15207. http://dx.doi.org/10.1371/journal.pone.0015207.
51. **Kramer G, Boehringer D, Ban N, Bukau B.** 2009. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat. Struct. Mol. Biol. **16:**589–597. http://dx.doi.org/10.1038/nsmb.1614.
52. **Ito K, Chiba S.** 2013. Arrest peptides: cis-acting modulators of translation. Annu. Rev. Biochem. **82:**171–202. http://dx.doi.org/10.1146/annurev-biochem-080211-105026.
53. **Pedersen K, Zavialov AV, Pavlov MY, Elf J, Gerdes K, Ehrenberg M.** 2003. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell **112:**131–140. http://dx.doi.org/10.1016/S0092-8674(02)01248-5.
54. **Zhang Y, Zhang J, Hoeflich KP, Ikura M, Qing G, Inouye M.** 2003. MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. Mol. Cell **12:**913–923. http://dx.doi.org/10.1016/S1097-2765(03)00402-7.
55. **Maisonneuve E, Shakespeare LJ, Jørgensen MG, Gerdes K.** 2011. Bacterial persistence by RNA endonucleases. Proc. Natl. Acad. Sci. U. S. A. **108:**13206–13211. http://dx.doi.org/10.1073/pnas.1100186108.
56. **Christensen SK, Gerdes K.** 2003. RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. Mol. Microbiol. **48:**1389–1400. http://dx.doi.org/10.1046/j.1365-2958.2003.03512.x.
57. **Christensen SK, Pedersen K, Hansen FG, Gerdes K.** 2003. Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. J. Mol. Biol. **332:**809–819. http://dx.doi.org/10.1016/S0022-2836(03)00922-7.
58. **Garza-Sánchez F, Schaub RE, Janssen BD, Hayes CS.** 2011. tmRNA regulates synthesis of the ArfA ribosome rescue factor. Mol. Microbiol. **80:**1204–1219. http://dx.doi.org/10.1111/j.1365-2958.2011.07638.x.
59. **Kobayashi K, Kuwana R, Takamatsu H.** 2008. kinA mRNA is missing a stop codon in the undomesticated *Bacillus subtilis* strain ATCC 6051. Microbiology **154:**54–63. http://dx.doi.org/10.1099/mic.0.2007/011783-0.

60. **Schaub RE, Poole SJ, Garza-Sánchez F, Benbow S, Hayes CS.** 2012. Proteobacterial ArfA peptides are synthesized from non-stop messenger RNAs. J. Biol. Chem. **287:**29765–29775. http://dx.doi.org/10.1074/jbc.M112.374074.
61. **Abo T, Inada T, Ogawa K, Aiba H.** 2000. SsrA-mediated tagging and proteolysis of LacI and its role in the regulation of lac operon. EMBO J. **19:**3762–3769. http://dx.doi.org/10.1093/emboj/19.14.3762.
62. **Gueneau de Novoa P, Williams KP.** 2004. The tmRNA website: reductive evolution of tmRNA in plastids and other endosymbionts. Nucleic Acids Res. **32:**D104–D108. http://dx.doi.org/10.1093/nar/gkh102.
63. **Huang C, Wolfgang MC, Withey J, Koomey M, Friedman DI.** 2000. Charged tmRNA but not tmRNA-mediated proteolysis is essential for *Neisseria gonorrhoeae* viability. EMBO J. **19:**1098–1107. http://dx.doi.org/10.1093/emboj/19.5.1098.
64. **Ramadoss NS, Zhou X, Keiler KC.** 2013. tmRNA is essential in *Shigella flexneri*. PLoS One **8:**e57537. http://dx.doi.org/10.1371/journal.pone.0057537.
65. **Personne Y, Parish T.** 2014. *Mycobacterium tuberculosis* possesses an unusual tmRNA rescue system. Tuberculosis **94:**34–42. http://dx.doi.org/10.1016/j.tube.2013.09.007.
66. **Thibonnier M, Thiberge J-M, De Reuse H.** 2008. Trans-translation in *Helicobacter pylori*: essentiality of ribosome rescue and requirement of protein tagging for stress resistance and competence. PLoS One **3:**e3810. http://dx.doi.org/10.1371/journal.pone.0003810.
67. **Fey PD, Endres JL, Yajjala VK, Widhelm TJ, Boissy RJ, Bose JL, Bayles KW.** 2013. A genetic resource for rapid and comprehensive phenotype screening of nonessential *Staphylococcus aureus* genes. mBio **4:**e00537–12. http://dx.doi.org/10.1128/mBio.00537-12.
68. **Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, Harrison M, Burgis TA, Lockyer M, Garcia-Lara J, Foster SJ, Pleasance SJ, Peters SE, Maskell DJ, Charles IG.** 2009. Comprehensive identification of essential *Staphylococcus aureus* genes using transposon-mediated differential hybridisation (TMDH). BMC Genomics **10:**291. http://dx.doi.org/10.1186/1471-2164-10-291.
69. **Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, Smith HO, Venter JC.** 2006. Essential genes of a minimal bacterium. Proc. Natl. Acad. Sci. U. S. A. **103:**425–430. http://dx.doi.org/10.1073/pnas.0510013103.
70. **Julio SM, Heithoff DM, Mahan MJ.** 2000. ssrA (tmRNA) plays a role in *Salmonella enterica* serovar *Typhimurium* pathogenesis. J. Bacteriol. **182:**1558–1563. http://dx.doi.org/10.1128/JB.182.6.1558-1563.2000.
71. **Bäumler AJ, Kusters JG, Stojiljkovic I, Heffron F.** 1994. *Salmonella typhimurium* loci involved in survival within macrophages. Infect. Immun. **62:**1623–1630.
72. **Okan NA, Mena P, Benach JL, Bliska JB, Karzai AW.** 2010. The *smpB-ssrA* mutant of *Yersinia pestis* functions as a live attenuated vaccine to protect mice against pulmonary plague infection. Infect. Immun. **78:**1284–1293. http://dx.doi.org/10.1128/IAI.00976-09.
73. **Keiler KC, Shapiro L.** 2003. tmRNA is required for correct timing of DNA replication in *Caulobacter crescentus*. J. Bacteriol. **185:**573–580. http://dx.doi.org/10.1128/JB.185.2.573-580.2003.
74. **Ebeling S, Kündig C, Hennecke H.** 1991. Discovery of a rhizobial RNA that is essential for symbiotic root nodule development. J. Bacteriol. **173:**6373–6382.
75. **Mann B, van Opijnen T, Wang J, Obert C, Wang Y-D, Carter R, McGoldrick DJ, Ridout G, Camilli A, Tuomanen EI, Rosch JW.** 2012. Control of virulence by small RNAs in *Streptococcus pneumoniae*. PLoS Pathog. **8:**e1002788. http://dx.doi.org/10.1371/journal.ppat.1002788.
76. **Muto A, Fujihara A, Ito KI, Matsuno J, Ushida C, Himeno H.** 2000. Requirement of transfer-messenger RNA for the growth of *Bacillus subtilis* under stresses. Genes Cells **5:**627–635. http://dx.doi.org/10.1046/j.1365-2443.2000.00356.x.
77. **Shin J-H, Price CW.** 2007. The SsrA-SmpB ribosome rescue system is important for growth of *Bacillus subtilis* at low and high temperatures. J. Bacteriol. **189:**3729–3737. http://dx.doi.org/10.1128/JB.00062-07.
78. **Munavar H, Zhou Y, Gottesman S.** 2005. Analysis of the *Escherichia coli* Alp phenotype: heat shock induction in *ssrA* mutants. J. Bacteriol. **187:**4739–4751. http://dx.doi.org/10.1128/JB.187.14.4739-4751.2005.
79. **Chadani Y, Ono K, Ozawa S-I, Takahashi Y, Takai K, Nanamiya H, Tozawa Y, Kutsukake K, Abo T.** 2010. Ribosome rescue by *Escherichia coli* ArfA (YhdL) in the absence of trans-translation system. Mol. Microbiol. **78:**796–808. http://dx.doi.org/10.1111/j.1365-2958.2010.07375.x.
80. **Chadani Y, Ito K, Kutsukake K, Abo T.** 2012. ArfA recruits release factor 2 to rescue stalled ribosomes by peptidyl-tRNA hydrolysis in *Escherichia coli*. Mol. Microbiol. **86:**37–50. http://dx.doi.org/10.1111/j.1365-2958.2012.08190.x.
81. **Chadani Y, Ono K, Kutsukake K, Abo T.** 2011. *Escherichia coli* YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways. Mol. Microbiol. **80:**772–785. http://dx.doi.org/10.1111/j.1365-2958.2011.07607.x.
82. **Hayes CS, Keiler KC.** 2010. Beyond ribosome rescue: tmRNA and co-translational processes. FEBS Lett. **584:**413–419. http://dx.doi.org/10.1016/j.febslet.2009.11.023.
83. **Handa Y, Inaho N, Nameki N.** 2011. YaeJ is a novel ribosome-associated protein in *Escherichia coli* that can hydrolyze peptidyl-tRNA on stalled ribosomes. Nucleic Acids Res. **39:**1739–1748. http://dx.doi.org/10.1093/nar/gkq1097.
84. **Gagnon MG, Seetharaman SV, Bulkley D, Steitz TA.** 2012. Structural basis for the rescue of stalled ribosomes: structure of YaeJ bound to the ribosome. Science **335:**1370–1372. http://dx.doi.org/10.1126/science.1217443.
85. **Kogure H, Handa Y, Nagata M, Kanai N, Güntert P, Kubota K, Nameki N.** 9 December 2013. Identification of residues required for stalled-ribosome rescue in the codon-independent release factor YaeJ. Nucleic Acids Res. http://dx.doi.org/10.1093/nar/gkt1280.
86. Reference deleted.
87. **Richter R, Rorbach J, Pajak A, Smith PM, Wessels HJ, Huynen MA, Smeitink JA, Lightowlers RN, Chrzanowska-Lightowlers ZM.** 2010. A functional peptidyl-tRNA hydrolase, ICT1, has been recruited into the human mitochondrial ribosome. EMBO J. **29:**1116–1125. http://dx.doi.org/10.1038/emboj.2010.14.
88. **Keiler KC, Shapiro L, Williams KP.** 2000. tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in *Caulobacter*. Proc. Natl. Acad. Sci. U. S. A. **97:**7778–7783. http://dx.doi.org/10.1073/pnas.97.14.7778.
89. **Cruz-Vera LR, Magos-Castro MA, Zamora-Romo E, Guarneros G.** 2004. Ribosome stalling and peptidyl-tRNA drop-off during translational delay at AGA codons. Nucleic Acids Res. **32:**4462–4468. http://dx.doi.org/10.1093/nar/gkh784.
90. **Gonzalez de Valdivia EI, Isaksson LA.** 2005. Abortive translation caused by peptidyl-tRNA drop-off at NGG codons in the early coding region of mRNA. FEBS J. **272:**5306–5316. http://dx.doi.org/10.1111/j.1742-4658.2005.04926.x.
91. **Ramadoss NS, Alumasa JN, Cheng L, Wang Y, Li S, Chambers BS, Chang H, Chatterjee AK, Brinker A, Engels IH, Keiler KC.** 2013. Small molecule inhibitors of trans-translation have broad-spectrum antibiotic activity. Proc. Natl. Acad. Sci. U. S. A. **110:**10282–10287. http://dx.doi.org/10.1073/pnas.1302816110.

JB Journals.ASM.org

CbbR, the Master Regulator for Microbial Carbon Dioxide Fixation

Andrew W. Dangel, F. Robert Tabita

Department of Microbiology, The Ohio State University, Columbus, Ohio, USA

Biological carbon dioxide fixation is an essential and crucial process catalyzed by both prokaryotic and eukaryotic organisms to allow ubiquitous atmospheric CO_2 to be reduced to usable forms of organic carbon. This process, especially the Calvin-Bassham-Benson (CBB) pathway of CO_2 fixation, provides the bulk of organic carbon found on earth. The enzyme ribulose 1,5-bisphosphate (RuBP) carboxylase/oxygenase (RubisCO) performs the key and rate-limiting step whereby CO_2 is reduced and incorporated into a precursor organic metabolite. This is a highly regulated process in diverse organisms, with the expression of genes that comprise the CBB pathway (the *cbb* genes), including RubisCO, specifically controlled by the master transcriptional regulator protein CbbR. Many organisms have two or more *cbb* operons that either are regulated by a single CbbR or employ a specific CbbR for each *cbb* operon. CbbR family members are versatile and accommodate and bind many different effector metabolites that influence CbbR's ability to control *cbb* transcription. Moreover, two members of the CbbR family are further post-translationally modified via interactions with other transcriptional regulator proteins from two-component regulatory systems, thus augmenting CbbR-dependent control and optimizing expression of specific *cbb* operons. In addition to interactions with small effector metabolites and other regulator proteins, CbbR proteins may be selected that are constitutively active and, in some instances, elevate the level of *cbb* expression relative to wild-type CbbR. Optimizing CbbR-dependent control is an important consideration for potentially using microbes to convert CO_2 to useful bioproducts.

The CbbR protein is a LysR-type transcriptional regulator (LTTR) that functions to control expression of genes of the CO_2 fixation (*cbb*) operons that specify enzymes of the Calvin-Bassham-Benson (CBB) pathway. CbbR-dependent regulation occurs in diverse organisms, including nonsulfur and sulfur purple bacteria, marine and freshwater chemoautotrophic bacteria, cyanobacteria, methylotrophic bacteria, several varieties of hydrogen-oxidizing bacteria, and different *Pseudomonas*, *Mycobacterium*, and *Clostridium* strains (1–14). In addition, CbbR regulates carbon fixation gene expression in chloroplasts of eukaryotic red algae (15).

For many prokaryotic and eukaryotic organisms, CO_2 is often the sole source of carbon, with the CBB pathway acting as the paramount metabolic pathway that enables such organisms to synthesize all the building blocks and macromolecules required for life. The net goal of the enzymes of the CBB cycle is to provide one triose phosphate molecule as the fundamental reduced form of useable carbon from an intake of three CO_2 molecules. Although several enzymes are dedicated to the CBB pathway, ribulose 1,5-bisphosphate (RuBP) carboxylase/oxygenase (RubisCO) is the enzyme that catalyzes the actual "fixation" of CO_2 onto the ene-diol form of RuBP, resulting in the production of two phosphorylated three-carbon molecules of 3-phosphoglyceric acid (3-PGA). Because this enzyme is a relatively poor catalyst and must contend with CO_2 and O_2, competing for the same active site under aerobic conditions, cells very often compensate by synthesizing huge amounts of RubisCO (e.g., up to 50% of the soluble protein under appropriate growth conditions [16]). Clearly, this is a highly regulated system, and under some physiological conditions, especially in bacteria, it is necessary for the organism to either upregulate or downregulate expression of genes of the CBB pathway. Transcriptional control of the *cbb* genes is thus vital because of the heavy energy demands and the burden of additional protein synthesis placed on the cell by CO_2 assimilation. In bacteria, the master regulator protein in all cases is CbbR.

Like all LTTR family members, CbbR binds as a tetramer to the promoter region of the *cbb* operon. The generalized consensus DNA binding sequence for LTTRs is T-N_{11}-A, and all CbbR proteins interact with this DNA binding motif (17). Typically, each of two DNA binding sites (within approximately 6 to 20 bp of each other) is bound by a CbbR dimer, creating a dimer of dimers (tetramer). Like all LTTRs, the CbbR protein structure is about 300 to 330 amino acids in length and is composed of three functional domains. There is a DNA binding domain (DBD) at the N terminus that binds to the *cbb* promoter region. The LTTR DBD is classified as a winged helix-turn-helix (HTH) motif (18, 19). All four HTH motifs within a CbbR tetramer interact with DNA when bound to the *cbb* promoter. A linker helix domain functions to connect the DBD and the recognition domains (RD) of the protein (RD-I and RD-II; also referred to as effector domains). The linker domain is a 30-amino-acid α-helix that operates as a rigid linker helix to prevent interaction between the DBD and the RD. The linker helix also contributes to dimer formation through coiled-coil interactions (20–26). Figure 1 illustrates a generalized structure for all CbbR proteins. Regions of conservation that distinguish CbbR proteins from other LTTRs are found in the recognition domains and are discussed below with respect to effector interactions.

Accepted manuscript posted online 31 August 2015

Citation Dangel AW, Tabita FR. 2015. CbbR, the master regulator for microbial carbon dioxide fixation. J Bacteriol 197:3488–3498. doi:10.1128/JB.00442-15.

Editor: W. Margolin

Address correspondence to F. Robert Tabita, tabita.1@osu.edu.

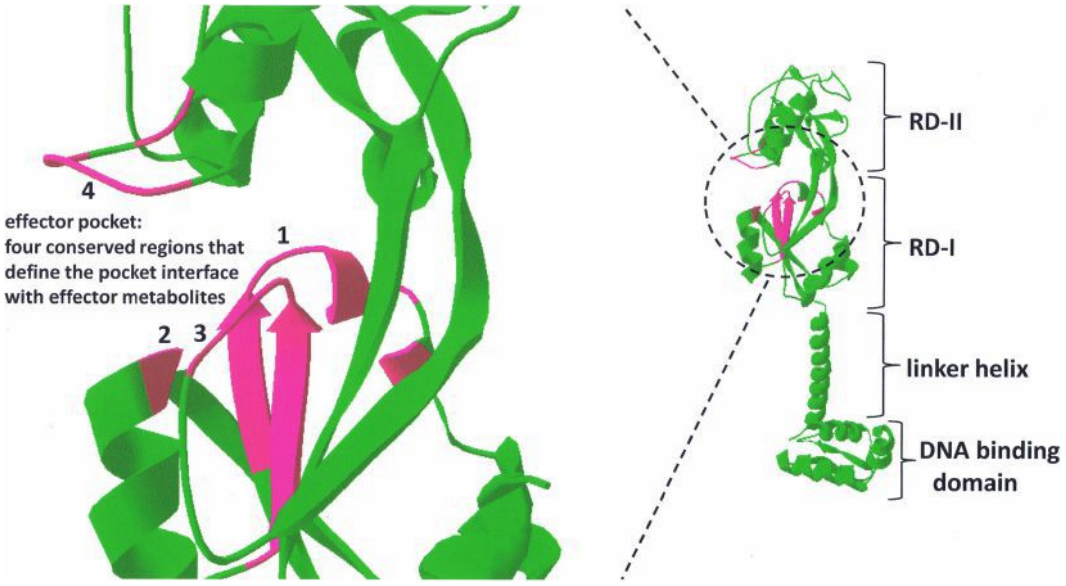


FIG 1 Proposed structure of CbbR. On the right is the generalized ribbon structure of the CbbR monomer based on the structure of the LTTR family member CbnR (20), illustrating the four major domains of the LTTR. On the left is the enlarged structure of the CbbR effector pocket. Four regions of the effector pocket, denoted 1 to 4, are highlighted in magenta, with residues of regions 1 to 4 conserved among CbbR family members. These four conserved regions define the effector pocket and are positioned at the interface between the effector metabolite(s) and CbbR. The conserved amino acid sequences (magenta) for each region are as follows: for region 1, GVVSTAKYFXP; for region 2, NR; for region 3, DLAIMGRPP; and for region 4, REXGSGTR ("X" represents a residue position that is not conserved). The effector pocket conservation also applies to CmpR, CcmR, and QscR, which are CbbR subfamily members. Conservation is high within the four regions of the effector pocket for all CbbRs examined: 81.5% for region 1, 95% for region 2, 86% for region 3, and 97% for region 4. DeepView/Swiss-Pdb Viewer software was used to generate the structural model.

cbbR AND *cbb* GENE ORGANIZATION

The *cbbR* gene, like most LTTR family members, is usually located directly upstream of the *cbb* operon that it regulates but in the opposite orientation. There are some notable exceptions to this general rule of gene organization. For example, *cbbR* of *Rhodospirillum rubrum* is in the same orientation as and adjacent to *cbbM* (encoding form II RubisCO); however, the orientation of *cbbR* is opposite that of the remainder of the *cbb* operon (4, 27). Another interesting exception is provided by *Hydrogenophilus thermoluteolus*; there, *cbbR* is located within a split *cbb* operon in an orientation opposite that of all the *cbb* genes. *Rhodobacter capsulatus* also presents an interesting situation where this organism contains two *cbbR* genes encoding two distinct CbbR proteins ($CbbR_I$ and $CbbR_{II}$) that regulate two separate *cbb* operons, one of which, along with its cognate *cbbR* gene, was apparently derived from a chemoautotrophic ancestor (7, 28). In the chemoautotroph *Hydrogenovibrio marinus*, there also are two *cbbR* genes (*cbbR1* and *cbbRm*) and three *cbb* operons (*cbbLS-1*, *cbbLS-2*, and *cbbM*), with *cbbR1* located upstream and in an orientation opposite that of *cbbLS-1*, while the *cbbRm* gene is located upstream of but in the same orientation as *cbbM* (10, 29–31). The *cbbLS-2* operon contains genes encoding carboxysomes and is expressed under conditions of low CO_2 concentrations, independently of CbbR1 or CbbRm regulation (30, 31). On the other hand, CbbR1 and CbbRm of *H. marinus* may be involved in repressing expression of carboxysome genes contained within the *CbbLS-2* operon at high levels of CO_2 (31). Finally, CbbR has also been shown to regulate the expression of carboxysome genes in *Acidithiobacillus ferrooxidans*, a chemoautotrophic gammaproteobacterium that characteristically grows in acidic environments (3, 13). Notably, the single CbbR from *A. ferrooxidans* regulates four distinct *cbb* operons (3, 13).

While the foregoing represent some very interesting situations where CbbR plays an important physiological role, in addition to regulating *cbb* gene expression, the usual situation is that a single *cbbR* gene is used to exclusively regulate the two major *cbb* operons that are found in most prokaryotic organisms. Many phototrophic and chemoautotrophic organisms contain multiple RubisCO genes, usually encoding distinct form I (CbbLS) and form II (CbbM) enzymes that function to fix CO_2 at low and high CO_2 levels, respectively (32–35). The most thoroughly studied examples where a single CbbR regulates *cbb* operons containing distinct form I and form II RubisCO genes are *Rhodobacter sphaeroides* (5, 36, 37) and *Rhodopseudomonas palustris* (11). Several additional autotrophic bacterial species, including *Ralstonia eutropha* strain H16, contain two *cbb* operons regulated by the product of a single *cbbR* gene. *R. eutropha* has a well-characterized *cbb* regulon where a single *cbbR* gene controls both chromosomal and megaplasmid-borne *cbb* genes (2); however, the RubisCO enzymes encoded by these separate operons are virtually identical (38). *Mycobacterium* sp. strain JC1 DSM 3803 also has two *cbb* operons regulated by one CbbR, but the *cbbR* gene is directly downstream of and in the same orientation as *cbbLS-1* (12, 39).

PHYLOGENETIC RELATEDNESS OF CbbRS AND CLOSELY RELATED LTTRS

Amino acid identities accurately reflect the general relatedness of CbbR proteins from different organisms. Yet there is a striking drift of amino acid homologies among the CbbR family similar to the general lack of amino acid identity within the LTTR family as

a whole. Despite these differences in primary sequence, there is a high degree of structural and conformational similarity of the monomeric, dimeric, and tetrameric states of all LTTR proteins. Regions of residue similarity and identity within the LTTR and CbbR families include the DBD (HTH motif), regions defining the effector pocket, and areas of the protein important for the formation of dimers and tetramers. As determined on the basis of amino acid identities, the CbbR subfamily also includes QscR, CmpR, and CcmR, three closely related LTTRs that are more similar to some CbbRs than some CbbRs are to each other. QscR regulates the expression of two operons involved in the one-carbon serine assimilatory pathway of some methylotrophic bacteria (40, 41). CmpR regulates transcription of operons involved with bicarbonate transport in cyanobacteria (42) and specifically regulates expression of certain genes involved in the CO_2-concentrating mechanism (CCM) (43–46). The CCM allows cyanobacteria to actively transport HCO_3^- into the cytoplasm and then into the carboxysome. Subsequently, carboxysomal carbonic anhydrase catalyzes the conversion of HCO_3^- to CO_2 such that CO_2 becomes highly concentrated in this microcompartment and is readily made available and saturates the active site of RubisCO (47). In *Synechocystis* sp. strain PCC 6803, CmpR activates transcription of the *cmpABCD* operon (high-affinity bicarbonate transporter), but another CbbR-like protein, CcmR, represses expression of the *ndhD3*, *ndhF3*, and *chpY* genes which are required for expression of the inducible high-affinity CO_2 transporter, NDH-1_3 (48, 49).

Total amino acid identities for individual CbbRs range from 22% to 56%, with the majority of identities falling between 35% and 45%. The CbbRs with the highest identities include CbbR of *Bradyrhizobium japonicum* and CbbR of *R. palustris* at 56%, CbbR of *Allochromatium vinosum* and CbbR of *Methylococcus capsulatus* at 55%, CbbR of *R. sphaeroides* and $CbbR_{II}$ of *R. capsulatus* at 54%, and CmpR of *Synechococcus elongatus* PCC 7942 and CcmR of *Synechocystis* PCC 6803 at 54%. The phylogenetic analysis of CbbR proteins is in good agreement with the overall phylogenetic relationship of microorganisms that possess *cbbR* genes.

INTERACTIONS WITH CbbR: THE CASE OF DUELING TRANSCRIPTION FACTORS

There are several studies that show LTTR interactions with RNA polymerase or sigma factors (50–55), but there are few examples of interactions of LTTR proteins (not including CbbRs) with other transcriptional regulators (56, 57). By and large, this is a testament to the ability of LTTRs to independently and adequately regulate their operons in the prokaryotic kingdom. However, in the case of some phototrophic bacteria, regulation of *cbb* expression is much more complex, imposing additional layers of regulation on the energetically costly process of CO_2 assimilation. There are two well-studied systems that illustrate this regulatory complexity: interaction of CbbR with additional (and different) transcription regulators in *R. sphaeroides* and *R. palustris* (58–61).

In *R. sphaeroides*, CbbR interacts with the RegA response regulator, which is part of a global two-component system (RegA/RegB) that controls expression of both the cbb_I and cbb_{II} operons of this organism (58, 62). In addition to the *cbb* regulon, RegA and its cognate sensor kinase, RegB, maintain control over several operons involved with energy-related (redox) metabolism in phototrophic nonsulfur purple bacteria (63, 64). Thus, the response regulator of this two-component system, RegA, binds to multiple sites within the promoter regions of both *cbb* operons of *R. sphaeroides* (37, 65). This scenario is similar to that seen with the *hemA* gene in *R. sphaeroides*, where RegA and FnrL bind the *hemA* promoter at positions where FnrL takes the place of CbbR in the *hemA* promoter and phosphorylation of RegA changes the affinity of RegA for the *hemA* promoter (66). For the cbb_I operon of *R. sphaeroides*, it has been demonstrated that RegA has a higher affinity for the promoter when RegA binding site 3 is present, and site 3 is also necessary for optimal expression of the cbb_I operon *in vivo* (58, 65). It was further shown that RegA greatly enhances the stability of the CbbR/promoter DNA complex and that CbbR must be bound to the *cbb* promoter in order to interact with RegA but that it is not necessary for RegA to be bound to DNA to interact with CbbR (58). This scenario presumably prevents interactions between the two regulator proteins unless CbbR is bound to the *cbb* promoter, pointing to a finely tuned attenuation strategy limiting potential nonspecific interactions between the two proteins. It may be surmised that a deleterious situation would be avoided if the proteins did not interact unless bound to the *cbb* operon promoter. Adding to the complexity, some studies have shown that phosphorylation of RegA greatly increases its DNA binding stability, while other studies have illustrated that phosphorylation of RegA enhances activation of transcription (64). Cross-linking experiments, using a bifunctional binding compound, indicated that RegA and CbbR form a stoichiometric complex, results that were buttressed by gel mobility shift assays that also showed specific interactions between RegA and CbbR-bound DNA (58). In addition, extensive mutational analyses provided evidence that CbbR and RegA interact with each other through their DBDs (59). A model for this rather complex regulatory scenario, which is based on several lines of evidence and provides a likely explanation for how CbbR and RegA interact to regulate *cbb* gene expression and, subsequently, CO_2 fixation in *R. sphaeroides*, is presented in Fig. 2. Also considered in this model is the involvement of various small-molecule effectors that influence CbbR function (Fig. 2) (discussed below).

R. palustris represents an interesting and even more complex regulatory system involving CbbR and several additional protein regulators. While *R. palustris* also contains a Reg-type two-component system (67), it is not clear what this system regulates in this organism, as it apparently does not control photosystem biosynthesis as in *R. sphaeroides* and *R. capsulatus* (J. T. Beatty and F. R. Tabita, unpublished observations), nor is there any evidence to date to indicate that the Reg system controls *cbb* gene expression in *R. palustris* (J. L. Gibson and F. R. Tabita, unpublished observations). However, the regulation of the cbb_I operon in *R. palustris* has proven to be extremely complex, involving a novel three-protein two-component system (11, 68). The regulatory proteins involved, CbbR, CbbRR1, CbbRR2, and CbbSR, are all encoded by genes that are closely juxtaposed within the cbb_I operon region, with *cbbR* divergently transcribed from *cbbRR1*, *cbbRR2*, and *cbbSR*, which are immediately upstream from the *cbbL* and *cbbS* genes encoding the large and small subunits, respectively, of form I RubisCO (11). The CbbSR protein is a large transmembrane sensor kinase which, like many sensor kinases, is capable of autophosphorylation. In addition, CbbSR contains a consensus phosphate acceptor site, with a conserved aspartate-containing motif typical of many response regulators. Studies have shown that, upon autophosphorylation at a specific histidine residue, the phosphate may be transferred to the acceptor site of CbbSR (11, 68). Thus, this large protein basically acts as its own two-compo-

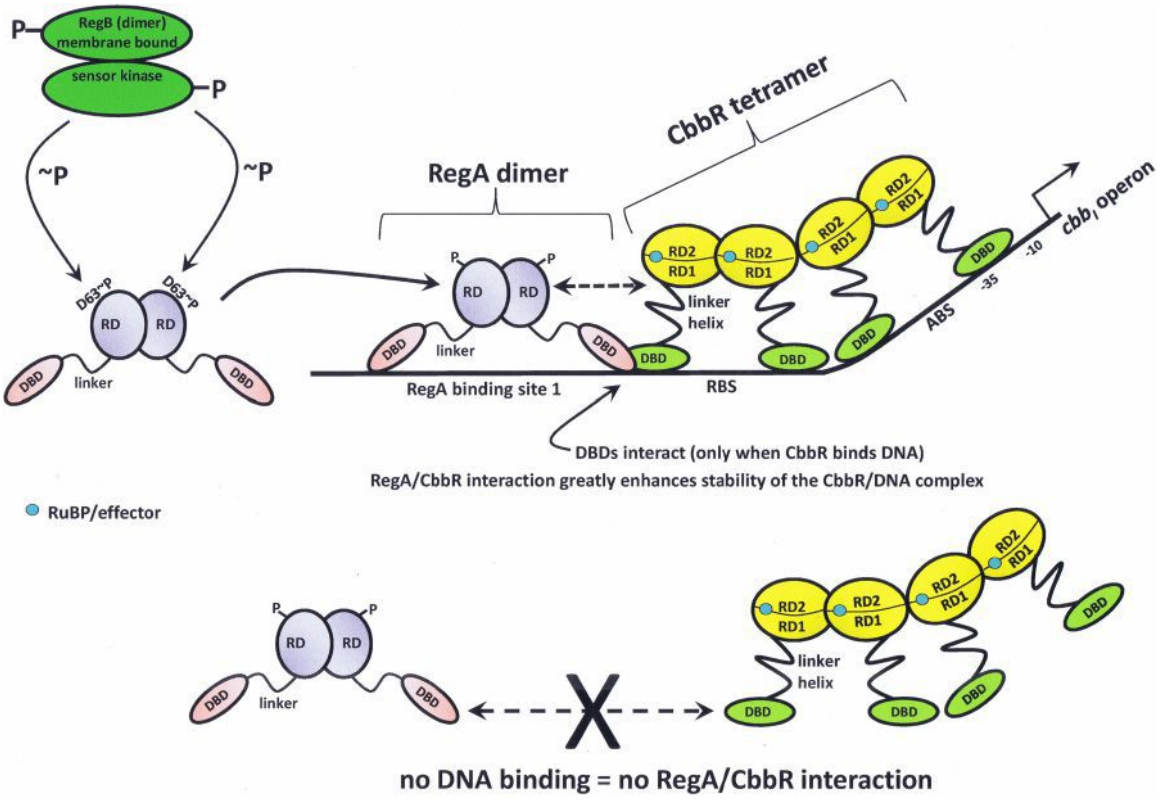


FIG 2 Transcriptional regulation of the *cbb* operons of *Rhodobacter sphaeroides*. CbbR and RegA interact on the cbb_I and cbb_{II} promoters. There are four RegA binding sites upstream of the cbb_I operon and six RegA binding sites upstream of the cbb_{II} operon (37, 65). RegA DNA binding site 1 and the RegA binding site (RBS) overlap upstream of the cbb_I transcriptional start site. The DBDs of CbbR and RegA interact and generate a CbbR/RegA/DNA complex at the cbb_I promoter (59). The interaction of RegA with CbbR greatly increases the stability of the CbbR/DNA complex (58). CbbR does not interact with RegA if CbbR is not bound to the *cbb* promoter (58). RuBP (positive effector) is shown within the effector pocket of a CbbR monomer; the pocket is a small cleft formed between RD-1 and RD-2. Dashed arrows represent interactions with CbbR. ABS, activation binding site. RD-1 and RD-2, recognition domain 1 and recognition domain 2, respectively. DBD, DNA binding domain. RD (RegA), receiver domain. ~P, phosphorylation at residue D63 of RegA. The −10 and −35 regions of the cbb_I promoter are indicated.

nent system. However, CbbSR also catalyzes phosphorylation of both response regulators, CbbRR1 and CbbRR2 (11), the specificity for which is influenced by specific PAS domains on CbbSR (68). Both physiological/genetic and *in vitro* studies indicate that CbbRR1 and CbbRR2 bind to CbbR and influence cbb_I gene transcription (60). In addition, various effector molecules influence these interactions in a concentration-dependent fashion and stabilize CbbR binding to the *cbbLS* promoter. CbbR/CbbRR1 interactions enhance the binding affinity of CbbR to the promoter, and CbbRR1, in concert with effectors ATP, RuBP, and fructose-1,6-bisphosphate, stabilizes the CbbR/promoter complex (69). A model for CbbR/CbbRR1/CbbRR2 interaction proposes that CbbRR2 acts as an antiactivator in the absence of effectors and that CbbRR1, by binding to CbbR and altering the conformation of CbbR, thus prevents CbbR from binding the *cbbLS* promoter (61). The presence of CbbRR1 and effectors negates the effect that CbbRR2 has on CbbR and allows binding of CbbR to the promoter and subsequent expression of the cbb_I operon (61). A model summarizing the information relative to CbbR involvement with all these additional factors in *R. palustris* is presented (Fig. 3). Surface plasmon resonance (SPR) experiments were crucial to providing quantitative results concerning the effects of effectors on protein interactions as well as forming the basis for interpreting the interplay between CbbR, CbbRR1, CbbRR2, and effectors and how these factors all influence the regulation of the *cbbLS* promoter in *R. palustris* (60, 61, 69). Indeed, it is interesting that such a very complex system has been adopted by *R. palustris* to ensure control of the expression of form I RubisCO (*cbbL* and *cbbS*) and the genes of the cbb_I operon. Of note, the cbb_{II} operon, including the gene (*cbbM*) encoding form II RubisCO, does not appear to be controlled by CbbR or by CbbRR1, CbbRR2, or CbbSR (11).

POSTRANSLATIONAL REGULATION OF CbbR FUNCTION: THE ROLE OF EFFECTOR METABOLITES

As prototypical LTTRs, CbbRs require a bound coinducer molecule or effector to activate transcription from the *cbb* promoter (17, 19). Common to most of the members of the LTTR family, the effector usually is an intermediate metabolite of the pathway that is regulated (17, 19). Effector binding occurs in a small cleft formed between RD-I and RD-II (Fig. 1) and is a hallmark of the LTTR family (22, 70–80). A recent study using nondenaturing mass spectrometry has illustrated for the first time that an LTTR tetramer binds four molecules of its effector in a stepwise pattern while bound to DNA (81). Binding of the effector produces a change in the angle at which an LTTR bends the promoter DNA to

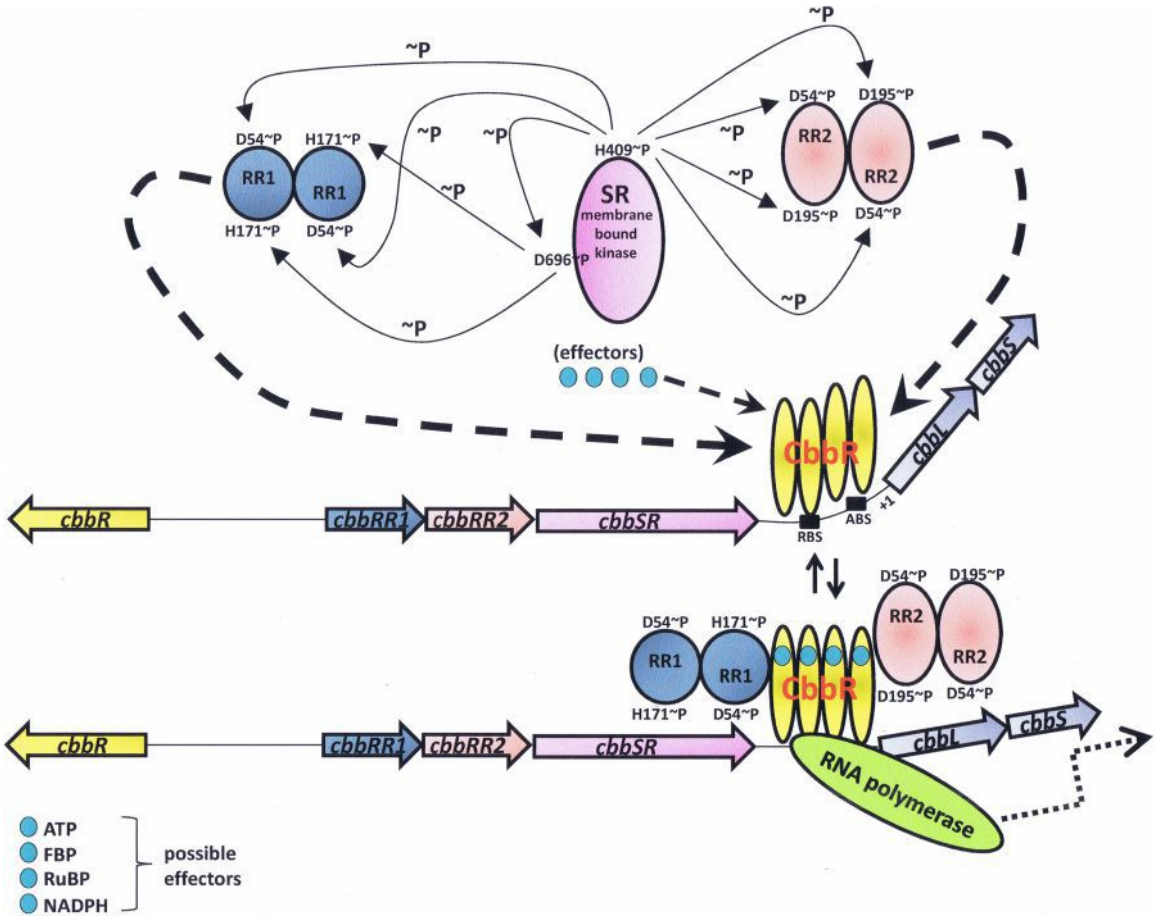


FIG 3 Transcriptional regulation of the cbb_I operon of *Rhodopseudomonas palustris* and the role of the four regulatory factors in the expression of cbb_I under autotrophic conditions. SR is the membrane-bound sensor kinase (CbbSR) that autophosphorylates and catalyzes phosphorylation of the two response regulators, RR1 (CbbRR1) and RR2 (CbbRR2). RR1 and RR2 subsequently interact with CbbR. CbbR binds the cbb_I promoter at the recognition binding site (RBS) and the activation binding site (ABS). Potential positive effectors ATP, FBP, RuBP, and NADPH (68) are shown. Dashed arrows represent interactions with CbbR (60, 61, 69). Oppositely pointing solid arrows represent reversible interactions with CbbR. Dotted arrows represent transcriptional activation. Relaxation of the bend angle that CbbR imposes on the cbb_I promoter is brought about by effector binding and precedes transcription. ~P, phosphorylation at specific residues of CbbSR, CbbRR1, and CbbRR2 (68). The transcriptional start site is denoted +1.

which it is bound (17, 19). For most scenarios, this change in DNA bend angle initiates contact or alters contact with the RNA polymerase, activating transcription. Effector binding to the LTTR can also inhibit transcription (19, 82). Studies illustrating DNA bend angle modification or initiation of transcription brought about by effector binding have been reported for some CbbRs (8, 76, 82, 83). The effector molecules may be different for each CbbR, depending on the organism. Figure 4 illustrates the relaxation of the *cbb* promoter bend angle imposed by CbbR subsequent to conformational changes elicited by effector (RuBP) binding after the switch to autotrophic growth for *R. sphaeroides*, *R. capsulatus*, *Xanthobacter flavus*, *and H. thermoluteolus* (8, 76, 82, 83).

In *R. sphaeroides*, RuBP, which of course is a unique metabolite of the CBB pathway, was suggested to be the effector for CbbR. The suggestion was based on a study employing a strain deleted for form I and form II RubisCO, leading to an accumulation of RuBP and a subsequent increase in transcription for both the cbb_I and cbb_{II} operons (84). Additionally, *in vitro*, the CbbR from *R. sphaeroides* was shown to alter the angle at which it bends the cbb_I promoter DNA in the presence of RuBP, as illustrated by a change in the mobility of the CbbR/DNA complex in gel mobility shift assays (76). DNase I footprint analysis also demonstrates that RuBP improves protection of the cbb_I promoter by CbbR (76).

For *R. capsulatus*, the metabolites that may influence CbbR-mediated expression present a more complex situation. Since *R. capsulatus* has two CbbRs ($CbbR_I$ and $CbbR_{II}$) regulating two *cbb* operons (cbb_I and cbb_{II}, respectively) (7), each CbbR has its own set of effector molecules. In gel mobility shift assays, expression of $CbbR_I$ was shown to result in a significant increase in binding to the cbb_I promoter DNA in the presence of 3-phosphoglycerate, 2-phosphoglycolate, ATP, KH_2PO_4, and RuBP and a small increase in binding to the cbb_I promoter in the presence of NADPH, fructose-6-phosphate, fructose-1,6-bisphosphate (FBP), and ribose-5-phosphate (85). DNase I footprint analyses illustrated modified protection of the cbb_I promoter DNA by $CbbR_I$ in the presence of RuBP, indicating a change in conformation of $CbbR_I$ and suggesting an altered bend angle of the cbb_I promoter DNA (85). For $CbbR_{II}$, gel mobility shifts demonstrated enhanced binding to the cbb_{II} promoter in the presence of RuBP, 2-phosphoglycolate, 3-phosphoglycerate, phosphoenolpyruvate, and FBP (85).

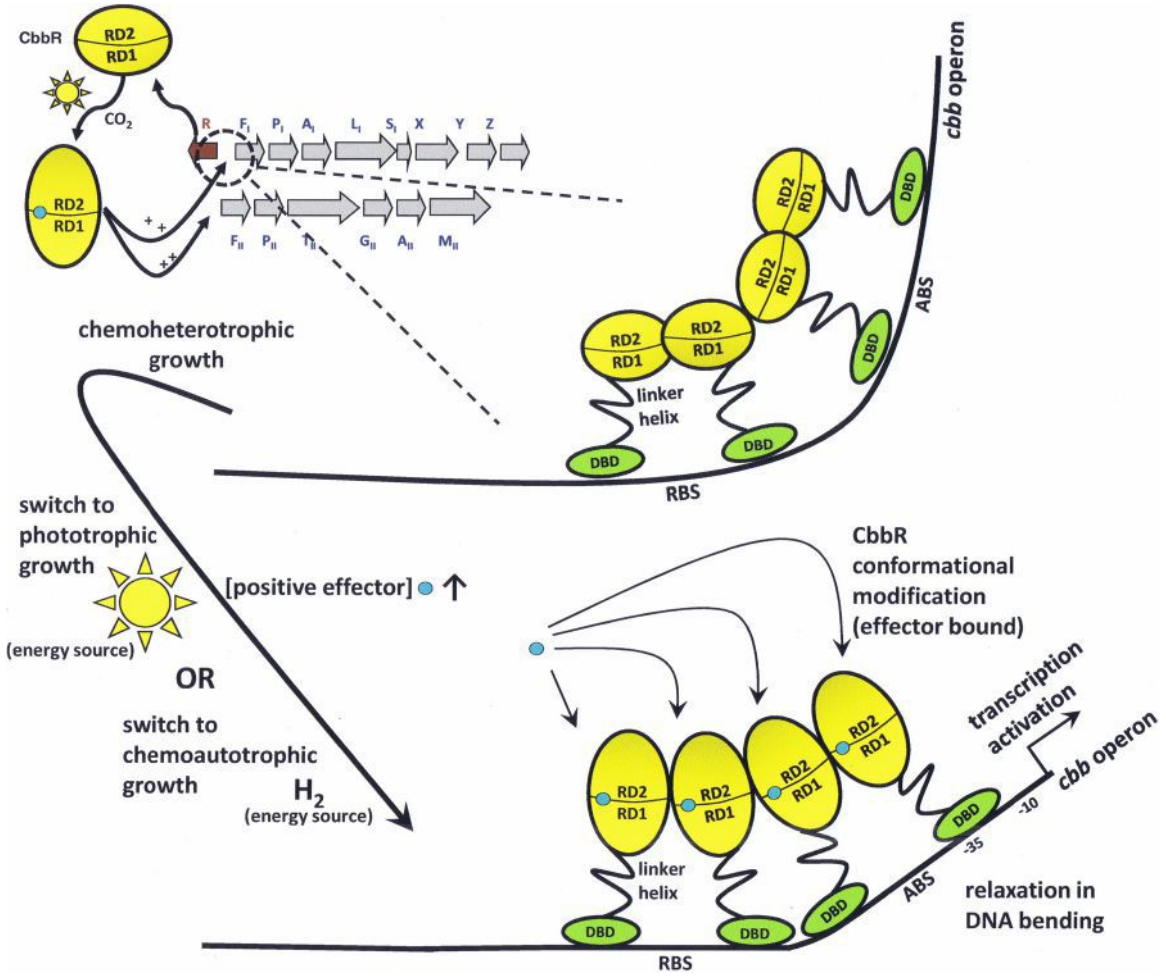


FIG 4 Consequences of relaxation of the promoter DNA bend angle upon effector binding (blue circle) to CbbR. For four CbbR-containing bacteria, a change in the promoter DNA bend angle has been shown to occur upon effector binding to the effector pocket of the protein. In *R. sphaeroides*, as illustrated here, CbbR undergoes a conformational change upon binding of a positive effector such as RuBP as a result of switching from heterotrophic to photo- or chemoautotrophic CO_2-dependent growth conditions, thus allowing a relaxation of DNA bending and subsequent transcription of the *cbb* genes. Under appropriate physiological conditions, RuBP concentrations greatly increase and RuBP binds to CbbR of *R. sphaeroides* (76) and $CbbR_I$ of *R. capsulatus* (84), while phosphoenolpyruvate (PEP), FBP, and 3-PGA concentrations greatly increase and each binds to $CbbR_{II}$ of *R. capsulatus* during photoautotrophic growth (85). NADPH concentrations rapidly increase in *X. flavus* (83, 87) and *H. thermoluteolus* (8) during chemoautotrophic growth. This relaxation of the DNA bend angle leads to appropriate contact with RNA polymerase and activates transcription of the *cbb* operon. RBS, recognition binding site; ABS, activation binding site. RD-1 and RD-2, recognition domains 1 and 2, respectively.

DNase I footprint analyses illustrated modified protection of the cbb_{II} promoter DNA by $CbbR_{II}$ in the presence of FBP, phosphoenolpyruvate, and 3-phosphoglycerate, indicating a change in conformation of $CbbR_{II}$ and suggesting an altered bend angle of the cbb_{II} promoter DNA (85).

Several compounds act as effectors of CbbR from *R. palustris*. Gel mobility shift assays demonstrated that ATP, FBP, RuBP, and NADPH all enhance binding of CbbR to the *cbbLS* promoter and that phosphoenolpyruvate inhibits binding of CbbR to the promoter (69). Quantitative SPR studies provided rate constant information and verified that *R. palustris* CbbR exhibits greater affinity for the *cbbLS* promoter in the presence of RuBP, FBP, ATP, and NADPH (61, 69).

Based on *in vitro* transcription experiments, the presence of phosphoenolpyruvate was shown to severely inhibit transcription of the *cbb* promoter by CbbR in *R. eutropha* strain H16 (82). This makes the effector metabolite, phosphoenolpyruvate, a corepressor for CbbR in *R. eutropha*, binding CbbR in the effector pocket but with the opposite effect on transcription. Gel mobility shift studies indicated that phosphoenolpyruvate enhances binding of CbbR to the *cbb* promoter (82, 86). Recent results also indicate that RuBP, ATP, and NADPH increase binding of wild-type CbbR to the *cbb* promoter of *R. eutropha* (86). Several mutant CbbRs with single-amino-acid substitutions near the effector pocket have reduced binding affinities in the presence of phophoenolpyruvate, RuBP, and ATP (86). Two other organisms, *Xanthobacter flavus* and *Hydrogenophilus thermoluteolus*, have CbbRs that show altered promoter DNA bending or increased promoter affinity in the presence of NADPH (8, 83, 87).

The CbbR from *Cyanidioschyzon merolae* (referred to as plastid-encoded transcription factor Ycf30) displays increased binding affinity for its promoter in the presence of NADPH and RuBP, as reported using gel mobility assays (15). *In vivo* experiments in permeabilized chloroplasts also indicated that RubisCO gene

transcription is activated by 3-phosphoglyceric acid, RuBP, and NADPH (15). Ycf30 controls expression of the nucleus-independent RubisCO operon in chloroplasts in this red alga (15).

Studies have demonstrated that CmpR and CcmR can use the same effector molecules that are utilized by many cyanobacterial CbbRs. CmpR from *S. elongatus* PCC 7942 has high affinity for 2-phosphoglycolate and low affinity for RuBP, as illustrated in gel mobility shift studies that demonstrated enhanced binding of CmpR to the *cmp* operon regulatory region (45). CcmR from *Synechocystis* PCC 6803 regulates the promoter regions of a variety of genes involved in the CCM through the use of $NADP^+$ and α-ketoglutarate as effectors (49). SPR studies illustrated that both $NADP^+$ and α-ketoglutarate enhanced binding of CcmR to the *ndhF3* promoter, the regulatory region for several genes involved in high-affinity CO_2 uptake (49). Similarly to the case with *R. eutropha*, where CbbR utilizes phosphoenolpyruvate as a corepressor (82), CcmR binds $NADP^+$ and α-ketoglutarate as corepressors to repress expression of the genes involved in the inducible high-affinity CCM of *Synechocystis* sp. strain PCC 6803 (49).

Finally, in the methylotrophic bacterium *Methylobacterium extorquens* AM1, QscR regulates two serine-cycle pathway operons, *qsc1* and *qsc2*, and also regulates the expression of a third gene, *glyA* (40, 41). Intermediate metabolites of the serine-cycle and traditional-energy metabolites (effectors of CbbRs) were found not to be effectors of QscR (40). Formyl-tetrahydrofolate, an intermediate for formaldehyde assimilation which is linked to the serine cycle, was shown to be a candidate effector for QscR (41). Gel mobility shift assays demonstrated that formyl-tetrahydrofolate enhances binding of QscR to the promoters of both the *qsc1* and *qsc2* serine-cycle operons (41).

To better understand how effector molecules interact with CbbR, it is instructive to consider an enlargement of the effector pocket structure of the CbbR subfamily as a distinguishing feature to separate CbbRs from other LTTRs (Fig. 1). The four conserved regions that define the effector pocket contain positively charged residues, usually arginine (sometimes lysine), and polar residues that attract and accommodate negatively charged effectors. The conserved amino acid sequences (highlighted in Fig. 1 in magenta) for each region are as follows: for region 1, GVVSTAKY FXP; for region 2, NR; for region 3, DLAIMGRPP; and for region 4, REXGSGTR ("X" represents a residue position that is not conserved) (Fig. 1). All analyzed bacterial CbbRs utilize similar effector metabolites that have negatively charged phosphate moieties, usually two phosphate moieties, or that may be organic acids that contain two negatively charged acid groups (Table 1). Many of the CbbR effectors, such as RuBP, 3-PGA, FBP, 2-phosphoglycolate, and 2-phosphoglycerate, are metabolites of the CBB pathway and would be expected to be present at higher concentrations in the cell during active biosynthetic CO_2 assimilation.

TABLE 1 CbbRs and subfamily members from various organisms and their effectors

Source and protein	Effector metabolite(s) (reference[s])
R. sphaeroides CbbR	RuBP[a,b] (76)
R. capsulatus	
$CbbR_I$	RuBP,[c,b] PEP,[b] 3-PGA,[b] 2-phosphoglycolate,[b] 2-phosphoglycerate,[b] ATP,[b] KH_2PO_4[b] (85)
$CbbR_{II}$	FBP,[c] 3-PGA,[c,b] PEP,[c,b] RuBP,[b] 2-phosphoglycolate[b] (85)
R. eutropha CbbR	PEP[b,d] (corepressor), RuBP,[b] ATP,[b] NADPH[b] (82, 86)
R. palustris CbbR	RuBP,[b,e] ATP,[b,e] FBP,[b,e] NADPH,[b] PEP (inhibits DNA binding) (61, 69)
X. flavus CbbR	NADPH[a,b] (83, 87)
H. thermoluteolus CbbR	NADPH[a] (8)
C. merolae CbbR	RuBP,[b,d] NADPH,[b,d] 3-PGA[d] (15)
S. elongatus CmpR	RuBP,[b] 2-phosphoglycolate[b] (45)
Synechocystis PCC 6803 CcmR	$NADP^+$,[e] α-ketoglutarate[e] (49)
M. extorquens QscR	Formyl-tetrahydrofolate[b] (41)

[a] Metabolite that allows a change in the bend angle CbbR imposes on promoter DNA via gel mobility shift assay.
[b] Metabolite that increases CbbR binding affinity for promoter DNA (i.e., enhances stability of CbbR on promoter DNA) via gel mobility shift assay.
[c] Metabolite that alters the region of promoter DNA protected by CbbR via DNase I footprinting/protection assay.
[d] Metabolite that allows CbbR to activate (or inhibit) transcription from the *cbb* promoter via *in vitro* transcription assay.
[e] Metabolite that increases CbbR binding affinity for promoter DNA (i.e., enhances stability of CbbR on promoter DNA) via SPR.

CONSTITUTIVELY ACTIVE CbbR PROTEINS

LTTR constitutive activity may be defined as activation of gene expression under conditions that normally repress gene transcription, typically in the absence of the LTTR's effector. When certain residues were altered, various LTTR proteins were found to constitutively activate gene expression; each LTTR appears to be unique with respect to which amino acid substitutions confer constitutive activity. This is probably a logical adaptation, as one might assume that the residues that are important for effector binding or for specific interactions with target DNA might be specific for each LTTR. ("LTTR*" denotes an LTTR variant with constitutive activity.) Many of these amino acid substitutions are centered at the effector pocket, but substitutions in other areas of the LTTR, such as at residues within the linker helix or hinge region or throughout RD-I and RD-II, can generate constitutive activity (70, 72, 76, 86, 88–94). Typically, single-amino-acid substitutions encompass the vast majority of the reported changes identified for constitutive proteins, and most constitutive proteins bind their effectors but may behave differently from the wild-type LTTR in gel mobility shift assays, DNase I footprinting assays, or *in vitro* transcription assays (76, 88, 89, 92–94). Amino acid substitutions that confer constitutive activity are thought to change the conformation of the LTTR tetramer to mimic the conformation seen when it is bound with the effector or to change the conformation of the LTTR/promoter complex to produce a favorable interaction with RNA polymerase to activate transcription. Studies of LTTR* proteins from several LTTR family members, including NodD, AmpR, OccR, CysB, OxyR, NahR, GtlR, XapR, and Cbl, have been previously published (70–72, 88–96).

A large set of CbbR* variants from both *R. sphaeroides* and *R. eutropha* have been isolated (76, 86). Constitutive proteins were generated by specific biological selection strategies involving the use of a reporter construct containing the *cbb* promoter fused to the *lacZ* open reading frame (ORF) integrated into the genome of a *cbbR* deletion strain for both organisms. The mutated *cbbR* proteins contain mutations that encode CbbR* proteins. For *R. sphaeroides*, several of the amino acid substitutions that confer constitutive activity clustered around the effector pocket which proved to be critical in defining the effector pocket for LTTRs (76). Sev-

eral of the CbbR* proteins interact differently with promoter DNA in the presence of RuBP (effector) compared to wild-type CbbR. Interestingly, several of the CbbR* proteins activate expression of the cbb_I promoter to a much greater extent than wild-type CbbR under conditions of *cbb* activation (76). Under conditions repressive for *cbb* activation (chemoheterotrophic growth), the amounts of *cbb* expression produced by the CbbR* proteins differed greatly; that is to say, some CbbR* proteins were better than others at activating gene expression.

CbbR* proteins from *R. eutropha* with amino acid substitutions located in all regions of the protein except the DBD were isolated. Substitutions were localized in the effector pocket, throughout RD-I and RD-II, and in the C terminus, and several residue changes were located in the linker helix. All of these were previously characterized (86). One particular CbbR* of interest is a truncation, leaving only the DBD and the linker helix to act as a transcriptional regulator. Nonetheless, this truncated CbbR* was able to support growth under chemoautotrophic conditions and to activate the *cbb* operons under repressive conditions in *R. eutropha* (86). This truncated protein illustrates that the DBD/linker helix region of CbbR is sufficient to activate expression from the *cbb* promoter, demonstrating that either the DBD or the linker helix or both make contact with the RNA polymerase (86). Similarly to some CbbR* proteins from *R. sphaeroides*, some of the CbbR* proteins from *R. eutropha* activated expression at levels severalfold greater than those seen with wild-type CbbR under autotrophic growth conditions (76, 86). The CbbRs of *R. sphaeroides* and *R. eutropha* exhibit only 35.6% identity, and each species appears to have a specific suite of residue changes that lead to CbbR* activity. Indeed, conserved residues whose presence is known to result in constitutive activity in *R. sphaeroides* CbbR did not confer constitutive activity when similar residues were changed in *R. eutropha* CbbR (86).

CONCLUDING REMARKS

CbbR controls the assimilation of carbon in autotrophic bacteria. It is the master regulator of the CBB CO_2 assimilation pathway, playing an essential role to ensure that the *cbb* genes are actively transcribed. It is clear that CbbR must be posttranslationally modified, and there are various ways in which this is accomplished, including the binding of small-molecule effectors as well as interactions with other transcription factors. Studies that investigate the interaction of other proteins with CbbR will advance understanding of how CO_2 fixation is regulated and of how LTTRs regulate transcription in general. CbbR also plays an important role in ensuring that CO_2-assimilatory organisms generate essential carbon metabolic intermediates that can be subsequently diverted into the synthesis of economically and globally meaningful biological molecules, such as biofuels. Constitutive CbbR variants have been proven to greatly increase the level of expression from the *cbb* promoter; thus, additional modification of the CbbR protein will further enhance the power of the *cbb* promoter as a tool for the production of biological compounds (76, 86). Other LTTRs are also amenable to constitutive modification, which may be important for the enhanced expression of other pathways in bacteria, since LTTRs are the most common transcriptional regulators in prokaryotes.

ACKNOWLEDGMENTS

This work was supported by grant DE-FG02-08ER15976 from the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, and ARPA-E grant DE-AR000009 from the U.S. Department of Energy.

REFERENCES

1. **Viale AM, Kobayashi H, Akazawa T, Henikoff S.** 1991. rbcR, a gene coding for a member of the LysR family of transcriptional regulators, is located upstream of the expressed set of ribulose 1,5-bisphosphate carboxylase/oxygenase genes in the photosynthetic bacterium *Chromatium vinosum*. J Bacteriol **173:**5224–5229.
2. **Windhövel U, Bowien B.** 1991. Identification of *cfxR*, an activator gene of autotrophic CO_2 fixation in *Alcaligenes eutrophus*. Mol Microbiol **5:**2695–2705. http://dx.doi.org/10.1111/j.1365-2958.1991.tb01978.x.
3. **Kusano T, Sugawara K.** 1993. Specific binding of *Thiobacillus ferrooxidans* RbcR to the intergenic sequence between the *rbc* operon and the *rbcR* gene. J Bacteriol **175:**1019–1025.
4. **Falcone DL, Tabita FR.** 1993. Complementation analysis and regulation of CO_2 fixation gene expression in a ribulose 1,5-bisphosphate carboxylase-oxygenase deletion strain of *Rhodospirillum rubrum*. J Bacteriol **175:**5066–5077.
5. **Gibson JL, Tabita FR.** 1993. Nucleotide sequence and functional analysis of CbbR, a positive regulator of the Calvin cycle operons of *Rhodobacter sphaeroides*. J Bacteriol **175:**5778–5784.
6. **van den Bergh ERE, Dijkhuizen L, Meijer WG.** 1993. CbbR, a LysR-type transcriptional activator, is required for expression of the autotrophic CO_2 fixation enzymes of *Xanthobacter flavus*. J Bacteriol **175:**6097–6104.
7. **Paoli GC, Vichivanives P, Tabita FR.** 1998. Physiological control and regulation of the *Rhodobacter capsulatus cbb* operons. J Bacteriol **180:**4258–4269.
8. **Terazono K, Hayashi NR, Igarashi Y.** 2001. CbbR, a LysR-type transcriptional regulator from *Hydrogenophilus thermoluteolus*, binds two *cbb* promoter regions. FEMS Microbiol Lett **198:**151–157. http://dx.doi.org/10.1111/j.1574-6968.2001.tb10635.x.
9. **Wei X, Sayavedra-Soto LA, Arp DJ.** 2004. The transcription of the *cbb* operon in *Nitrosomonas europaea*. Microbiology **150:**1869–1879. http://dx.doi.org/10.1099/mic.0.26785-0.
10. **Toyoda K, Yoshizawa Y, Arai H, Ishii M, Igarashi Y.** 2005. The role of two CbbRs in the transcriptional regulation of three ribulose-1,5-bisphosphate carboxylase/oxygenase genes in *Hydrogenovibrio marinus* strain MH-110. Microbiology **151:**3615–3625. http://dx.doi.org/10.1099/mic.0.28056-0.
11. **Romagnoli S, Tabita FR.** 2006. A novel three-protein two-component system provides a regulatory twist on an established circuit to modulate expression of the cbb_I region of *Rhodopseudomonas palustris* CGA010. J Bacteriol **188:**2780–2791. http://dx.doi.org/10.1128/JB.188.8.2780-2791.2006.
12. **Lee JH, Park DO, Park SW, Hwang EH, Oh JI, Kim YM.** 2009. Expression and regulation of ribulose 1,5-bisphosphate carboxylase/oxygenase genes in *Mycobacterium* sp. strain JC1 DSM 3803. J Microbiol **47:**297–307. http://dx.doi.org/10.1007/s12275-008-0210-3.
13. **Esparza M, Cárdenas JP, Bowien B, Jedlicki E, Holmes DS.** 2010. Genes and pathways for CO_2 fixation in the obligate, chemolithoautotrophic acidophile, Acidithiobacillus ferrooxidans, carbon fixation in A. ferrooxidans. BMC Microbiol **10:**229. http://dx.doi.org/10.1186/1471-2180-10-229.
14. **Wang D, Zhang Y, Pohlmann EL, Li J, Roberts GP.** 2011. The poor growth of *Rhodospirillum rubrum* mutants lacking RubisCO is due to the accumulation of ribulose-1,5-bisphophate. J Bacteriol **193:**3293–3303. http://dx.doi.org/10.1128/JB.00265-11.
15. **Minoda A, Weber APM, Tanaka K, Miyagishima SY.** 2010. Nucleus-independent control of the rubisco operon by the plastid-encoded transcription factor Ycf30 in the red alga *Cyanidioschyzon merolae*. Plant Physiol **154:**1532–1540. http://dx.doi.org/10.1104/pp.110.163188.
16. **Ellis RJ.** 1979. The most abundant protein in the world. Trends Biochem Sci **4:**241–244. http://dx.doi.org/10.1016/0968-0004(79)90212-3.
17. **Schell MA.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu Rev Microbiol **47:**597–626. http://dx.doi.org/10.1146/annurev.mi.47.100193.003121.
18. **Harrison SC, Aggarwal AK.** 1990. DNA recognition by proteins with the

helix-turn-helix motif. Annu Rev Biochem **59:**933–969. http://dx.doi.org/10.1146/annurev.bi.59.070190.004441.
19. **Maddocks SE, Oyston PCF.** 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology **154:**3609–3623. http://dx.doi.org/10.1099/mic.0.2008/022772-0.
20. **Muraoka S, Okumura R, Ogawa N, Nonaka T, Miyashita K, Senda T.** 2003. Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. J Mol Biol **328:**555–566. http://dx.doi.org/10.1016/S0022-2836(03)00312-7.
21. **Monferrer D, Tralau T, Kertesz MA, Dix I, Sola M, Uson I.** 2010. Structural studies on the full-length LysR-type regulator TsaR from *Comamonas testosteroni* T-2 reveal a novel open conformation of the tetrameric LTTR fold. Mol Microbiol **75:**1199–1214. http://dx.doi.org/10.1111/j.1365-2958.2010.07043.x.
22. **Zhou X, Lou Z, Fu S, Yang A, Shen H, Li Z, Feng Y, Bartlam M, Wang H, Rao Z.** 2010. Crystal structure of ArgP from *Mycobacterium tuberculosis* confirms two distinct conformations of full-length LysR transcriptional regulators and reveals its function in DNA binding and transcriptional regulation. J Mol Biol **396:**1012–1024. http://dx.doi.org/10.1016/j.jmb.2009.12.033.
23. **Taylor JL, De Silva RS, Kovacikova G, Lin W, Taylor RK, Skorupski K, Kull FJ.** 2012. The crystal structure of AphB, a virulence gene activator from *Vibrio cholerae*, reveals residues that influence its response to oxygen and pH. Mol Microbiol **83:**457–470. http://dx.doi.org/10.1111/j.1365-2958.2011.07919.x.
24. **Ruangprasert A, Craven SH, Neidle EL, Momany C.** 2010. Full-length structures of BenM and two variants reveal different oligomerization schemes for LysR-type transcriptional regulators. J Mol Biol **404:**568–586. http://dx.doi.org/10.1016/j.jmb.2010.09.053.
25. **Vinson C, Acharya A, Taparowsky EJ.** 2006. Deciphering B-zip transcription factor interactions *in vitro* and *in vivo*. Biochim Biophys Acta **1759:**4–12. http://dx.doi.org/10.1016/j.bbaexp.2005.12.005.
26. **Batchelor JD, Lee PS, Wang AC, Doucleff M, Wemmer DE.** 2013. Structural mechanism of GAF-regulated σ^{54} activators from *Aquifex aeolicus*. J Mol Biol **425:**156–170. http://dx.doi.org/10.1016/j.jmb.2012.10.017.
27. **Nargang F, McIntosh L, Somerville C.** 1984. Nucleotide sequence of the ribulose bisphosphate carboxylase gene from *Rhodospirillum rubrum*. Mol Gen Genet **193:**220–224. http://dx.doi.org/10.1007/BF00330671.
28. **Paoli GC, Morgan NS, Tabita FR, Shively JM.** 1995. Expression of the *cbbLcbbS* and *cbbM* genes and distinct organization of the *cbb* Calvin cycle structural genes of *Rhodobacter capsulatus*. Arch Microbiol **164:**396–405.
29. **Chung SY, Yaguchi T, Nishihara H, Igarashi Y, Kodama T.** 1993. Purification of form L_2 rubisCO from a marine obligately autotrophic hydrogen-oxidizing bacterium, Hydrogenovibrio marinus strain MH-110. FEMS Microbiol Lett **109:**49–54.
30. **Yaguchi T, Chung SY, Igarashi Y, Kodama T.** 1994. Cloning and sequence of the L_2 form of RubisCO from a marine obligately autotrophic hydrogen-oxidizing bacterium, Hydrogenovibrio marinus strain MH-110. Biosci Biotech Biochem **58:**1733–1737. http://dx.doi.org/10.1271/bbb.58.1733.
31. **Yoshizawa Y, Toyoda K, Arai H, Ishii M, Igarashi Y.** 2004. CO_2-responsive expression and gene organization of three ribulose-1,5-bisphosphate carboxylase/oxygenase enzymes and carboxysomes in *Hydrogenovibrio marinus* strain MH-110. J Bacteriol **186:**5685–5691. http://dx.doi.org/10.1128/JB.186.17.5685-5691.2004.
32. **Watson GMF, Tabita FR.** 1997. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation. FEMS Microbiol Lett **146:**13–22. http://dx.doi.org/10.1111/j.1574-6968.1997.tb10165.x.
33. **Kusian B, Bowien B.** 1997. Organization and regulation of *cbb* CO_2 assimilation genes in autotrophic bacteria. FEMS Microbiol Rev **21:**135–155. http://dx.doi.org/10.1111/j.1574-6976.1997.tb00348.x.
34. **Tabita FR, Hanson TE, Li H, Satagopan S, Singh J, Chan S.** 2007. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. Microbiol Mol Biol Rev **71:**576–599. http://dx.doi.org/10.1128/MMBR.00015-07.
35. **Tabita FR, Satagopan S, Hanson TE, Kreel NE, Scott SS.** 2008. Distinct form I, II, III, and IV RubisCO proteins from the three kingdoms of life provide clues about RubisCO evolution and structure/function relationships. J Exp Bot **59:**1515–1514.
36. **Dubbs JM, Tabita FR.** 1998. Two functionally distinct regions upstream of the cbb_I operon of *Rhodobacter sphaeroides* regulate gene expression. J Bacteriol **180:**4903–4911.
37. **Dubbs JM, Tabita FR.** 2003. Interactions of the cbb_{II} promoter-operator region with CbbR and RegA (PrrA) regulators indicate distinct mechanisms to control expression of the two *cbb* operons of *Rhodobacter sphaeroides*. J Biol Chem **278:**16443–16450. http://dx.doi.org/10.1074/jbc.M211267200.
38. **Kusian B, Bednarski R, Husemann M, Bowien B.** 1995. Characterization of the duplicate ribulose-1,5-bisphosphate carboxylase genes and *cbb* promoters of *Alcaligenes eutrophus*. J Bacteriol **177:**4442–4450.
39. **Park SW, Hwang EH, Jang HS, Lee JH, Kang BS, Oh JI, Kim YM.** 2009. Presence of duplicate genes encoding a phylogenetically new subgroup of form I ribulose 1,5-bisphosphate carboxylase/oxygenase in *Mycobacterium* sp. strain JC1 DSM 3803. Res Microbiol **160:**159–165. http://dx.doi.org/10.1016/j.resmic.2008.12.002.
40. **Kalyuzhnaya MG, Lidstrom ME.** 2003. QscR, a LysR-type transcriptional regulator and CbbR homolog, is involved in regulation of the serine cycle genes in *Methylobacterium extorquens* AM1. J Bacteriol **185:**1229–1235. http://dx.doi.org/10.1128/JB.185.4.1229-1235.2003.
41. **Kalyuzhnaya MG, Lidstrom ME.** 2005. QscR-mediated transcriptional activation of serine cycle genes in *Methylobacterium extorquens* AM1. J Bacteriol **187:**7511–7517. http://dx.doi.org/10.1128/JB.187.21.7511-7517.2005.
42. **Omata T, Gohta S, Takahashi Y, Harano Y, Maeda S-I.** 2001. Involvement of a CbbR homolog in low CO_2-induced activation of the bicarbonate transporter operon in cyanobacteria. J Bacteriol **183:**1891–1898. http://dx.doi.org/10.1128/JB.183.6.1891-1898.2001.
43. **Woodger FJ, Badger MR, Price GD.** 2003. Inorganic carbon limitation induces transcripts encoding components of the CO_2-concentrating mechanism in *Synechococcus* sp. PCC7942 through a redox-independent pathway. Plant Physiol **133:**2069–2080.
44. **Takahashi Y, Yamaguchi O, Omata T.** 2004. Roles of CmpR, a LysR family transcriptional regulator, in acclimation of the cyanobacterium *Synechococcus* sp. strain PCC 7942 to low-CO_2 and high-light conditions. Mol Microbiol **52:**837–845. http://dx.doi.org/10.1111/j.1365-2958.2004.04021.x.
45. **Nishimura T, Takahashi Y, Yamaguchi O, Suzuki H, Maeda S-I, Omata T.** 2008. Mechanism of low CO_2-induced activation of the *cmp* bicarbonate transporter operon by a LysR family protein in the cyanobacterium *Synechococcus elongatus* strain PCC7942. Mol Microbiol **68:**98–109. http://dx.doi.org/10.1111/j.1365-2958.2008.06137.x.
46. **López-Igual R, Picossi S, Lopéz-Garrido J, Flores E, Herrero A.** 2012. N and C control of ABC-type bicarbonate transporter Cmp and its LysR-type transcriptional regulator CmpR in a heterocyst-forming cyanobacterium, Anabaena sp. Environ Microbiol **14:**1035–1048. http://dx.doi.org/10.1111/j.1462-2920.2011.02683.x.
47. **Rae BD, Long BM, Badger MR, Price GD.** 2013. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. Microbiol Mol Biol Rev **77:**357–379. http://dx.doi.org/10.1128/MMBR.00061-12.
48. **Woodger FJ, Bryant DA, Price GD.** 2007. Transcriptional regulation of the CO_2-concentrating mechanism in a euryhaline, coastal marine cyanobacterium, Synechococcus sp. strain PCC 7002: role of NdhR/CcmR. J Bacteriol **189:**3335–3347. http://dx.doi.org/10.1128/JB.01745-06.
49. **Daley SME, Kappell AD, Carrick MJ, Burnap RL.** 2012. Regulation of the cyanobacterial CO_2-concentrating mechanism involves internal sensing of $NADP^+$ and α-ketoglutarate levels by transcription factor CcmR. PLoS One **7:**e41286. http://dx.doi.org/10.1371/journal.pone.0041286.
50. **Tao K, Fujita N, Ishihama A.** 1993. Involvement of the RNA polymerase alpha subunit C-terminal region in co-operative interaction and transcriptional activation with OxyR protein. Mol Microbiol **7:**859–864. http://dx.doi.org/10.1111/j.1365-2958.1993.tb01176.x.
51. **McFall SM, Chugani SA, Chakrabarty AM.** 1998. Transcriptional activation of the catechol and chlorocatechol operons: variations on a theme. Gene **223:**257–267. http://dx.doi.org/10.1016/S0378-1119(98)00366-7.
52. **Park W, Jeon CO, Madsen EL.** 2002. Interaction of NahR, a LysR-type transcriptional regulator, with the alpha subunit of RNA polymerase in the naphthalene degrading bacterium, Pseudomonas putida NCIB 9816-4. FEMS Microbiol Lett **213:**159–165. http://dx.doi.org/10.1111/j.1574-6968.2002.tb11300.x.
53. **Deghmane AE, Giorgini D, Maigre L, Taha MK.** 2004. Analysis *in vitro* and *in vivo* of the transcriptional regulator CrgA of *Neisseria meningitidis*

upon contact with target cells. Mol Microbiol **53:**917–927. http://dx.doi.org/10.1111/j.1365-2958.2004.04167.x.

54. **Lochowska A, Iwanicka-Nowicka R, Zaim J, Witkowska-Zimny M, Bolewska K, Hryniewicz MM.** 2004. Identification of activating region (AR) of *Escherichia coli* LysR-type transcription factor CysB and CysB contact site on RNA polymerase alpha subunit at the *cysP* promoter. Mol Microbiol **53:**791–806. http://dx.doi.org/10.1111/j.1365-2958.2004.04161.x.
55. **Stauffer LT, Stauffer GV.** 2005. GcvA interacts with both the α and σ subunits of RNA polymerase to activate the *Escherichia coli gcvB* gene and the *gcvTHP* operon. FEMS Microbiol Lett **242:**333–338. http://dx.doi.org/10.1016/j.femsle.2004.11.027.
56. **Ghrist A, Heil G, Stauffer GV.** 2001. GcvR interacts with GcvA to inhibit activation of the *Escherichia coli* glycine cleavage operon. Microbiology **147:**2215–2221. http://dx.doi.org/10.1099/00221287-147-8-2215.
57. **Kovacikova G, Lin W, Skorupski K.** 2004. *Vibrio cholerae* AphA uses a novel mechanism for virulence gene activation that involves interaction with the LysR-type regulator AphB at the *tcpPH* promoter. Mol Microbiol **53:**129–142. http://dx.doi.org/10.1111/j.1365-2958.2004.04121.x.
58. **Dangel AW, Tabita FR.** 2009. Protein-protein interactions between CbbR and RegA (PrrA), transcriptional regulators of the *cbb* operons of *Rhodobacter sphaeroides*. Mol Microbiol **71:**717–729. http://dx.doi.org/10.1111/j.1365-2958.2008.06558.x.
59. **Dangel AW, Luther A, Tabita FR.** 2014. Amino acid residues of RegA important for interactions with the CbbR-DNA complex of *Rhodobacter sphaeroides*. J Bacteriol **196:**3179–3190. http://dx.doi.org/10.1128/JB.01842-14.
60. **Joshi GS, Bobst CE, Tabita FR.** 2011. Unravelling the regulatory twist—regulation of CO_2 fixation in *Rhodopseudomonas palustris* CGA010 mediated by atypical response regulators. Mol Microbiol **80:**756–771. http://dx.doi.org/10.1111/j.1365-2958.2011.07606.x.
61. **Joshi GS, Zianni M, Bobst CE, Tabita FR.** 2013. Regulatory twist and synergistic role of metabolic coinducer- and response regulator-mediated CbbR-cbb_I interactions in *Rhodopseudomonas palustris* CGA010. J Bacteriol **195:**1381–1388. http://dx.doi.org/10.1128/JB.02060-12.
62. **Gibson JL, Dubbs JM, Tabita FR.** 2002. Differential expression of the CO_2 fixation operons of *Rhodobacter sphaeroides* by the Prr/RegA two-component system during chemoautotrophic growth. J Bacteriol **184:**6654–6664. http://dx.doi.org/10.1128/JB.184.23.6654-6664.2002.
63. **Joshi HM, Tabita FR.** 1996. A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. Proc Natl Acad Sci U S A **93:**14515–14520. http://dx.doi.org/10.1073/pnas.93.25.14515.
64. **Elsen S, Swem LR, Swem DL, Bauer CE.** 2004. RegB/RegA, a highly conserved redox-responding global two-component regulatory system. Microbiol Mol Biol Rev **68:**263–279. http://dx.doi.org/10.1128/MMBR.68.2.263-279.2004.
65. **Dubbs JM, Bird TH, Bauer CE, Tabita FR.** 2000. Interaction of CbbR and RegA* transcription regulators with the *Rhodobacter sphaeroides* cbb_I promoter-operator region. J Biol Chem **275:**19224–19230. http://dx.doi.org/10.1074/jbc.M002125200.
66. **Ranson-Olson B, Zeilstra-Ryalls JH.** 2008. Regulation of the *Rhodobacter sphaeroides* 2.4.1 *hemA* gene by PrrA and FnrL. J Bacteriol **190:**6769–6778. http://dx.doi.org/10.1128/JB.00828-08.
67. **Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, Land ML, Pelletier DA, Beatty JT, Lang AS, Tabita FR, Gibson JL, Hanson TE, Bobst C, Torres y Torres JL, Peres C, Harrison FH, Gibson J, Harwood CS.** 2004. Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. Nat Biotechnol **22:**55–61. http://dx.doi.org/10.1038/nbt923.
68. **Romagnoli S, Tabita FR.** 2007. Phosphotransfer reactions of the CbbRRS three-protein two-component system from *Rhodopseudomonas palustris* CGA010 appear to be controlled by an internal molecular switch on the sensor kinase. J Bacteriol **189:**325–335. http://dx.doi.org/10.1128/JB.01326-06.
69. **Joshi GS, Zianni M, Bobst CE, Tabita FR.** 2012. Further unravelling the regulatory twist by elucidating metabolic coinducer-mediated CbbR-cbb_I promoter interactions in *Rhodopseudomonas palustris* CGA010. J Bacteriol **194:**1350–1360. http://dx.doi.org/10.1128/JB.06418-11.
70. **Burn JE, Hamilton WD, Wooten JC, Johnston AW.** 1989. Single and multiple mutations affecting properties of the regulatory gene *nodD* of *Rhizobium*. Mol Microbiol **3:**1567–1577. http://dx.doi.org/10.1111/j.1365-2958.1989.tb00142.x.
71. **Cebolla A, Sousa C, de Lorenzo V.** 1997. Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading *Pseudomonas* define protein sites involved in binding of aromatic inducers. J Biol Chem **272:**3986–3992. http://dx.doi.org/10.1074/jbc.272.7.3986.
72. **Jørgensen C, Dandanell G.** 1999. Isolation and characterization of mutations in the *Escherichia coli* regulatory protein XapR. J Bacteriol **181:**4397–4403.
73. **Tyrrell R, Verschueren KH, Dodson EJ, Murshudov GN, Addy C, Wilkinson AJ.** 1997. The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure **5:**1017–1032. http://dx.doi.org/10.1016/S0969-2126(97)00254-2.
74. **Choi H, Kim S, Mukhopadhyay P, Cho S, Woo J, Storz G, Ryu S.** 2001. Structural basis of the redox switch in the OxyR transcription factor. Cell **105:**103–113. http://dx.doi.org/10.1016/S0092-8674(01)00300-2.
75. **Smirnova IA, Dian C, Leonard GA, McSweeney S, Birse D, Brzezinski P.** 2004. Development of a bacterial biosensor for nitrotoluenes: the crystal structure of the transcriptional regulator DntR. J Mol Biol **340:**405–418. http://dx.doi.org/10.1016/j.jmb.2004.04.071.
76. **Dangel AW, Gibson JL, Janssen AP, Tabita FR.** 2005. Residues that influence *in vivo* and *in vitro* CbbR function in *Rhodobacter sphaeroides* and identification of a specific region critical for co-inducer recognition. Mol Microbiol **57:**1397–1414. http://dx.doi.org/10.1111/j.1365-2958.2005.04783.x.
77. **Stec E, Witkowska-Zimny M, Hryniewicz MM, Neumann P, Wilkinson AJ, Brzozowski AM, Verma CS, Zaim J, Wysocki S, Bujacz GD.** 2006. Structural basis of the sulphate starvation response in *E. coli*: crystal structure and mutational analysis of the cofactor-binding domain of the Cbl transcriptional regulator. J Mol Biol **364:**309–322. http://dx.doi.org/10.1016/j.jmb.2006.06.033.
78. **Ezezika OC, Haddad S, Clark TJ, Neidle EL, Momany C.** 2007. Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. J Mol Biol **367:**616–629. http://dx.doi.org/10.1016/j.jmb.2006.09.090.
79. **Balcewich MD, Reeve TM, Orlikow EA, Donald LJ, Vocadlo DJ, Mark BL.** 2010. Crystal structure of the AmpR effector binding domain provides insight into the molecular regulation of inducible AmpC β-lactamase. J Mol Biol **400:**998–1010. http://dx.doi.org/10.1016/j.jmb.2010.05.040.
80. **Hayes RP, Moural TW, Lewis KM, Onofrei D, Xun L, Kang C.** 2014. Structures of the inducer-binding domain of pentachlorophenol-degrading gene regulator PcpR from *Sphingobium chlorophenolicum*. Int J Mol Sci **15:**20736–20752. http://dx.doi.org/10.3390/ijms151120736.
81. **Vadlamani G, Thomas MD, Patel TR, Donald LJ, Reeve TM, Stetefeld J, Standing KG, Vocadlo DJ, Mark BL.** 2015. The β-lactamase gene regulator AmpR is a tetramer that recognizes and binds the D-Ala-D-Ala motif of its repressor UDP-*N*-acetylmuramic acid (MurNAc)-pentapeptide. J Biol Chem **290:**2630–2643. http://dx.doi.org/10.1074/jbc.M114.618199.
82. **Grzeszik C, Jeffke T, Schaferjohann J, Kusian B, Bowien B.** 2000. Phosphoenolpyruvate is a signal metabolite in transcriptional control of the *cbb* CO_2 fixation operons in *Ralstonia eutropha*. J Mol Microbiol Biotechnol **2:**311–320.
83. **van Keulen G, Girbal L, van den Bergh ERE, Dijkhuizen L, Meijer WG.** 1998. The LysR-type transcriptional regulator CbbR controlling autotrophic CO_2 fixation by *Xanthobacter flavus* is an NADPH sensor. J Bacteriol **180:**1411–1417.
84. **Smith SA, Tabita FR.** 2002. Up-regulated expression of the cbb_I and cbb_{II} operons during photoheterotrophic growth of a ribulose 1,5-bisphosphate carboxylase-oxygenase deletion mutant of *Rhodobacter sphaeroides*. J Bacteriol **184:**6721–6724. http://dx.doi.org/10.1128/JB.184.23.6721-6724.2002.
85. **Dubbs P, Dubbs JM, Tabita FR.** 2004. Effector-mediated interaction of $CbbR_I$ and $CbbR_{II}$ regulators with target sequences in *Rhodobacter capsulatus*. J Bacteriol **186:**8026–8035. http://dx.doi.org/10.1128/JB.186.23.8026-8035.2004.
86. **Dangel AW, Tabita FR.** 9 July 2015, posting date. Amino acid substitutions in the transcriptional regulator CbbR lead to constitutively active CbbR proteins that elevate expression of the *cbb* CO_2 fixation operons in *Ralstonia eutropha* (*Cupriavidus necator*) and identify regions of CbbR necessary for gene activation. Microbiology http://dx.doi.org/10.1099/mic.0.000131.
87. **van Keulen G, Ridder ANJA, Dijkhuizen L, Meijer WG.** 2003. Analysis of DNA binding and transcriptional activation by the LysR-type transcrip-

tional regulator CbbR of *Xanthobacter flavus*. J Bacteriol **185:**1245–1252. http://dx.doi.org/10.1128/JB.185.4.1245-1252.2003.

88. **Lochowska A, Iwanicka-Nowicka R, Plochocka D, Hryniewicz MM.** 2001. Functional dissection of the LysR-type CysB transcriptional regulator. Regions important for DNA binding, inducer response, oligomerization, and positive control. J Biol Chem **276:**2098–2107.
89. **Akakura R, Winans SC.** 2002. Constitutive mutations of the OccR regulatory protein affect DNA bending in response to metabolites released from plant tumors. J Biol Chem **277:**5866–5874. http://dx.doi.org/10.1074/jbc.M110555200.
90. **Bartowsky E, Normark S.** 1991. Purification and mutant analysis of *Citrobacter freundii* AmpR, the regulator for chromosomal AmpC β-lactamase. Mol Microbiol **5:**1715–1725. http://dx.doi.org/10.1111/j.1365-2958.1991.tb01920.x.
91. **Cho K, Winans SC.** 1993. Altered-function mutations in the *Agrobacterium tumefaciens* OccR protein and in an OccR-regulated promoter. J Bacteriol **175:**7715–7719.
92. **Kullik I, Toledano MB, Tartaglia LA, Storz G.** 1995. Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for oxidation and transcriptional activation. J Bacteriol **177:**1275–1284.
93. **Belitsky BR, Sonenshein AL.** 1997. Altered transcription activation specificity of a mutant form of *Bacillus subtilis* GltR, a LysR family member. J Bacteriol **179:**1035–1043.
94. **Bykowski T, van der Ploeg JR, Iwanicka-Nowicka R, Hryniewicz MM.** 2002. The switch from inorganic to organic sulphur assimilation in *Escherichia coli*: adenosine 5′-phosphosulphate (APS) as a signalling molecule for sulphate excess. Mol Microbiol **43:**1347–1358. http://dx.doi.org/10.1046/j.1365-2958.2002.02846.x.
95. **Hou B, Li F, Yang X, Hong G.** 2009. A small functional intramolecular region of NodD was identified by mutation. Acta Biochim Biophys Sin **41:**822–830. http://dx.doi.org/10.1093/abbs/gmp073.
96. **Colyer TE, Kredich NM.** 1994. Residue threonine-149 of the *Salmonella typhimurium* CysB transcription activator: mutations causing constitutive expression of positively regulated genes of the cysteine regulon. Mol Microbiol **13:**797–805. http://dx.doi.org/10.1111/j.1365-2958.1994.tb00472.x.

JB Journals.ASM.org

Fur-Mediated Global Regulatory Circuits in Pathogenic *Neisseria* Species

Chunxiao Yu[a] and Caroline Attardo Genco[a,b]

Department of Medicine, Section of Infectious Diseases,[a] and Department of Microbiology,[b] Boston University School of Medicine, Boston, Massachusetts, USA

The ferric uptake regulator (Fur) protein has been shown to function as a repressor of transcription in a number of diverse microorganisms. However, recent studies have established that Fur can function at a global level as both an activator and a repressor of transcription through both direct and indirect mechanisms. Fur-mediated indirect activation occurs via the repression of additional repressor proteins, or small regulatory RNAs, thereby activating transcription of a previously silent gene. Fur mediates direct activation through binding of Fur to the promoter regions of genes. Whereas the repressive mechanism of Fur has been thoroughly investigated, emerging studies on direct and indirect Fur-mediated activation mechanisms have revealed novel global regulatory circuits.

Iron homeostasis is a highly regulated process in bacteria, as iron both is an essential nutrient and when in excess, can lead to toxicity via the production of hydroxyl or peroxide radicals. To maintain iron homeostasis, many bacteria utilize iron-binding transcriptional regulators, which, upon binding to free iron, are triggered to regulate transcription of genes involved in maintaining intracellular iron levels. Iron-binding proteins are members of a large class of metal ion-binding transcriptional regulators, includin MntR, DtxR, and Zur (58). These regulators act as environmental sensors of essential metals, including iron, and modulate gene expression accordingly. The master regulator of iron homeostasis is an iron-binding transcription factor termed the Ferric Uptake Regulator (Fur).

The first indication that bacterial iron homeostasis relied on a single central regulator was published by Ernst et al. in 1978 (31). Those studies showed that a *Salmonella enterica* serovar Typhimurium mutant lacked iron-responsive regulation of many genes, including those involved in iron-enterochelin and ferrochrome uptake. That mutant was termed an iron (Fe) uptake regulation (*fur*) mutant (31). A similar mutant was soon isolated in *Escherichia coli*, and the *E. coli fur* gene was subsequently cloned and sequenced (45, 84). Orthologues of *fur* have since been identified in numerous Gram-negative and Gram-positive species, and it has been shown that Fur proteins share a high degree of sequence homology between species (Fig. 1). In most organisms, Fur is present as a 15-to-17-kDa protein that forms dimers in the presence of iron (II) or other divalent cations (6, 24, 74, 91). As revealed by analysis of several crystal structures of the Fur protein from various pathogens, the amino terminus of Fur has been shown to bind to DNA whereas the carboxyl terminus is involved in dimer formation (26, 81, 91). Analysis of the crystal structure of Fur has also identified multiple metal-binding sites. These metal-binding sites contain a conserved histidine-histidine-aspartic acid-histidine (HHDH) motif (a specific region shown to be involved in cofactor binding), and mutagenesis studies have shown these four amino acids are crucial for Fur function (59, 83). As a repressor, the iron-bound Fur dimer binds to the −10 and −35 promoter regions to exclude binding of RNA polymerase, which results in the inhibition of transcriptional initiation (32) (Fig. 2A). The DNA sequence recognized by repressive Fur (designated a Fur box) was initially defined as a conserved 19-bp sequence, GATAA TGATAATCATTATC, in *E. coli* (24). Fur boxes determined in other Gram-positive or Gram-negative bacteria are similar to this consensus sequence (7, 24, 27, 37, 44, 69, 78, 80, 96, 98, 104, 105) (Table 1). Subsequently, the Fur box was interpreted as 9-1-9 inverted repeats (GATAATGAT-A-ATCATTATC), or a hexameric repetition of nATwAT (24, 33). The Fur boxes in *Bacillus subtilis* and *Helicobacter pylori* are represented by shorter inverted repeats of 7-1-7 (TGATAATnATTATCA and TAATAATnATTATTA, respectively) (37, 80). Based on these studies and the crystal structure of Fur, it is predicted that two Fur dimers simultaneously bind one Fur box and the two Fur dimers may occupy three to four hexameric repeats (81). This consensus sequence has been used to successfully predict novel Fur-regulated genes *in silico* (27, 44, 53).

Fur FUNCTIONS AS A GLOBAL REGULATORY PROTEIN

Recently, global analyses of iron- and/or Fur-responsive transcriptomes of diverse bacterial pathogens, such as *H. pylori*, *Pseudomonas syringae*, *Vibrio cholerae*, *Yersinia pestis*, *Haemophilus influenzae*, *S. enterica* serovar Typhimurium, and *Listeria monocytogenes*, have revealed a number of novel regulatory roles for Fur (12, 20, 40, 56, 68, 94, 101, 102, 107). First, Fur has been demonstrated to repress transcription even in the absence of iron, a process termed apo-Fur regulation (13) (Fig. 2B). apo-Fur repression has been primarily characterized in *H. pylori* and has not yet been well described in other bacteria (13, 72). Although iron may not be important for apo-Fur function, it is possible that other metal ions play a role in apo-Fur-mediated transcriptional control (24, 74, 91). Second, in addition to its role as a repressor, Fur can also function as an activator in both the iron-bound form and apo form. In *V. vulnificus*, apo-Fur-mediated activation was shown to positively regulate the *fur* gene itself (57) (Fig. 2C). Transcriptome analysis has also identified additional genes that are activated by iron-bound Fur (12, 56, 68, 101, 102, 107).

Theoretically, transcriptional activation through Fur can be fulfilled by several pathways, including both direct (Fig. 2D) and indirect (Fig. 2E) mechanisms. A few examples of Fur-mediated

Published ahead of print 10 August 2012

Address correspondence to Caroline Attardo Genco, cgenco@bu.edu.

doi:10.1128/JB.00262-12

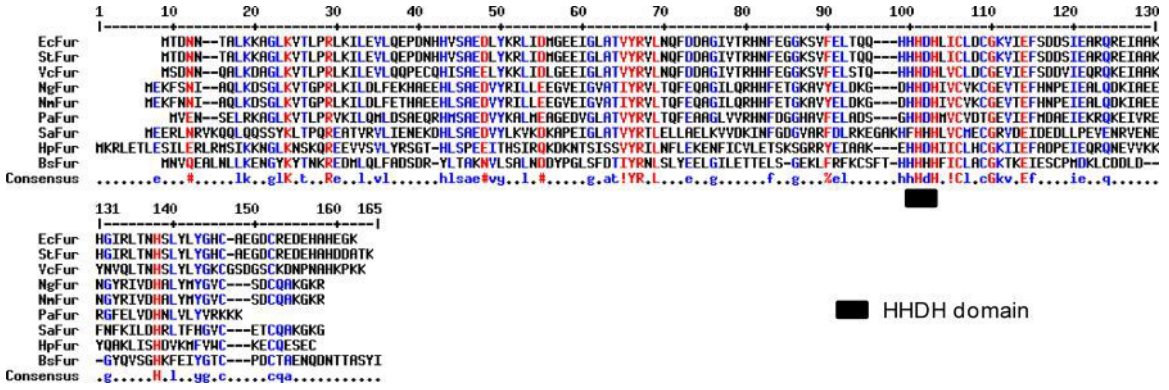


FIG 1 Alignment of Fur orthologues from both Gram-negative and Gram-positive bacteria.

indirect activation have been reported. Fur may activate transcription indirectly via repression of a small RNA (sRNA), such as RyhB of *E. coli* (5, 62, 99). As a result, the repressed targets of RyhB, consisting of *sdhCDAB*, *acnA*, *fumA*, and *sodB*, are activated when Fur is present (61–64). Fur-repressed sRNAs and their target genes have been identified in the Fur regulons of other bacteria, such as RyhB of *V. cholerae* (21), PrrF1 and PrrF2 of *P. aeruginosa* (103), and FsrA of *B. subtilis* (38), as well as NrrF in *Neisseria meningitidis* (65, 66). In addition to sRNAs, Fur may repress a proteinaceous repressor to activate downstream genes. For example, transcriptional activation of *hilA* in *S. enterica* serovar Typhimurium was demonstrated to result from the direct repression of a negative regulator of *hilA*, H-NS (histone-like nucleotide binding protein), by Fur (95). Fur has also been shown to directly bind to the promoter regions of Fur-activated genes. In *E. coli*, Fur and H-NS compete for overlapping binding sites within the promoter regions of *ftnA*, resulting in derepression of *ftnA* transcription (77). The transcription of the *V. cholerae* porin *ompT* gene was positively regulated by iron-bound Fur through the direct binding of Fur to the promoter region (17). However, in contrast to the wealth of studies describing Fur-mediated repression and despite the few previous examples, Fur-mediated direct activation via binding to defined promoter regions has been less studied.

In the remainder of this review, we discuss recent studies of Fur-mediated global regulatory circuits in the pathogenic *Neisseria* (*N. meningitidis* and *N. gonorrhoeae*). Since genomic analysis has revealed that there are fewer than 60 predicted regulatory proteins in the *Neisseria* genomes compared to ~200 in the *E. coli* genome (85), we propose that Fur-mediated regulation in these organisms may have more global and versatile consequences for gene expression and associated pathogenic mechanisms.

Fur REGULON OF *N. MENINGITIDIS*

An *N. meningitidis* global microarray analysis has identified 233 genes whose transcription levels were affected by growth under iron-replete versus -depleted conditions (44). Of these 233 genes, ~ 50% were predicted by *in silico* analysis to contain Fur boxes in their promoter regions (44). In addition, the majority of the predicted Fur binding promoter regions were experimentally demonstrated to bind Fur *in vitro* (44). A subsequent study examining iron regulation in a meningococcal *fur* mutant strain identified 83 genes whose iron-responsive regulation required Fur (22). Interestingly, 44 of those genes were repressed and 38 were activated, defining a new role for *Neisseria* Fur in activation of gene expression (22).

As demonstrated by electrophoretic mobility shift assay (EMSA) and/or DNase I footprinting results, genes and operons directly repressed by Fur in *N. meningitidis* encode proteins which can be classified into four major groups based on their functions: iron uptake and transport, energy metabolism and biosynthesis, toxin and stress responses, and regulation (Table 2) (22, 24, 44, 65, 66, 90). As expected, genes encoding iron uptake and transport proteins such as *tbpA*, *tbpB*, *lbpA*, and *lbpB* are repressed by Fur under iron-replete conditions, in agreement with the primary role of Fur as a maintainer of iron homeostasis (Table 2). The second group of Fur-repressed genes includes a large number of genes involved in energy metabolism and biosynthesis (Table 2). The protein products of these genes appear to enable bacterial growth, whereas their roles in pathogenesis have not yet been investigated. Interestingly, these genes have few homologs in *N. gonorrhoeae* (Table 2). The third group includes genes involved in virulence and bacterial adaption. It has been shown that FrpC-like proteins of *N. meningitidis* may play a role in pathogenesis (35, 79), and several *frpA- and frpC*-related gene loci, including NMB0364, NMB0584, NMB1405, and NMB1412 to -1414, were repressed by Fur directly (Table 2) (44). Several chaperone proteins and putative transposases, in addition to RecN, which are involved in DNA recombination and repair processes are also proposed to support bacterial adaption to the host environment and are directly repressed by Fur (Table 2) (44). The last group of Fur-repressed genes can be classified as regulators. A large percentage of Fur-dependent genes did not appear to be directly regulated by Fur, as demonstrated by the inability of Fur to bind to the promoter regions, which suggests the involvement of secondary regulators (22, 44). So far, only the small RNA NrrF has been identified as a transcriptional regulator directly controlled by Fur in *N. meningitidis* (Table 2) (65, 66). It is therefore logical to predict that additional, as-yet-uncharacterized Fur-controlled regulators could exist. Conversely, nine genes or operons directly activated by Fur fall into two major groups: iron storage and oxidative stress resistance genes and gene loci such as *sodB*, *kat*, *norB*, *aniA*, and NMB1438 to -1436 and energy metabolism loci such as the *nuo* complex (NMB0242 to -0244) (Table 2) (22, 44). In addition, a large number of hypothetical proteins under Fur regulation await further investigation (22, 44).

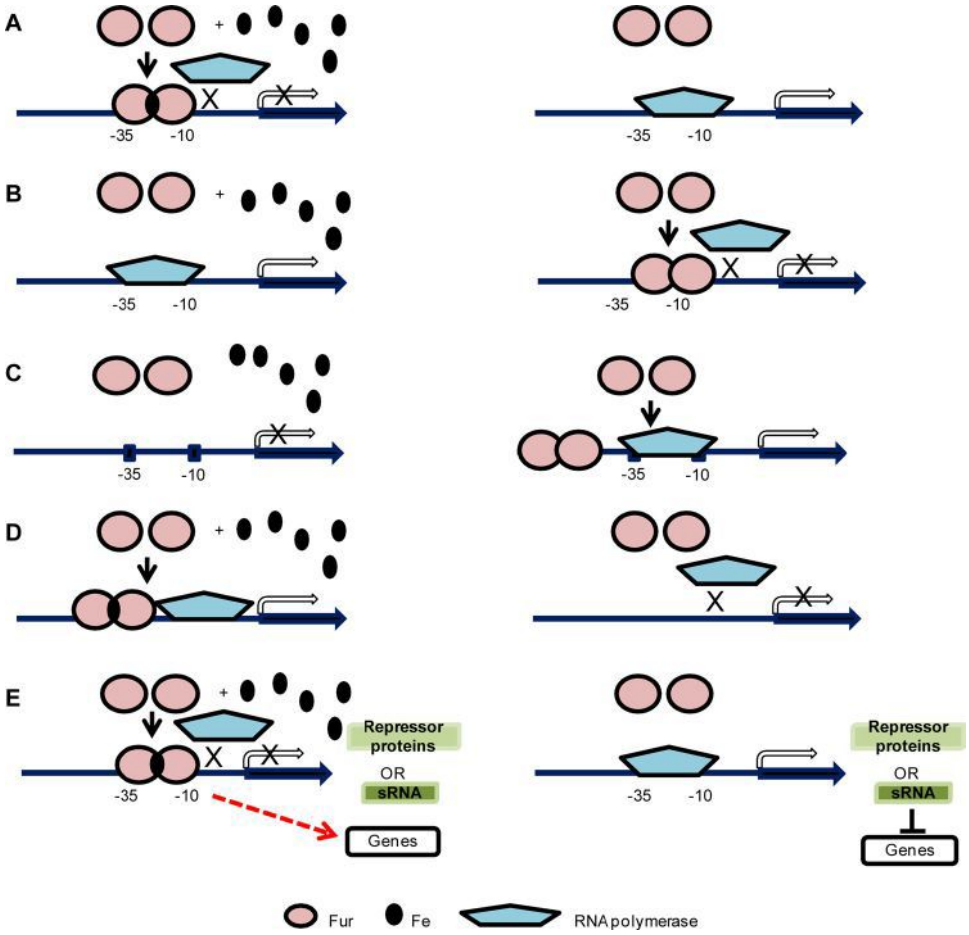


FIG 2 Mechanisms of Fur-mediated regulation. (A) Iron-bound Fur dimer binds to the −10 and −35 motifs in the promoter region, blocking binding of RNA polymerase and thus reducing transcription (left panel). Without iron, Fur does not bind to the promoter region, thereby allowing transcription via RNA polymerase (right panel). This mechanism has been well established by experimental evidence (32). (B) In the apo-Fur repression mechanism, when iron is present, Fur does not bind to the promoter region, resulting in transcription of the gene by RNA polymerase (left panel). Without iron, the apo-Fur dimer binds to the −10 and −35 motifs in the promoter region to inhibit binding of RNA polymerase and repress transcription (right panel). Apo-Fur repression was experimentally demonstrated in *H. pylori* (13, 72). (C) In the apo-Fur activation mechanism, Fur does not bind to the promoter region; thus, gene transcription is not active when iron is present (left panel). Without iron, the apo-Fur dimer binds to a site further upstream of the −10 and −35 motifs in the promoter region and upregulates transcription (right panel). This mechanism was first described in *V. vulnificus* (57). (D) Iron-bound Fur dimer binds to the promoter region and upregulates transcription (left panel). Without iron, Fur does not bind to the promoter region to initiate transcription (right panel). A few cases have been reported that indicate Fur-mediated direct activation through binding to defined promoter regions (50, 77, 105). (E) Iron-bound Fur dimer binds to −10 and −35 motifs in the promoter region and represses a negative regulator such as a protein repressor or a small RNA. Subsequently, genes that are repressed by the negative regulators are then transcribed and these genes show indirect Fur activation (left panel). Without iron and Fur repression, the negative regulators are transcribed and repress their target genes (right panel). Fur-repressed small RNAs have been reported in *E. coli* (61–64), *V. cholerae* (21), *P. aeruginosa* (103), and *B. subtilis* (38) as well as in the two pathogenic *Neisseria* species (28, 65, 66). Known Fur-repressed protein regulators include H-NS in *S. enterica* serovar Typhimurium (95) and MpeR in *N. gonorrhoeae* (53).

TABLE 1 Fur box consensus sequences of various Gram-positive or Gram-negative bacteria

Species	Consensus sequence of Fur box[a]	Reference(s)
Bacillus subtilis	TGATAATnATTATCA	7, 37
Staphylococcus aureus	GTTCATGATAATCATTATC	104
Escherichia coli	GATAATGATAATCATTATC	23
Pseudomonas aeruginosa	GATAATGATAATCATTATC	78
Salmonella enterica serovar Typhimurium	GATAATGATAATCATTATC	96
Vibrio cholerae	GATAATGATAATCATTATC	69
Helicobacter pylori	TAATAATnATTATTA	80, 98
Neisseria meningitidis	nATwATnATwATnATwATn	44
Neisseria gonorrhoeae	T-ATAAT-ATTATCA	27, 105

[a] See text for explanation of the significance of the hyphens and the uppercase and lowercase characters in the sequences.

Fur REGULON OF *N. GONORRHOEAE*

Microarray studies in *N. gonorrhoeae* determined that ~20% of the gonococcal genome is regulated in response to growth under iron-replete versus -depleted conditions (27, 53). When examined by *in silico* analysis, 92 genes or operons were predicted to contain a Fur box (27, 53). However, only a small percentage of these putative operator regions were demonstrated to bind Fur by a Fur titration assay (FurTa) (99), EMSA, and/or footprinting (Table 2) (25, 36, 41, 53, 105). Similar to *N. meningitidis*, gonococcal Fur-repressed genes include a large number involved in iron acquisition (Table 2). However, only three genes or loci encoding proteins for energy metabolism and biosynthesis (*fumC*, NGO0108, and NGO0114) are identified as being Fur repressed in *N. gonor-*

TABLE 2 Genes directly regulated by Fur in *N. meningitidis* and *N. gonorrhoeae*[a]

	N. meningitidis MC58				*N. gonorrhoeae* FA1090			
Category	Gene	Function	Expt(s)	Reference(s)	Gene	Function	Expt(s)	Reference(s)
Direct repression								
Iron acquisition	NMB0205	*fur*, ferric uptake regulator protein	EM, FP	44	NGO1779	*fur*	EM, FT	53, 86
	NMB0634	*fbpA*, iron binding protein	EM	44	NGO0215 to -0217	*fbpABC*	EM, FP, FT	25, 36, 53
	NMB1668	*hmbR*, hemoglobin receptor	FP	22	NGO1318	*hemO-hemR*, heme utilization protein	EM, FT	53, 86
	NMB0460	*tbp2*, transferrin binding protein B	EM	44	NGO1496	*tbpB*	FT	53
	NMB0461	*tbp1*, transferrin binding protein A	EM	44	NGO1495	*tbpA*	FT	53
	NMB1540	*lbpA*, lactoferrin binding protein A	EM	44		*lbpA*[b]	P[c]	9, 41
	NMB1541	*lbpB*, lactoferrin binding protein B	EM, FP	22, 44		*lbpB*[b]	P	8
	NMB1730	*tonB*, energy transducer	FP	22	NGO2176	*tonB*	EM	86
	NMB1728	*exbD*, biopolymer transport protein	EM	44				
	NMB0175	*zupT*, zinc transporter	FP	22				
	NMB1988	*frpB* (*fetA*), ferric enterobactin receptor	EM, FP	22, 44	NGO2093	*fetA*	FT	53
					NGO2092	*fetB*, ferric enterobactin periplasmic binding protein	EM, FT	53
					NGO0024	putative FetB2 protein	FT	53
					NGO0553	*tdfG*, putative TonB-dependent receptor	FT	53
					NGO2109	*hpuB*, hemoglobin-haptoglobin utilization protein B	FT	53
RTX toxin/virulence	NMB0364	FrpA/C-related protein	EM	44				
	NMB0584	FrpA/C-related protein	EM	44				
	NMB1405	FrpA/C-related protein	EM	44				
	NMB1412 to -1414	FrpA/C-related protein	EM	44				
					NGO0275	IgA1 protease	FT	53
						Opa (opacity-associated protein) A–K	EM	86
					NGO1822	*secY*	EM	86
Adaption/stress response					NGO0449	*sodB*	EM	86
					NGO0652	Thioredoxin I	FT	53
	NMB0544	*dnaK*, heat shock protein, chaperone	FP	22				
	NMB1472	*clpB*, chaperone	FP	22				
	NMB0740	*recN*, DNA repair protein	EM	44	NGO0318	*recN*	EM	86
	NMB0101	Putative transposase	EM	44				
	NMB1798	Putative transposase	EM	44				
Energy metabolism	NMB1395 and -1396	Alcohol dehydrogenase/*mutY*, A/G-specific adenine glycosylase	EM, FP	22, 44				
	NMB1377	*lldD*, L-lactate dehydrogenase	EM, FP	22, 44				
	NMB1458	*fumC*, fumarase C in TCA cycle	EM, FP	22, 44	NGO1029	*fumC*, fumarase C in TCA cycle	EM	86
					NGO0108	Putative oxidoreductase	FT	53
					NGO0114	Putative glutaredoxin	FT	53
Biosynthesis	NMB1898	*mlp*, lipoprotein	EM	44				
	NMB0317	7-cyano-7-deazaguanine reductase	EM	44				
	NMB0294	*dsbA-2*, disulfide interchange protein	EM	44				
	NMB0343	YciI-like protein	EM	44				
	NMB0394	*nadA*, quinolinate synthetase	EM	44				
	NMB0396	*nadC*, nicotinate-nucleotide pyrophosphorylase	EM	44				
	NMB1381	HesB/YadR/YfhF family protein	EM	44				
	NMB1380	*nifU*, nitrogen fixation	EM	44				
Regulators		*nrrF*, small RNA, transcriptional regulator	EM	65, 66		*nrrF*	P	28
					NGO0025	*mpeR*, AraC-like regulator	FT	53
Hypothetical	NMB0034 to -0036	Hypothetical protein	EM, FP	22, 44				
	NMB0744	Hypothetical protein	EM	44	NGO0322	Hypothetical protein	FT	53
	NMB0821	Hypothetical protein	EM	44				
	NMB0865 and -0864	Hypothetical protein	EM	44				

(Continued on following page)

TABLE 2 (Continued)

Category	*N. meningitidis* MC58				*N. gonorrhoeae* FA1090			
	Gene	Function	Expt(s)	Reference(s)	Gene	Function	Expt(s)	Reference(s)
	NMB1340	Hypothetical protein	EM	44				
	NMB1491	Hypothetical protein	EM	44				
	NMB1796	Hypothetical protein	FP	22				
	NMB1879 and -1880	Hypothetical protein	FP	22				
					NGO0554	Hypothetical protein	FT	53
Direct activation								
Iron acquisition/ storage					NGO1205	Putative TonB-dependent receptor	EM, FP	105
					NGO0794	*bfrA*, bacterioferritin	EM, FP, FT	53, 105
Energy metabolism	NMB1613	*fumB*, fumarate hydratase	FP	22				
	NMB0242 to -0244	*nuoB-nuoD*, NADH dehydrogenase subunits	EM, FP	23, 44	NGO1748 to -1751	*nuo* operon	EM, FP, FT	53, 105
					NGO0711	Alcohol dehydrogenase	EM, FP, FT	53, 105
					NGO2116	ATP-binding protein	EM, FP, FT	53, 105
					NGO0076	Putative phosphatase	EM, FP, FT	53, 105
Adaption/stress response	NMB0663	*nspA*, neisserial surface protein A	EM, FP	44, 90	NGO0233	*nspA*	EM, FP	105
	NMB1622	*norB*, nitric oxide reductase	FP	23	NGO1275	*norB*	EM, FP, FT	53, 105
	NMB1623	*aniA*, nitrite reductase	EM, FP	23, 44	NGO1276	*aniA*	EM, FP, FT	53, 105
	NMB0884	*sodB*, superoxide dismutase	EM	44				
	NMB0216	*kat*, catalase	EM	44				
	NMB1436 to -1438	Hypothetical proteins	EM, FP	22, 44	NGO0904 to -0906	Hypothetical Fe-S protein; hypothetical protein; Fe-S oxidoreductase	EM, FP, FT	53, 105
					NGO1317	Transposase	EM, FP, FT	53, 105
Transcription/ regulation					NGO0199	Transcription termination factor Rho	EM, FP	105
					NGO1851	DNA direct RNA polymerase subunit β	EM, FP	105
Hypothetical	NMB0298	Hypothetical protein	EM	44				
					NGO1207 to -1209	Excinuclease ABC subunit A; restriction endonuclease R.NgoMIII; DNA cytosine methyltransferase M.NgoMIII	EM, FP	105
					NGO1282	Hypothetical protein	EM, FP	105
					NGO1430	Hypothetical protein	EM, FP, FT	53, 105

[a] EM, electrophoresis mobility shift assay (EMSA): purified protein is incubated with radiolabeled probes (DNA or RNA) and subsequently run on a native polyacrylamide gel. Protein-bound probes show less mobility and shift up compared to free probes. FP, footprinting: DNase I is used to cut one end of labeled DNA, and the resulting patterns are analyzed by gel electrophoresis. The protein-bound site on the DNA is protected from cleavage and results in a clear area. FT (FurTa) (99), Fur titration assay: a bacterial genomic DNA library is constructed on a multicopy plasmid, such as puc18. The plasmids are transformed into an *E. coli* strain deficient in enterochelin synthesis and containing a fusion construct of the promoter of *fhuF*::*lacZ* in the chromosome. The promoter of *fhuF* has weak affinity to the $FurFe^{2+}$ repressor. If the multicopy plasmids do not contain a Fur box, $FurFe^{2+}$ represses the promoter region of *fhuF*::*lacZ* fusion construct so that the strain produces Lac (white) colonies on MacConkey plates supplemented with iron. In contrast, if the multicopy plasmids contain a Fur box, then the high number of these Fur boxes competes with the binding of $FurFe^{2+}$ to derepress the promoter region of *fhuF*::*lacZ* fusion construct, which results in Lac-positive colonies on MacConkey plates supplemented with iron.

[b] *lbpA* and *lbpB* genes are not present in *N. gonorrhoeae* strain FA1090 but were identified in strain FA19.

[c] Predicted to be directly regulated by Fur according to the presence of an *in silico* Fur box in the promoter region of the gene.

rhoeae (Table 2) (53, 86). The Fur-repressed virulence-associated genes and stress response genes are also different from those in *N. meningitidis* (Table 2). Genes encoding IgA1 protease, which cleaves human IgA on the mucosal surface (88), SecY, a putative preprotein translocase (86), and the opacity-associated Opa proteins (A to K) are directly repressed by Fur only in *N. gonorrhoeae* (Table 2), although *N. meningitidis* also contains these genes (18, 90). Interestingly, *sodB*, which is one of the Fur-activated genes in *N. meningitidis*, is repressed by Fur in the gonococcus (Table 2). In addition to *sodB*, only one gene and one locus, *recN* and NGO0652, are identified as stress response genes in *N. gonorrhoeae* (Table 2). NGO0652 encodes a putative thioredoxin pro-

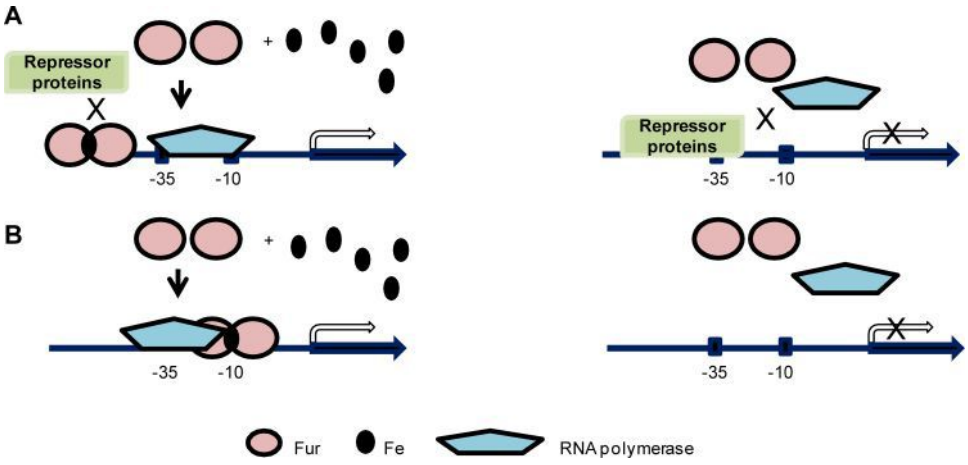


FIG 3 Fur mediates direct activation via variable mechanisms. (A) An iron-bound Fur dimer binds to a site further upstream of −10 and −35 motifs in the promoter region and out-competes binding of a repressor protein, thus activating transcription (left panel). Without iron, Fur does not function and the repressor protein binds to the promoter region to repress gene transcription (right panel). Three examples have been reported in *E. coli* and *N. gonorrhoeae* (50, 77). (B) An iron-bound Fur dimer binds to a site overlapping the −10 and −35 motifs in the promoter region in order to recruit RNA polymerase to activate transcription (left panel). Without iron, Fur does not bind to the promoter region and RNA polymerase is not recruited to initiate transcription (right panel). This mechanism is still a hypothesized model without experimental evidence (105).

tein (53), and a thioredoxin-like protein in the gonococcus has been suggested to play a role in defense against oxidative stress (1). More Fur-activated genes and operons have been identified in the gonococci than in the meningococci to date (Table 2). In addition to *nspA*, *norB*, *aniA*, and the *nuo* operon that have already been reported in the meningococcus, Fur-activated genes and loci in the gonococcus include those involved in iron storage and transport (NGO1205 and *bfrA*), transcription and regulation (NGO0199 and NGO1851), energy metabolism (NGO0711 and NGO2116), and adaption (NGO0076 and NGO1317) (Table 2) (53, 105). Their roles in gonococcal pathogenesis have yet to be investigated.

MECHANISMS OF Fur-MEDIATED DIRECT AND INDIRECT REGULATION IN PATHOGENIC *NEISSERIA*

Direct Fur repression. The repression mechanism of Fur, as well as the Fur box consensus sequence, is similar to that determined in other bacterial species (22, 25, 36, 43, 44, 86). Fur dimers bind to the −10 and −35 motifs in the promoter region to prevent the binding of RNA polymerase, which results in transcriptional repression of the gene. The apo-Fur repression mechanism may also be present in *Neisseria*, since several genes of *N. gonorrhoeae* have been shown to be repressed by Fur under iron-depleted conditions (105).

Direct Fur activation. The mechanism by which Fur functions to directly activate gene transcription is poorly understood. In general, proteins which function to activate transcription utilize one of the following three pathways: (i) the activator forces a repressor out of a potential binding site, allowing the initiation of transcription; (ii) the activator recruits RNA polymerase to enhance transcription; or (iii) the activator binding alters DNA morphology, allowing RNA polymerase binding (11). In addition, multiple mechanisms may be utilized simultaneously (11). Thus, in contrast to the simple mechanism of Fur-mediated repression, Fur-mediated direct activation may result from several different pathways.

Examples of the first pathway have been characterized in *Neisseria*. In *N. gonorrhoeae*, Fur-mediated activation of the *norB* gene is accomplished via the Fur box overlapping with the binding site of another repressor, ArsR. In this scenario, Fur binding to the promoter region of *norB* competes with ArsR and results in derepression of transcription (50). The position of the Fur box in the promoter region is at a distance further upstream from the −10 and −35 motifs than those in the genes repressed by Fur (Fig. 3A). Similar positioning of Fur boxes were reported in *aniA*, *norB*, and *nuoA* in *N. meningitidis* and in NGO0199, NGO1275 (*norB*), NGO1276 (*aniA*), and NGO1282 in *N. gonorrhoeae* (22, 105), suggesting that these genes could be activated by Fur via exclusion of a repressor. Interestingly, recent identifications of genes directly activated by Fur and characterization of their respective Fur boxes have shown that the Fur boxes of a majority of these genes in *N. gonorrhoeae* are in fact localized close to the −10 and −35 motifs, either overlapping both motifs or downstream of the −10 motif (105). This suggests an alternative mechanism utilized by Fur to directly initiate transcription. It is hypothesized that binding of Fur to a position close to the −10 and −35 motifs may recruit RNA polymerase binding in order to enhance transcription initiation (Fig. 3B) (105). Depending on the subtle differences of the positions of Fur box relative to the promoter motifs, Fur may interact with different subunits of RNA polymerase, although evidence for the direct interaction between Fur and subunits of RNA polymerase has not yet been established. However, we cannot rule out the possibility that Fur may utilize unidentified mechanisms other than ones discussed above to activate gene transcription.

Indirect Fur regulation. Fur-mediated indirect regulation in pathogenic *Neisseria* can occur via sRNA-mediated regulation or a regulator protein. Fur-mediated activation has been demonstrated to function indirectly through repressing a sRNA, NrrF, in *N. meningitidis*, which negatively regulates its target genes (65, 66). Nearly all sRNAs described in a variety of Gram-positive and Gram-negative bacteria share four common characteristics: (i) they are mainly localized in intergenic regions; (ii) the sequences are highly conserved among most genetically similar species; (iii)

the 3′ termini of sRNAs usually contain a Rho-independent terminator structure, which is composed of a stem-loop followed by a polyuridine region; and (iv) the lengths of sRNAs range from 50 to 300 nucleotides (100). Generally, sRNAs base pair to the ribosome-binding site on target mRNAs to interfere with ribosome binding and block the initiation of translation (100). The complementarity between a sRNA and mRNA requires a "seed" region of at least 8 to 9 continuous base pairs (100). Alternatively, sRNAs may lead to decreased stability of target mRNAs, resulting in reduced translation (100). In addition, many of the known sRNAs require a protein cofactor, Host Factor Qβ-phage (Hfq), to facilitate the binding of sRNAs to mRNA (42, 97).

The meningococcal sRNA NrrF (between loci NMB2073 and NMB2074) was identified by screening intergenic regions for Fur boxes upstream of Rho-independent terminators (65). One of 19 possible candidates, NrrF, was found to be iron regulated via Fur (65). NrrF is upregulated under iron-depleted conditions in a wild-type strain and derepressed in the *fur* mutant strain in a manner independent of the presence of iron, as determined by both Northern blot and reverse transcription-PCR (RT-PCR) experiments (65). The regulatory targets of NrrF were predicted using a bioinformatics approach which identified the *sdhA-sdhC* operon as a possible target (65). Subsequent experiments examining transcription of *sdhA-sdhC* in a wild-type strain and a *nrrF* mutant strain under iron-depleted or iron-replete conditions confirmed an NrrF-dependent repression pattern for *sdhA-sdhC* (65). Interestingly, unlike the Fur-regulated sRNAs in other bacteria, NrrF can function in the absence of Hfq. In an Δ*hfq* strain, the regulation of *sdhA-sdhC* in response to iron availability is unchanged, as is the stability of NrrF (66). NrrF is the first Fur-controlled negative regulator that has been discovered in pathogenic *Neisseria* (65, 66). Targets of NrrF, *sdhA-sdhC*, are in turn indirectly activated by Fur (65, 66). A NrrF homologue (between NGO2002 and NGO2004) has also been found in the *N. gonorrhoeae* genome (28).

To date, one *Neisseria* regulatory protein which is under Fur-mediated direct regulation has been reported. The *N. gonorrhoeae* Fur-repressed MpeR protein, an AraC-like regulator (53), activates *fetA*, an outer membrane transporter required for acquisition of xenosiderophore ferric enterobactin as an iron source (47) and represses MtrF and MtrR, which function in the *mtr* efflux pump and modulate antimicrobial resistance systems (34, 67, 89).

Fur GLOBAL REGULATORY CIRCUITS AND CROSS-TALK WITH NEWLY DEFINED REGULONS

With the discovery of increasing numbers of Fur-regulated genes, additional regulators have been found to cooperate with Fur to control the same gene or operon. For example, in *N. meningitidis*, Fur-regulated *hemO* and *hmbR* were shown to be positively regulated by the two-component system MisR/S under both iron-depleted and -replete conditions (106). Transcription of *N. meningitidis kat*, which is activated by Fur, is also repressed by OxyR in the absence of H_2O_2 and activated by OxyR with H_2O_2 (49). Perhaps the best-studied cooperative regulation has been shown for the gonococcal *aniA* and *norB* genes, involved in the anaerobic respiration pathway in *N. gonorrhoeae*. Several regulators are utilized for *aniA* and *norB* regulation, including Fur, NsrR, NarQ/P, FNR, and ArsR (50, 52). NsrR is a repressor containing a [2Fe-2S] cluster, which is involved in NO sensing (50, 52). In addition to iron limitation, the pathogenic *Neisseria* spp. may also encounter other stresses in the host environment, such as the simultaneous presence of NO, H_2O_2, and pH. The gonococcal Fur regulon has also been shown to overlap anaerobic and hydrogen peroxide regulons (51, 92). Thus, it is highly likely that Fur functions together with other regulators to enable *Neisseria* spp. to respond to complicated environmental conditions or stimuli within the human host.

We propose that the *Neisseria* Fur regulon encompasses a complicated network due to the ability of this protein to function as either a repressor or an activator in both direct and indirect pathways. To date, studies in the pathogenic *Neisseria* have examined only a small subset of Fur-regulated genes and have had limited value in deciphering a Fur-mediated global network. High-throughput bioinformatics techniques designed to assist in the analysis of the relationships among regulons of different regulators (19) should help to identify the entire *Neisseria* Fur regulon.

BIOLOGICAL ROLES OF Fur REGULATORY CIRCUITS

Although the genomes of *N. meningitidis* and *N. gonorrhoeae* are closely related, the Fur regulons of these two organisms are not completely overlapping (Table 2). These differences may relate to pathogen-specific requirements during human colonization and associated inflammatory pathologies.

N. gonorrhoeae mainly colonizes the human urethra, endocervix, fallopian tubes, and uterus. This pathogen causes urethritis in men, with obvious inflammatory symptoms such as a purulent discharge with an influx of polymorphonuclear leukocytes (PMN). In women, gonococcal infection presents as cervicitis, vaginitis, or a more serious pelvic inflammatory disease. Also, infection in women is typically asymptomatic and may lead to serious complications, including endometritis, salpingitis, and disseminated gonococcal infection (DGI) (30). *N. meningitidis* frequently colonizes the nasopharynx and causes meningitis or septicemia upon entering the cerebrospinal fluid or bloodstream, respectively. Generally, host niches for both organisms are iron-depleted environments, as free iron in the human host is scarce. Therefore, genes involved in iron acquisition, including *fur* itself, are upregulated in both pathogens during *N. meningitidis* infection in human blood and in the *N. gonorrhoeae* RNA isolated from cervical swab specimens from women with uncomplicated gonorrhea or urethral swab specimens from men with urethral infections (2, 3, 29). However, one of the differences between species-specific host niches is the iron source. In serum, transferrin is the major iron-carrying protein, while in the mucosal surfaces, lactoferrin is the major iron source (14). In addition, the concentration of lactoferrin can change with the menstrual cycle in human vaginal mucus (15). All *N. meningitidis* strains are able to utilize both transferrin and lactoferrin (70, 71), which may guarantee the survival of *N. meningitidis* in both types of host niches. In contrast, not all *N. gonorrhoeae* strains contain *lbpA* and *lbpB* genes, required for utilizing lactoferrin (8, 9, 41, 70). It has been proposed that gonococcal strains unable to utilize lactoferrin may be related to the asymptomatic infections often observed in women (10).

In *N. meningitidis*, the *aniA* (nitrite reductase), *kat* (catalase), and *nspA* (*Neisseria* surface protein A) genes are under direct Fur activation and have been demonstrated to be upregulated during *N. meningitidis* colonization in human blood (29). The *kat* and *aniA* genes have been postulated to play a role in *N. meningitidis* survival under conditions of stress from reactive oxygen and nitrogen species from neutrophils and macrophages (4, 87). Inter-

estingly, the *kat* gene is not under iron or Fur-mediated control in the gonococcus (53). The *nspA* gene encodes a human factor H binding protein that facilitates resistance to human complement, resulting in enhanced survival of *N. meningitidis* in blood (29, 60). Furthermore, genes under Fur-mediated direct activation, including *nspA* and *aniA*, were not detected in specimens isolated from women with gonococcal infection (3). Another significant difference between *N. meningitidis* and *N. gonorrhoeae* Fur regulons during infection may relate to the expression of Opa proteins, which are involved in bacterial adherence and invasion of human epithelial cells and neutrophils (18). *N. meningitidis* contains 3 to 4 Opa proteins, and at least one of them (NMB1636) is upregulated in human blood (29), although none of meningococcal *opa* genes are regulated via Fur (22, 44). In contrast, the putative promoter regions of all 11 gonococcal *opa* genes (A to K) have been shown to bind directly to Fur (86). Gonococcal Opa proteins have been shown to be expressed during natural infections as well as in experimentally infected volunteers (54, 55, 93). Furthermore, Opa–carcinoembryonic antigen-related cell adhesion molecule (CEACAM) interactions have been shown to promote gonococcal colonization in mouse models (16, 75, 76, 82). Above all, these results suggest that specific colonization niches and the associated pathogenic processes of the two pathogenic species serve to define the Fur regulons of these organisms and, in particular, those genes which are activated by Fur.

Moreover, the global effects resulting from Fur-mediated regulation determine the virulence of several bacterial pathogens. A large number of genes encoding iron-uptake protein homologues in the *Staphylococcus aureus* genome contain a predicted Fur box in their promoter regions, suggesting a major role of Fur in iron homeostasis (48). A *fur* mutant strain of *S. aureus* showed growth defects and higher sensitivity to H_2O_2 (48). These characteristics may have led to the reduced virulence (lower recovery of the *fur* mutant strain compared to the wild-type strain) seen in a murine skin abscess model (48). Fur of *Bacillus cereus*, an opportunistic human pathogen that causes food poisoning and endophthalmitis, may control at least 16 genes, according to the results of Fur-box prediction in the promoter regions of the genome (46). These genes include those involved in iron uptake and storage, secondary cellular metabolism, and virulence. The *fur* mutant strain of *B. cereus* also shows reduced virulence, with a 50% lethal dose (LD_{50}) value of 4,932 CFU compared to an LD_{50} value of 1,859 CFU of the wild-type strain in an insect model (46). Similarly, *V. cholerae* has been shown to have at least 65 Fur iron-repressed genes which are involved in iron acquisition and metabolism and two genes indirectly activated by Fur (69). In an infant mouse model of intestinal colonization, the *V. cholerae fur* mutant strain displayed significantly attenuated colonization when competing with the wild-type strain (69). Fur of the gastric pathogen *H. pylori* regulates genes critical for acid acclimation and oxidative stress (39, 80). The *fur* mutant strain of *H. pylori* displays a 100-fold-higher 50% infectious dose than the wild-type strain and lower colonization ability when competing with the wild-type strain in the Mongolian gerbil model (39, 73, 80). In addition, the *fur* mutant strain showed an attenuated ability to induce host inflammation and injury (73). All of the studies noted above emphasized the importance of the Fur regulon in bacterial pathogenesis.

CONCLUDING REMARKS

Recent studies have begun to define the Fur regulons of the pathogenic *Neisseria*. For both *N. meningitidis* and *N. gonorrhoeae*, global mechanisms of transcriptional control by Fur have been linked to the ability of these pathogens to cause disease and to respond to various stimuli within the human host. We predict that Fur-regulated circuits embrace broad components in the genomes and enable these organisms to respond to a variety of stress situations. Thus, understanding pathogenic aspects of Fur-mediated regulation is critical in revealing bacterial pathogenic mechanisms and will help to discover new therapeutic targets in these pathogens.

REFERENCES

1. **Achard ME, et al.** 2009. A periplasmic thioredoxin-like protein plays a role in defense against oxidative stress in Neisseria gonorrhoeae. Infect. Immun. **77**:4934–4939.
2. **Agarwal S, et al.** 2005. The gonococcal Fur-regulated tbpA and tbpB genes are expressed during natural mucosal gonococcal infection. Infect. Immun. **73**:4281–4287.
3. **Agarwal S, Sebastian S, Szmigielski B, Rice PA, Genco CA.** 2008. Expression of the gonococcal global regulatory protein Fur and genes encompassing the Fur and iron regulon during in vitro and in vivo infection in women. J. Bacteriol. **190**:3129–3139.
4. **Anjum MF, Stevanin TM, Read RC, Moir JW.** 2002. Nitric oxide metabolism in Neisseria meningitidis. J. Bacteriol. **184**:2987–2993.
5. **Argaman L, et al.** 2001. Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. Curr. Biol. **11**:941–950.
6. **Bagg A, Neilands JB.** 1987. Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in Escherichia coli. Biochemistry **26**:5471–5477.
7. **Baichoo N, Helmann JD.** 2002. Recognition of DNA by Fur: a reinterpretation of the Fur box consensus sequence. J. Bacteriol. **184**:5826–5832.
8. **Biswas GD, Anderson JE, Chen CJ, Cornelissen CN, Sparling PF.** 1999. Identification and functional characterization of the Neisseria gonorrhoeae lbpB gene product. Infect. Immun. **67**:455–459.
9. **Biswas GD, Sparling PF.** 1995. Characterization of lbpA, the structural gene for a lactoferrin receptor in Neisseria gonorrhoeae. Infect. Immun. **63**:2958–2967.
10. **Britigan BE, Cohen MS, Sparling PF.** 1985. Gonococcal infection: a model of molecular pathogenesis. N. Engl. J. Med. **312**:1683–1694.
11. **Browning DF, Busby SJ.** 2004. The regulation of bacterial transcription initiation. Nat. Rev. Microbiol. **2**:57–65.
12. **Butcher BG, et al.** 2011. Characterization of the Fur Regulon in Pseudomonas syringae pv. tomato DC3000. J. Bacteriol. **193**:4598–4611.
13. **Carpenter BM, Whitmire JM, Merrell DS.** 2009. This is not your mother's repressor: the complex role of fur in pathogenesis. Infect. Immun. **77**:2590–2601.
14. **Chen CY, Genco CA, Rock JP, Morse SA.** 1989. Physiology and metabolism of Neisseria gonorrhoeae and Neisseria meningitidis: implications for pathogenesis. Clin. Microbiol. Rev. **2**(Suppl):S35–S40.
15. **Cohen MS, Britigan BE, French M, Bean K.** 1987. Preliminary observations on lactoferrin secretion in human vaginal mucus: variation during the menstrual cycle, evidence of hormonal regulation, and implications for infection with Neisseria gonorrhoeae. Am. J. Obstet. Gynecol. **157**:1122–1125.
16. **Cole JG, Fulcher NB, Jerse AE.** 2010. Opacity proteins increase Neisseria gonorrhoeae fitness in the female genital tract due to a factor under ovarian control. Infect. Immun. **78**:1629–1641.
17. **Craig SA, Carpenter CD, Mey AR, Wyckoff EE, Payne SM.** 2011. Positive regulation of the Vibrio cholerae porin OmpT by iron and fur. J. Bacteriol. **193**:6505–6511.
18. **Criss AK, Seifert HS.** 2012. A bacterial siren song: intimate interactions between Neisseria and neutrophils. Nat. Rev. Microbiol. **10**:178–190.
19. **Danielli A, Scarlato V.** 2010. Regulatory circuits in Helicobacter pylori: network motifs and regulators involved in metal-dependent responses. FEMS Microbiol. Rev. **34**:738–752.
20. **Davies BW, Bogard RW, Mekalanos JJ.** 2011. Mapping the regulon of

Vibrio cholerae ferric uptake regulator expands its known network of gene regulation. Proc. Natl. Acad. Sci. U. S. A. **108**:12467–12472.
21. **Davis BM, Quinones M, Pratt J, Ding Y, Waldor MK.** 2005. Characterization of the small untranslated RNA RyhB and its regulon in Vibrio cholerae. J. Bacteriol. **187**:4005–4014.
22. **Delany I, Grifantini R, Bartolini E, Rappuoli R, Scarlato V.** 2006. Effect of Neisseria meningitidis fur mutations on global control of gene transcription. J. Bacteriol. **188**:2483–2492.
23. **Delany I, Rappuoli R, Scarlato V.** 2004. Fur functions as an activator and as a repressor of putative virulence genes in Neisseria meningitidis. Mol. Microbiol. **52**:1081–1090.
24. **de Lorenzo V, Wee S, Herrero M, Neilands JB.** 1987. Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (fur) repressor. J. Bacteriol. **169**:2624–2630.
25. **Desai PJ, Angerer A, Genco CA.** 1996. Analysis of Fur binding to operator sequences within the Neisseria gonorrhoeae fbpA promoter. J. Bacteriol. **178**:5020–5023.
26. **Dian C, et al.** 2011. The structure of the Helicobacter pylori ferric uptake regulator Fur reveals three functional metal binding sites. Mol. Microbiol. **79**:1260–1275.
27. **Ducey TF, Carson MB, Orvis J, Stintzi AP, Dyer DW.** 2005. Identification of the iron-responsive genes of Neisseria gonorrhoeae by microarray analysis in defined medium. J. Bacteriol. **187**:4865–4874.
28. **Ducey TF, Jackson L, Orvis J, Dyer DW.** 2009. Transcript analysis of nrrF, a Fur repressed sRNA of Neisseria gonorrhoeae. Microb. Pathog. **46**:166–170.
29. **Echenique-Rivera H, et al.** 2011. Transcriptome analysis of Neisseria meningitidis in human whole blood and mutagenesis studies identify virulence factors involved in blood survival. PLoS Pathog. **7**:e1002027. doi:10.1371/journal.ppat.1002027.
30. **Edwards JL, Apicella MA.** 2004. The molecular mechanisms used by Neisseria gonorrhoeae to initiate infection differ between men and women. Clin. Microbiol. Rev. **17**:965–981.
31. **Ernst JF, Bennett RL, Rothfield LI.** 1978. Constitutive expression of the iron-enterochelin and ferrichrome uptake systems in a mutant strain of Salmonella typhimurium. J. Bacteriol. **135**:928–934.
32. **Escolar L, de Lorenzo V, Perez-Martin J.** 1997. Metalloregulation in vitro of the aerobactin promoter of Escherichia coli by the Fur (ferric uptake regulation) protein. Mol. Microbiol. **26**:799–808.
33. **Escolar L, Perez-Martin J, de Lorenzo V.** 1998. Binding of the fur (ferric uptake regulator) repressor of Escherichia coli to arrays of the GATAAT sequence. J. Mol. Biol. **283**:537–547.
34. **Folster JP, Shafer WM.** 2005. Regulation of mtrF expression in Neisseria gonorrhoeae and its role in high-level antimicrobial resistance. J. Bacteriol. **187**:3713–3720.
35. **Forman S, et al.** 2003. Neisseria meningitidis RTX proteins are not required for virulence in infant rats. Infect. Immun. **71**:2253–2257.
36. **Forng RY, Ekechukwu CR, Subbarao S, Morse SA, Genco CA.** 1997. Promoter mapping and transcriptional regulation of the iron-regulated Neisseria gonorrhoeae fbpA gene. J. Bacteriol. **179**:3047–3052.
37. **Fuangthong M, Helmann JD.** 2003. Recognition of DNA by three ferric uptake regulator (Fur) homologs in Bacillus subtilis. J. Bacteriol. **185**:6348–6357.
38. **Gaballa A, et al.** 2008. The Bacillus subtilis iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins. Proc. Natl. Acad. Sci. U. S. A. **105**:11927–11932.
39. **Gancz H, Censini S, Merrell DS.** 2006. Iron and pH homeostasis intersect at the level of Fur regulation in the gastric pathogen Helicobacter pylori. Infect. Immun. **74**:602–614.
40. **Gao H, et al.** 2008. The iron-responsive Fur regulon in Yersinia pestis. J. Bacteriol. **190**:3063–3075.
41. **Genco CA, Desai PJ.** 1996. Iron acquisition in the pathogenic Neisseria. Trends Microbiol. **4**:179–184.
42. **Gottesman S.** 2004. The small RNA regulators of Escherichia coli: roles and mechanisms*. Annu. Rev. Microbiol. **58**:303–328.
43. **Grifantini R, et al.** 2004. Characterization of a novel Neisseria meningitidis Fur and iron-regulated operon required for protection from oxidative stress: utility of DNA microarray in the assignment of the biological role of hypothetical genes. Mol. Microbiol. **54**:962–979.
44. **Grifantini R, et al.** 2003. Identification of iron-activated and -repressed Fur-dependent genes by transcriptome analysis of Neisseria meningitidis group B. Proc. Natl. Acad. Sci. U. S. A. **100**:9542–9547.
45. **Hantke K.** 1984. Cloning of the repressor protein gene of iron-regulated systems in Escherichia coli K12. Mol. Gen. Genet. **197**:337–341.
46. **Harvie DR, Vilchez S, Steggles JR, Ellar DJ.** 2005. Bacillus cereus Fur regulates iron metabolism and is required for full virulence. Microbiology **151**:569–577.
47. **Hollander A, Mercante AD, Shafer WM, Cornelissen CN.** 2011. The iron-repressed, AraC-like regulator MpeR activates expression of fetA in Neisseria gonorrhoeae. Infect. Immun. **79**:4764–4776.
48. **Horsburgh MJ, Ingham E, Foster SJ.** 2001. In Staphylococcus aureus, fur is an interactive regulator with PerR, contributes to virulence, and is necessary for oxidative stress resistance through positive regulation of catalase and iron homeostasis. J. Bacteriol. **183**:468–475.
49. **Ieva R, et al.** 2008. OxyR tightly regulates catalase expression in Neisseria meningitidis through both repression and activation mechanisms. Mol. Microbiol. **70**:1152–1165.
50. **Isabella V, et al.** 2008. cis- and trans-acting elements involved in regulation of norB (norZ), the gene encoding nitric oxide reductase in Neisseria gonorrhoeae. Microbiology **154**:226–239.
51. **Isabella VM, Clark VL.** 2011. Deep sequencing-based analysis of the anaerobic stimulon in Neisseria gonorrhoeae. BMC Genomics **12**:51. doi:10.1186/1471-2164-12-51.
52. **Isabella VM, Lapek JD, Jr, Kennedy EM, Clark VL.** 2009. Functional analysis of NsrR, a nitric oxide-sensing Rrf2 repressor in Neisseria gonorrhoeae. Mol. Microbiol. **71**:227–239.
53. **Jackson LA, et al.** 2010. Transcriptional and functional analysis of the Neisseria gonorrhoeae Fur regulon. J. Bacteriol. **192**:77–85.
54. **James JF, Swanson J.** 1978. Studies on gonococcus infection. XIII. Occurrence of color/opacity colonial variants in clinical cultures. Infect. Immun. **19**:332–340.
55. **Jerse AE, et al.** 1994. Multiple gonococcal opacity proteins are expressed during experimental urethral infection in the male. J. Exp. Med. **179**:911–920.
56. **Ledala N, Sengupta M, Muthaiyan A, Wilkinson BJ, Jayaswal RK.** 2010. Transcriptomic response of Listeria monocytogenes to iron limitation and Fur mutation. Appl. Environ. Microbiol. **76**:406–416.
57. **Lee HJ, Bang SH, Lee KH, Park SJ.** 2007. Positive regulation of fur gene expression via direct interaction of fur in a pathogenic bacterium, Vibrio vulnificus. J. Bacteriol. **189**:2629–2636.
58. **Lee JW, Helmann JD.** 2007. Functional specialization within the Fur family of metalloregulators. Biometals **20**:485–499.
59. **Lewin AC, Doughty PA, Flegg L, Moore GR, Spiro S.** 2002. The ferric uptake regulator of Pseudomonas aeruginosa has no essential cysteine residues and does not contain a structural zinc ion. Microbiology **148**:2449–2456.
60. **Lewis LA, et al.** 2010. The meningococcal vaccine candidate neisserial surface protein A (NspA) binds to factor H and enhances meningococcal resistance to complement. PLoS Pathog. **6**:e1001027. doi:10.1371/journal.ppat.1001027.
61. **Massé E, Escorcia FE, Gottesman S.** 2003. Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. Genes Dev. **17**:2374–2383.
62. **Massé E, Gottesman S.** 2002. A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. **99**:4620–4625.
63. **Massé E, Majdalani N, Gottesman S.** 2003. Regulatory roles for small RNAs in bacteria. Curr. Opin. Microbiol. **6**:120–124.
64. **Massé E, Vanderpool CK, Gottesman S.** 2005. Effect of RyhB small RNA on global iron use in Escherichia coli. J. Bacteriol. **187**:6962–6971.
65. **Mellin JR, Goswami S, Grogan S, Tjaden B, Genco CA.** 2007. A novel fur- and iron-regulated small RNA, NrrF, is required for indirect fur-mediated regulation of the sdhA and sdhC genes in Neisseria meningitidis. J. Bacteriol. **189**:3686–3694.
66. **Mellin JR, et al.** 2010. Role of Hfq in iron-dependent and -independent gene regulation in Neisseria meningitidis. Microbiology **156**:2316–2326.
67. **Mercante AD, et al.** 2012. MpeR regulates the mtr efflux locus in Neisseria gonorrhoeae and modulates antimicrobial resistance by an iron-responsive mechanism. Antimicrob. Agents Chemother. **56**:1491–1501.
68. **Merrell DS, et al.** 2003. Growth phase-dependent response of Helicobacter pylori to iron starvation. Infect. Immun. **71**:6510–6525.
69. **Mey AR, Wyckoff EE, Kanukurthy V, Fisher CR, Payne SM.** 2005. Iron and fur regulation in Vibrio cholerae and the role of fur in virulence. Infect. Immun. **73**:8167–8178.
70. **Mickelsen PA, Blackman E, Sparling PF.** 1982. Ability of Neisseria

gonorrhoeae, Neisseria meningitidis, and commensal Neisseria species to obtain iron from lactoferrin. Infect. Immun. **35**:915–920.
71. **Mickelsen PA, Sparling PF.** 1981. Ability of Neisseria gonorrhoeae, Neisseria meningitidis, and commensal Neisseria species to obtain iron from transferrin and iron compounds. Infect. Immun. **33**:555–564.
72. **Miles S, Carpenter BM, Gancz H, Merrell DS.** 2010. Helicobacter pylori apo-Fur regulation appears unconserved across species. J. Microbiol. **48**: 378–386.
73. **Miles S, et al.** 2010. Detailed in vivo analysis of the role of Helicobacter pylori Fur in colonization and disease. Infect. Immun. **78**:3073–3082.
74. **Mills SA, Marletta MA.** 2005. Metal binding characteristics and role of iron oxidation in the ferric uptake regulator from Escherichia coli. Biochemistry **44**:13553–13559.
75. **Muenzner P, Bachmann V, Zimmermann W, Hentschel J, Hauck CR.** 2010. Human-restricted bacterial pathogens block shedding of epithelial cells by stimulating integrin activation. Science **329**:1197–1201.
76. **Muenzner P, Rohde M, Kneitz S, Hauck CR.** 2005. CEACAM engagement by human pathogens enhances cell adhesion and counteracts bacteria-induced detachment of epithelial cells. J. Cell Biol. **170**:825–836.
77. **Nandal A, et al.** 2010. Induction of the ferritin gene (ftnA) of Escherichia coli by Fe(2+)-Fur is mediated by reversal of H-NS silencing and is RyhB independent. Mol. Microbiol. **75**:637–657.
78. **Ochsner UA, Vasil ML.** 1996. Gene repression by the ferric uptake regulator in Pseudomonas aeruginosa: cycle selection of iron-regulated genes. Proc. Natl. Acad. Sci. U. S. A. **93**:4409–4414.
79. **Osicka R, Kalmusova J, Krizova P, Sebo P.** 2001. Neisseria meningitidis RTX protein FrpC induces high levels of serum antibodies during invasive disease: polymorphism of frpC alleles and purification of recombinant FrpC. Infect. Immun. **69**:5509–5519.
80. **Pich OQ, Carpenter BM, Gilbreath JJ, Merrell DS.** 2012. Detailed analysis of Helicobacter pylori Fur-regulated promoters reveals a Fur box core sequence and novel Fur-regulated genes. Mol. Microbiol. **84**:921–941.
81. **Pohl E, et al.** 2003. Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. Mol. Microbiol. **47**:903–915.
82. **Sadarangani M, Pollard AJ, Gray-Owen SD.** 2011. Opa proteins and CEACAMs: pathways of immune engagement for pathogenic Neisseria. FEMS Microbiol. Rev. **35**:498–514.
83. **Saito T, Duly D, Williams RJ.** 1991. The histidines of the iron-uptake regulation protein, Fur. Eur. J. Biochem. **197**:39–42.
84. **Schäffer S, Hantke K, Braun V.** 1985. Nucleotide sequence of the iron regulatory gene fur. Mol. Gen. Genet. **200**:110–113.
85. **Schielke S, Frosch M, Kurzai O.** 2010. Virulence determinants involved in differential host niche adaptation of Neisseria meningitidis and Neisseria gonorrhoeae. Med. Microbiol. Immunol. **199**:185–196.
86. **Sebastian S, Agarwal S, Murphy JR, Genco CA.** 2002. The gonococcal fur regulon: identification of additional genes involved in major catabolic, recombination, and secretory pathways. J. Bacteriol. **184**:3965–3974.
87. **Seib KL, Tseng HJ, McEwan AG, Apicella MA, Jennings MP.** 2004. Defenses against oxidative stress in Neisseria gonorrhoeae and Neisseria meningitidis: distinctive systems for different lifestyles. J. Infect. Dis. **190**:136–147.
88. **Senior BW, Stewart WW, Galloway C, Kerr MA.** 2001. Cleavage of the hormone human chorionic gonadotropin, by the Type 1 IgA1 protease of Neisseria gonorrhoeae, and its implications. J. Infect. Dis. **184**:922–925.
89. **Shafer WM, et al.** 2001. Genetic organization and regulation of antimicrobial efflux systems possessed by Neisseria gonorrhoeae and Neisseria meningitidis. J. Mol. Microbiol. Biotechnol. **3**:219–224.
90. **Shaik YB, et al.** 2007. Expression of the iron-activated nspA and secY genes in Neisseria meningitidis group B by Fur-dependent and -independent mechanisms. J. Bacteriol. **189**:663–669.
91. **Sheikh MA, Taylor GL.** 2009. Crystal structure of the Vibrio cholerae ferric uptake regulator (Fur) reveals insights into metal co-ordination. Mol. Microbiol. **72**:1208–1220.
92. **Stohl EA, Criss AK, Seifert HS.** 2005. The transcriptome response of Neisseria gonorrhoeae to hydrogen peroxide reveals genes with previously uncharacterized roles in oxidative damage protection. Mol. Microbiol. **58**:520–532.
93. **Swanson J, Barrera O, Sola J, Boslego J.** 1988. Expression of outer membrane protein II by gonococci in experimental gonorrhea. J. Exp. Med. **168**:2121–2129.
94. **Troxell B, Fink RC, Porwollik S, McClelland M, Hassan HM.** 2011. The Fur regulon in anaerobically grown Salmonella enterica sv. Typhimurium: identification of new Fur targets. BMC Microbiol. **11**:236. doi: 10.1186/1471-2180-11-236.
95. **Troxell B, et al.** 2011. Fur negatively regulates hns and is required for the expression of HilA and virulence in Salmonella enterica serovar Typhimurium. J. Bacteriol. **193**:497–505.
96. **Tsolis RM, Baumler AJ, Stojiljkovic I, Heffron F.** 1995. Fur regulon of Salmonella typhimurium: identification of new iron-regulated genes. J. Bacteriol. **177**:4628–4637.
97. **Urban JH, Vogel J.** 2007. Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res. **35**:1018–1037.
98. **van Vliet AH, et al.** 2003. Differential regulation of amidase- and formamidase-mediated ammonia production by the Helicobacter pylori fur repressor. J. Biol. Chem. **278**:9052–9057.
99. **Vassinova N, Kozyrev D.** 2000. A method for direct cloning of fur-regulated genes: identification of seven new fur-regulated loci in Escherichia coli. Microbiology **146**(Pt 12):3171–3182.
100. **Waters LS, Storz G.** 2009. Regulatory RNAs in bacteria. Cell **136**:615–628.
101. **Whitby PW, Seale TW, VanWagoner TM, Morton DJ, Stull TL.** 2009. The iron/heme regulated genes of Haemophilus influenzae: comparative transcriptional profiling as a tool to define the species core modulon. BMC Genomics **10**:6. doi:10.1186/1471-2164-10-6.
102. **Whitby PW, Vanwagoner TM, Seale TW, Morton DJ, Stull TL.** 2006. Transcriptional profile of Haemophilus influenzae: effects of iron and heme. J. Bacteriol. **188**:5640–5645.
103. **Wilderman PJ, et al.** 2004. Identification of tandem duplicate regulatory small RNAs in Pseudomonas aeruginosa involved in iron homeostasis. Proc. Natl. Acad. Sci. U. S. A. **101**:9792–9797.
104. **Xiong A, Singh VK, Cabrera G, Jayaswal RK.** 2000. Molecular characterization of the ferric-uptake regulator, fur, from Staphylococcus aureus. Microbiology **146**(Pt 3):659–668.
105. **Yu C, Genco CA.** 2012. Fur-mediated activation of gene transcription in the human pathogen Neisseria gonorrhoeae. J. Bacteriol. **194**:1730–1742.
106. **Zhao S, Montanez GE, Kumar P, Sannigrahi S, Tzeng YL.** 2010. Regulatory role of the MisR/S two-component system in hemoglobin utilization in Neisseria meningitidis. Infect. Immun. **78**:1109–1122.
107. **Zhou D, et al.** 2006. Global analysis of iron assimilation and fur regulation in Yersinia pestis. FEMS Microbiol. Lett. **258**:9–17.

JB Journals.ASM.org

MINIREVIEW

Peroxide-Sensing Transcriptional Regulators in Bacteria

James M. Dubbs[a] and Skorn Mongkolsuk[a,b,c,d]

Laboratory of Biotechnology, Chulabhorn Research Institute, Lak Si, Bangkok, Thailand[a]; Center of Emerging Bacterial Infections[b] and Department of Biotechnology,[c] Faculty of Science, Mahidol University, Bangkok, Thailand; and Center of Excellence for Environmental Health, Toxicology, and Management of Chemicals, Bangkok, Thailand[d]

The ability to maintain intracellular concentrations of toxic reactive oxygen species (ROS) within safe limits is essential for all aerobic life forms. In bacteria, as well as other organisms, ROS are produced during the normal course of aerobic metabolism, necessitating the constitutive expression of ROS scavenging systems. However, bacteria can also experience transient high-level exposure to ROS derived either from external sources, such as the host defense response, or as a secondary effect of other seemingly unrelated environmental stresses. Consequently, transcriptional regulators have evolved to sense the levels of ROS and coordinate the appropriate oxidative stress response. Three well-studied examples of these are the peroxide responsive regulators OxyR, PerR, and OhrR. OxyR and PerR are sensors of primarily H_2O_2, while OhrR senses organic peroxide (ROOH) and sodium hypochlorite (NaOCl). OxyR and OhrR sense oxidants by means of the reversible oxidation of specific cysteine residues. In contrast, PerR senses H_2O_2 via the Fe-catalyzed oxidation of histidine residues. These transcription regulators also influence complex biological phenomena, such as biofilm formation, the evasion of host immune responses, and antibiotic resistance via the direct regulation of specific proteins.

An effective oxidative stress defense response is a required item in the basic survival kit of all aerobic organisms as well as those anaerobes that exist in environments subject to transient exposures to oxygen. This is due to molecular oxygen's ability to accept electrons from cellular redox components to form toxic reactive oxygen species (ROS) (51). One such product is superoxide anion ($O_2^{\bullet -}$), which results from a one-electron reduction of O_2. Another, resulting from a two-electron reduction, is hydrogen peroxide (H_2O_2) (51). Both $O_2^{\bullet -}$ and H_2O_2 have the ability to oxidize exposed iron sulfur clusters, as has been observed for certain dehydratases (50), while H_2O_2 is known to oxidize Fe^{2+} in proteins containing mononuclear iron centers (100). This not only inactivates the enzymes but also results in the release of free ferric iron (Fe^{3+}), which is converted to ferrous iron (Fe^{2+}) in the intracellular reducing environment (51). This is significant since Fe^{2+} can reduce H_2O_2 to form hydroxyl radicals ($^{\bullet}OH$) in the Fenton reaction as follows: $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + {}^{\bullet}OH + OH^-$.

Hydroxyl radicals are highly potent oxidants of cellular macromolecules that react at a diffusion-limited rate (51). Thus, the increased intracellular iron levels resulting from oxidative damage to a subset of iron-containing proteins can lead to increased $^{\bullet}OH$-mediated cellular damage, particularly to DNA (55). $O_2^{\bullet -}$, H_2O_2, and $^{\bullet}OH$ are all capable of causing a variety of oxidative lesions in proteins, DNA, and lipids (51). An important aspect of some of these secondary reactions is the production of a variety of organic peroxides that can also mediate further oxidative damage (97).

In addition to endogenous sources, bacteria also encounter ROS from external sources. Macrophages actively produce large amounts of $O_2^{\bullet -}$ and nitric oxide ($NO^{\bullet}$) in order to kill invading pathogens (92). The host defense response in plants is similar and also includes the synthesis of organic peroxides (62), while some microorganisms excrete ROS to inhibit the growth of competitors (42). Exposure to redox cycling compounds, such as the herbicide paraquat, or naturally occurring quinones can also serve as a source of ROS. These molecules can participate in cyclic reactions in which they transfer an electron from cellular electron donors, such as NADH oxidases, to O_2 to generate $O_2^{\bullet -}$ (22).

Bacteria maintain a basal level of protective enzymes to detoxify $O_2^{\bullet -}$ and H_2O_2, thereby keeping their concentrations within safe limits (95). These include superoxide dismutases that use a metal center to catalyze the dismutation of $O_2^{\bullet -}$ to H_2O_2 and catalases that usually employ a heme cofactor to convert H_2O_2 to O_2 and H_2O (50). Peroxiredoxins, a class of peroxidases, function to reduce H_2O_2 and organic peroxides (ROOH) to either H_2O or H_2O and the corresponding alcohol (ROH) via the peroxide-mediated oxidation of cysteine thiols that are subsequently reduced using cellular electron donors to regenerate the active enzyme (88, 94).

When ROS levels exceed safe limits, bacteria have the ability to mount an inducible response, resulting in increased expression of ROS detoxification enzymes along with additional protective systems that repair oxidative damage, protect vulnerable enzymes from inactivation, and control the levels of free Fe^{2+} (87). For example, oxidative stress increases the demand for reducing equivalents, necessitating the induction of metabolic pathways to increase the reductant supply (32). The expression of iron-binding proteins, like Dps, is often induced in response to oxidative stress to reduce the levels of free Fe in order to prevent production of $^{\bullet}OH$ via the Fenton reaction (12). Strategies are also employed to protect enzymes that contain vulnerable Fe-containing centers from oxidation by increasing the uptake of Mn^{2+}, which is able to replace Fe^{2+} at some active sites, rendering them resistant to oxidation (50). In some cases, Fe-containing enzymes are replaced with analogs that do not require Fe (70).

The regulation of the expression of genes involved in the bacterial oxidative stress defense response is complex and often under the control of regulators that can directly sense the levels of spe-

Published ahead of print 13 July 2012

Address correspondence to James M. Dubbs, djames@cri.or.th.

doi:10.1128/JB.00304-12

cific ROS and activate or derepress target gene transcription. This review summarizes work concerning the mechanisms of oxidant sensing and transcriptional regulation by OxyR, PerR, and OhrR, three well-studied peroxide responsive regulators that have the ability to distinguish between ROOH and H_2O_2 (75). Two distinct mechanisms for peroxide sensing that utilize either the oxidation of cysteine residues (OxyR and OhrR) or the metal-catalyzed oxidation of histidine residues (PerR) have evolved in these regulators. In each case, oxidative modification of the regulator alters its DNA binding properties.

OxyR

OxyR is an H_2O_2-sensing transcriptional regulator of the LysR family (93) that is generally found in Gram-negative bacteria but is also known to occur in a few Gram-positive bacteria (78, 83). Like other regulators of this family, it contains a conserved N-terminal helix-turn-helix DNA binding domain, a central coinducer recognition and response domain which senses the regulatory signal, and a C-terminal domain that functions in multimerization and activation (59–61, 93, 115). OxyR functions primarily as a global regulator of the peroxide stress response that maintains intracellular H_2O_2 levels within safe limits (1) and also plays a role in the cellular response to thiol depletion (2). The OxyR regulon of *Escherichia coli* is comprised of over 20 genes, including genes involved in H_2O_2 detoxification (*katE*, *ahpCF*), heme biosynthesis (*hemH*), reductant supply (*grxA*, *gor*, *trxC*), thiol-disulfide isomerization (*dsbG*), Fe-S center repair (*sufA-E*, *sufS*), iron binding (*yaaA*), repression of iron import systems (*fur*), and manganese import (71, 124, 125). OxyR also upregulates the expression of OxyS, a small regulatory RNA that integrates peroxide stress with general stress pathways (34, 122). Although there are significant differences, the OxyR regulons of other organisms tend to include similar classes of genes (39, 82, 96, 104, 116).

While OxyR is primarily thought of as a transcriptional activator under oxidizing conditions that acts through direct interaction with the RNA polymerase α subunit (61, 103, 106), OxyR can function as either a repressor or activator under both oxidizing and reducing conditions (19, 45, 49, 96, 108, 121).

In *E. coli*, tetrameric OxyR binds to the 5′ promoter-operator regions of target genes at a conserved sequence motif (61, 108, 123) (Fig. 1). The oxidized ($OxyR^{ox}$) and reduced ($OxyR^{red}$) forms of the protein adopt different conformational states that can sometimes result in changes in the DNA binding contacts, with the reduced form of OxyR contacting two pairs of major grooves separated by one helical turn while the DNA binding contacts of the oxidized form shift to contact four consecutive major grooves (108). These redox-dependent changes also alter its affinity for target promoters and can affect promoter conformation and, thereby, contacts to RNA polymerase (103, 107, 108).

Sensing of H_2O_2 occurs via direct oxidation of OxyR at a specific "sensing" cysteine residue (Fig. 1). In *E. coli*, in which the normal intracellular H_2O_2 level is ~20 nM, OxyR is present in its reduced form (95, 123). Rapid oxidation of $OxyR^{red}$ occurs when intracellular H_2O_2 levels reach ~100 nM, which is well below the level at which growth is inhibited (~2 μM) (33, 95). $OxyR^{ox}$ is slowly reduced via glutaredoxin 1 using electrons supplied by reduced glutathione (2, 105, 123) (Fig. 1). This provides feedback regulation of the system, since expression of *grxA* (glutaredoxin 1) and *gor* (glutathione reductase) is induced during oxidative stress as part of the OxyR regulon (105, 123, 125). This dependence on

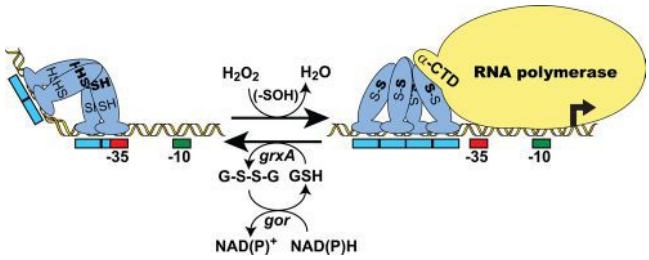


FIG 1 Mechanism of transcription activation via intramolecular disulfide formation in 2-Cys OxyR. A model for H_2O_2-dependent, OxyR-mediated transcription activation of a target gene in *E. coli* is shown. Activation begins with the oxidation of the sensing cysteine (SH) residue of OxyR to sulfenic acid (-SOH), followed by the rapid formation of an intramolecular disulfide bond with the resolving cysteine (SH). The resulting conformational change often causes a shift in the DNase I footprint and can also affect DNA binding affinity and promoter conformation as well as render OxyR capable of interacting with RNA polymerase. Transcription activation involves the direct interaction of OxyR with the alpha subunit C-terminal domain (α-CTD) of RNA polymerase. Oxidized OxyR is reduced, using reduced glutathione (GSH) as the electron donor, via the glutaredoxin (*grxA*)/glutathione reductase (*gor*) system, with reducing equivalents ultimately supplied by NAD(P)H. Red and green boxes indicate RNA polymerase σ^{70} −35 and −10 promoter elements, respectively. Blue boxes indicate OxyR DNA binding contacts. Activation can also occur via oxidative modification of the sensing cysteine alone.

reduction by GrxA also renders OxyR sensitive to the thiol-disulfide redox status (glutathione disulfide [GSSG]/reduced glutathione [GSH] ratio) of the cell (2, 123). Recent evidence indicates that a similar reductive recycling of OxyR occurs in *Pseudomonas aeruginosa*, which utilizes a thioredoxin/thioredoxin reductase system that is part of the OxyR regulon (116).

The exact mechanism(s) of oxidant sensing by OxyR is still a subject of active investigation and has generated some debate (44, 84). Two basic models have been proposed. The first, supported by clear evidence, indicates that H_2O_2-mediated activation of *E. coli* OxyR constitutes a simple on/off switch that occurs through the formation of a specific disulfide bond between the conserved cysteine residues C199 and C208 (65, 105, 123). Activation involves a two-step oxidation that begins when H_2O_2 reacts with a thiolate ion of the sensing cysteine, C199, to form a sulfenic acid (C199-SOH) (Fig. 1). C199-SOH rapidly reacts with C208-SH to form an intramolecular disulfide bond (65). This induces structural changes in the regulatory domain (amino acids 80 to 305) (18, 65) that result in altered associations between the subunits within the tetramer, leading to altered DNA binding properties and allowing productive interaction between OxyR and RNA polymerase (18, 106, 108).

The second "molecular code" hypothesis suggests that modification of C199 alone is sufficient to activate OxyR in the absence of disulfide bond formation and that the type of modification determines the regulatory outcome (56). In this case, a more selective regulatory response in which distinct subsets of OxyR target genes are differentially regulated depending on the particular oxidized form of OxyR involved is expected. Initially, this was based primarily on *in vitro* studies indicating that oxidative modification of only C199 was sufficient for activation, since hydroxylated Cys199-SOH, *S*-nitrosylated (Cys199-SNO), and *S*-glutathionylated (Cys199-S-S-G) forms of OxyR were able to activate transcription, and that the different modifications elicited distinct changes in the circular dichroism (CD) spectra of OxyR-DNA

complexes (40, 56). The idea was subsequently reinforced by the observation that Cys199-thiol-esterification of OxyR resulted in activation of target genes both *in vitro* and *in vivo* (38). Until recently, there were little available data to adequately evaluate the molecular code model's prediction of modification-specific patterns of gene activation. In fact, microarray studies had indicated that OxyR did not play a significant role in gene activation in response to the *S*-nitrosylating agents nitrosylated glutathione (GSNO) and sodium nitrite under aerobic conditions (80). However, a recent analysis in *E. coli* has shown that *S*-nitrosylation of proteins occurs naturally during anaerobic respiratory growth on nitrate, resulting in the nitrosylation of OxyR at C199 that causes the activation of a set of genes that is distinct from those upregulated in response to H_2O_2 (98). Thus, it appears that in the case of *E. coli* OxyR, both intramolecular disulfide bond formation and modifications to Cys199 alone can elicit distinct OxyR-mediated regulatory outcomes.

Several recent studies also indicate that there is diversity in the activation mechanisms of OxyR orthologs in other organisms. For example, *in vivo* studies of the *E. coli*-like OxyR from *Pseudomonas aeruginosa*, involving free thiol labeling of OxyR in extracts of wild-type and C199/C208 single and double mutants exposed to H_2O_2, detected several oxidized forms of OxyR in addition to the one containing a disulfide bond between C199 and C208 (45). The resulting suggestion that *Pseudomonas* OxyR has a different redox cycle than the *E. coli* protein has been supported by a recent mutagenesis study indicating that a third cysteine residue (Cys296) may be involved in peroxide sensing in *P. aeruginosa* (4). This cysteine is not conserved in the *E. coli* protein and has so far been found only in a small group within the betaproteobacteria as well as in several *Pseudomonas* species (4).

Finally, a second OxyR structural class, containing only one cysteine, has been identified in *Deinococcus radiodurans* (13, 120). Two novel OxyR proteins (DrOxyR and DrOxyR$_2$) are found in this organism, with each containing one essential sensing cysteine residue that roughly corresponds to C208 in the *E. coli* protein. DrOxyR and DrOxyR$_2$ are 31% and 28% identical to *E. coli* OxyR, respectively, and contain conserved residues that have been implicated in DNA binding (R4, L32 S33, R50), activation (D142, R273), and multimerization (A233) (60, 115). However, residues in the *E. coli* protein that are predicted to play a role in sulfenic acid formation at C199 and subsequent disulfide bond formation with C208 (H198, R201, R266, T238) (59, 61) are absent. The differences in the sensing mechanisms between the 1-Cys and 2-Cys OxyR proteins have yet to be explored. However, DrOxyR is able to complement an *E. coli* OxyR mutant, thus indicating functional similarity, while *in vitro* chemical modification studies have shown that H_2O_2-dependent activation of DrOxyR involves the formation of a cysteine sulfenic acid at the essential sensing cysteine (13).

PerR

PerR functions as a peroxide responsive repressor and is a member of the Fur family of small, dimeric, metal-responsive transcriptional regulators (1). PerR is a global regulator that responds primarily to H_2O_2. It substitutes for OxyR in many Gram-positive bacteria (75). However, its distribution is not limited to Gram-positive bacteria (69, 78, 91, 112), and in several cases, PerR is found along with OxyR (36, 37, 111, 118). In *Bacillus subtilis*, PerR regulates itself (30) and other genes, including those involved in oxidative stress defense (*katA*, *ahpCF*, *mrgA*) (10, 14), metal homeostasis (*zosA*, *fur*, *hemAXCDBL*) (14, 30, 31), and surfactant production (*srfA*) (41), in response to peroxide exposure. As with OxyR, PerR homologs in other organisms tend to regulate similar classes of genes (7, 8, 36, 48, 79, 91, 117).

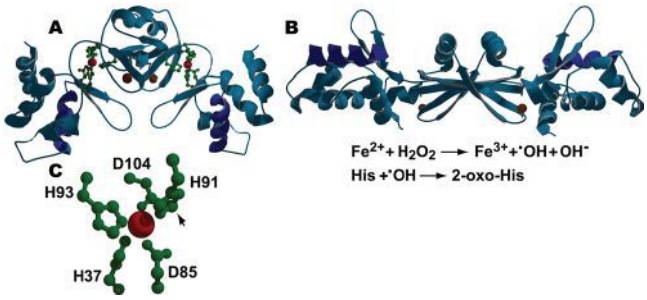


FIG 2 H_2O_2-mediated inactivation of PerR. (A) The ribbon structure of the carbon backbone of reduced *B. subtilis* PerR-Zn^{2+}-Mn^{2+}, showing the side chains of the amino acids coordinating the regulatory metal (green) and the DNA binding helices (dark blue), is shown. The structural and regulatory metals, Zn^{2+} and Mn^{2+}, are shown as brown and red spheres, respectively. In its reduced state, PerR-Zn^{2+}-Fe^{2+}/Mn^{2+} binds to sites overlapping target promoter/operators and blocks transcription. (B) The ribbon structure of oxidized PerR-Zn^{2+}-Mn^{2+} is shown. The steps in the Fenton-mediated oxidation of His residues in PerR-Zn^{2+}-Fe^{2+} by H_2O_2 are shown below the panel. Exposure to H_2O_2 results in the iron-catalyzed production of $OH^{\bullet}$, followed by the oxidation of the 2-carbon of the imidazole ring of one of two histidine residues (H37 and H91) that participate in coordinating the bound Fe^{2+} in each monomer. While H37 is the preferential target for oxidation, oxidation of either histidine to 2-oxo-histidine results in the disruption of normal Fe coordination, resulting in the destabilization of the DNA binding domain of the monomer and leading to the release of oxidized PerR from the DNA. (C) Close-up view of the Mn^{2+}-containing regulatory metal binding site in reduced PerR that shows the relative positions of the metal-coordinating side groups of H37, H91, H93, D85, and D104 (positions of the coordinating side groups roughly correspond to those on the right side of panel A). The arrow indicates the 2-carbon of the imidazole group of H91, which is the site of oxidation in H91 and H37. In all cases, DNA binding helices are shown in dark blue. The Swiss Protein Data Bank (PDB) identifications (IDs) for reduced and oxidized PerR are 3F8N and 2RGV, respectively.

Each PerR monomer contains a binding site for a structural Zn^{2+} as well as a regulatory site that, in *B. subtilis*, binds either Fe^{2+} or Mn^{2+} (46, 52) (Fig. 2A). Repression is achieved when PerR, containing metals bound at both sites, binds DNA at a consensus Per-box (29) that either overlaps part of the promoter or is immediately downstream from it (36, 46). PerR-mediated positive regulation has also been observed in at least two cases and appears to involve PerR binding to distant upstream sites; however, the mechanism of activation has not been investigated (7, 41).

While coordination of either Fe^{2+} or Mn^{2+} at the regulatory site is required for DNA binding of the PerR dimer (46), the Fe^{2+}-containing form (PerR-Zn^{2+}-Fe^{2+}) is thought to be responsible for H_2O_2 sensing *in vivo*. This is based on the observations that PerR-Zn^{2+}-Fe^{2+} is $>10^3$-fold more sensitive to oxidation than PerR-Zn^{2+}-Mn^{2+} (67) and that Fe^{2+} has a higher affinity (~30-fold) for binding to the regulatory site than Mn^{2+}. The exact physiological relevance of Mn^{2+} binding at the regulatory site is not fully understood, but Mn^{2+} can compete with Fe^{2+} for binding to PerR *in vivo*, thus altering the relative sensitivity of genes in the PerR regulon to H_2O_2 induction (14, 26, 66). Not only does Mn compete with Fe, but in some cases, Mn, not Fe, appears to func-

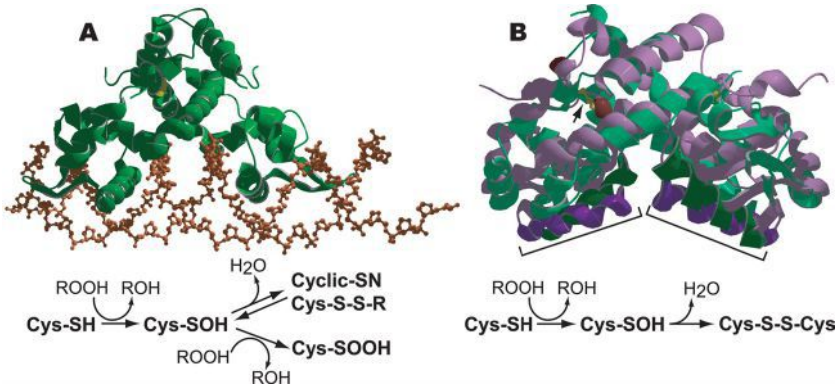


FIG 3 Organic hydroperoxide (ROOH)-mediated derepression of OhrR-regulated promoters. (A) The structure of reduced 1-Cys *B. subtilis* OhrR (green ribbon) bound to the *ohrA* operator (DNA sugar phosphate backbone in brown) is shown. The redox cycle of *B. subtilis* OhrR is depicted below. Reduced 1-Cys OhrR binds to the target promoter/operator through the interaction of winged helix-turn-helix DNA binding domains with the DNA major groove, thereby blocking transcription. In the presence of organic hydroperoxide (ROOH), the single sensing cysteine in reduced OhrR, C15 (yellow), is oxidized to cysteine sulfenic acid (Cys-SOH). The Cys-SOH derivative remains bound to the promoter and must undergo one of several further modifications to induce the conformational change necessary to release the repressor. Cys-SOH can react either with a reduced cellular thiol to form a mixed thiol (Cys-S-S-R) or with the amino group of a neighboring amino acid residue to form a cyclic amide (Cys-SN). Both Cys-S-S-R and Cys-SN are recycled *in vivo* via reduction. Derepression can also occur through further oxidation of the Cys-SOH, for example, to cysteine sulfinic acid (Cys-SOOH). Overoxidized derivatives of OhrR are likely degraded. (B) An overlay of the ribbon structures of the reduced (purple) and oxidized (green) forms of the 2-Cys *X. campestris* OhrR showing the peroxide-induced shift in the positions of the DNA binding helices (dark purple and dark green, indicated by brackets). The redox-active cysteines in reduced OhrR are colored brown, while the disulfide bonds in the oxidized form are rendered in yellow (also indicated by an arrow in the right-hand monomer). The redox cycle of a 2-Cys OhrR is depicted below. The sensing cysteine (C22) of reduced 2-Cys OhrR bound to a target promoter/operator is oxidized in the presence of ROOH to cysteine sulfenic acid (Cys-SOH). The sensing Cys-SOH rapidly reacts with a second "resolving" cysteine (C127) residue to form an intersubunit disulfide bond that induces a conformational change that repositions the DNA binding helices and releases OhrR from the promoter. The oxidized disulfide bond-containing form of 2-Cys OhrR is likely actively rereduced *in vivo*. The PDB IDs for reduced *B. subtilis* OhrR proteins bound to the *ohrA* operator and oxidized and reduced *X. campestris* OhrR proteins are 1Z9C, 2PFB, and 2PEX, respectively.

tion as the corepressor (30). Mn^{2+} may also serve to protect PerR from oxidative inactivation while still allowing PerR-mediated derepression of target genes in response to Fe^{2+} released due to oxidative damage (66).

The oxidation-dependent disruption of proper iron coordination at the regulatory site is the basis of oxidant sensing by PerR (Fig. 2). Accumulated evidence indicates that Fe^{2+} (or Mn^{2+}) bound at the regulatory site is coordinated by three histidine (H37, H91, H93) and two aspartate (D104, D85) residues. These constitute a four-sided pyramid with alternating histidine (H37 and H91) and aspartate residues at the corners of the base and a His (H93) positioned at the apex (52) (Fig. 2C). Binding of the regulatory metal locks the dimer in a caliper-like configuration that stabilizes the positions of the N-terminal DNA binding domains of the monomers such that they can interact with the DNA (52) (compare Fig. 2A and B). H_2O_2-mediated inactivation of PerR involves the reduction of H_2O_2 by Fe^{2+} at the regulatory site to form the hydroxyl radical ($^{\bullet}OH$) (Fig. 2C). This highly reactive species oxidizes H37 and, to a lesser degree, H91 to 2-oxo-histidine, resulting in the disruption of normal metal coordination (67, 110). This causes an opening of the caliper configuration that negates DNA binding, a result which is seen in the crystal structures of demetalated, nonoxidized apo-PerR-Zn^{2+} as well as oxidized PerR-Zn^{2+}-Fe^{2+} (52, 109, 110) (Fig. 2B). The region that allows access to Fe^{2+} bound at the regulatory site is rich in hydrophilic residues, thus explaining PerR's preference for H_2O_2 (52). The fact that there is no known physiological system to repair 2-oxo-histidine has led to the proposal that oxidized PerR is probably not recycled (67).

A second distinct mechanism of PerR inactivation occurs via exposure to nitrosylating agents such as nitrous oxide (NO). Exposure of *B. subtilis* to either an NO bolus or the NO generator Na-nitroprusside results in PerR-mediated induction of the PerR regulon during growth under aerobic and anaerobic conditions (77). The observation that anaerobic growth in the presence of Mn^{2+} reduces NO-induced expression of the PerR regulon while high iron levels enhance it suggests that induction is mediated by PerR-Zn^{2+}-Fe^{2+}. And it has been proposed that NO-mediated inactivation of PerR may be due to nitrosylation of bound Fe^{2+} (77).

The structural site that binds Zn^{2+} does not appear to play a role in oxidant sensing. The structural Zn^{2+} is coordinated by the cysteines C96, C99, C136, and C139 (67). In both PerR-Zn^{2+}-Fe^{2+} and PerR-Zn^{2+}-Mn^{2+}, these cysteines have been shown to be highly resistant to oxidation with either H_2O_2 or diamide at levels higher than would be encountered *in vivo* (109).

OhrR

OhrR is a transcriptional repressor belonging to the MarR family of regulators and senses both organic peroxides and NaOCl (17). Generally, in their reduced form, OhrR dimers bind cooperatively to conserved, AT-rich, inverted-repeat sequences that overlap target gene promoters via winged-helix DNA binding domains (20, 28, 47, 76, 83) (Fig. 3A). OhrR is found in both Gram-positive and Gram-negative bacteria (1, 86) and can coreside with either OxyR (85) or PerR (27) or both (83). The first examples of OhrR to be characterized were shown to regulate themselves along with a gene encoding an organic peroxide-specific peroxiredoxin, Ohr (20, 27, 76, 83). Other orthologs that regulate a larger and more diverse set of target genes have since been characterized (16, 63, 72).

Organic peroxide- or sodium hypochlorite (NaOCl)-mediated inactivation of OhrR, leading to derepression of OhrR-repressed

genes, occurs via the oxidation of a sensing cysteine that undergoes subsequent disulfide formation (either intersubunit or mixed). This usually renders OhrR incapable of binding DNA, leading to derepression of target genes (1, 28, 85, 86). One exception to this general scheme occurs in *Streptomyces coelicolor*, in which reduced OhrR binds cooperatively to multiple sites overlapping the divergently transcribed promoters of *ohrR* and *ohrA*. Upon oxidation, $OhrR^{ox}$ remains bound to a single high-affinity binding site within the *ohrR-ohrA* intergenic region that is upstream of the *ohrR* promoter and overlaps the *ohrA* −35 sequence. Interestingly, $OhrR^{ox}$ bound at this site activates *ohrR* transcription but does not interfere with the transcription of *ohrA* (83).

The preferred organic peroxide inducer for OhrR can vary between organisms. For example, OhrR proteins of *B. subtilis* and *Xanthomonas campestris* are more sensitive to complex organic peroxides, such as linoleic acid hydroperoxide, while *Agrobacterium tumefaciens* OhrR preferentially senses less-complex organic peroxides like cumene hydroperoxide (57, 85, 102).

OhrR is now recognized to be a member of an oxidant-responsive subgroup of regulators of the MarR family that share a peroxide-sensing mechanism. Other members of the group include *Pseudomonas aeruginosa* OspR (63) and MgrA and SarZ in *Staphylococcus aureus* (54, 73). All sense peroxides or other oxidants through the initial oxidation of a sensing cysteine residue (15, 16, 63). However, OspR, MgrA, and SarZ differ from OhrR in that each appears to play a much more global regulatory role.

Like OhrR, *Pseudomonas aeruginosa* OspR appears to preferentially sense organic peroxides and mediates peroxide resistance through the regulation of a glutathione peroxidase gene (*PA2826*) while also directly regulating genes involved in quorum sensing and tyrosine metabolism. Alterations in OspR activity also affect β-lactam resistance, pigment production, and virulence in a murine acute pneumonia model (63).

MgrA and SarZ are global regulators in *S. aureus* that use peroxide as a signal for an adaptive shift to growth inside a host. MgrA is known to regulate, directly or indirectly, over 340 genes, including those encoding a variety of virulence factors as well as several antibiotic resistance determinants (72). SarZ, a close homolog of MgrA, also acts as a global regulator, controlling at least 80 genes involved in virulence, peroxide and antibiotic resistance, and the metabolic shift to anaerobic growth (16).

OhrR and its orthologs fall into one of two structural classes depending on the number of cysteines that participate in peroxide sensing (Fig. 3). Members of the first class, known as the 1-Cys class, which also includes OspR, MgrA, and SarZ, contain a single N-terminal sensing cysteine (C15 in *B. subtilis* OhrR) that is conserved in all OhrR proteins (47, 86). A second 2-Cys class contains an additional cysteine residue (C127 in *X. campestris* OhrR) in the C-terminal region that is also involved in peroxide sensing (86).

The current model for the mechanism of organic peroxide sensing in OhrR/MgrA family members is similar in several respects to peroxide sensing in OxyR. In both 1-Cys and 2-Cys OhrR proteins, the initial step in organic peroxide-mediated derepression involves the oxidation of the N-terminal sensing cysteine to a sulfenic acid (C-SOH) (28, 86) (Fig. 3). Organic peroxide specificity is due to a hydrophobic region lining the access channel to the sensing cysteine (47). This OhrR sulfenic acid derivative is still capable of DNA binding (68). At least one of the oxidized cysteines in the dimer must undergo additional modification in order for derepression to occur (25). This second stage of the sensing mechanism is where the 1-Cys and 2-Cys classes differ. Studies of *B. subtilis* OhrR have determined that inactivation of a 1-Cys OhrR can occur in several ways (Fig. 3A). The C15 sulfenic acid can either rapidly form a mixed disulfide with a cellular free thiol (i.e., bacillithiol, a recently identified 398-Da thiol) or react with the amino group of an adjacent amino acid residue to form a cyclic sulfenamide (17, 68). Mixed disulfide and protein sulfenamide formation is reversible since DNA binding activity of these derivatives of OhrR was able to be recovered *in vitro* by reduction in the presence of dithiothreitol (DTT) (68, 102). Irreversible inactivation of OhrR results when the C15 sulfenic acid is further oxidized to either a sulfinic (C15-SOOH) or sulfonic ($C15\text{-}SO_3H$) acid, a process which occurs either in cells exposed to highly efficient inducers, like linoleic acid hydroperoxide, or in organic peroxide-exposed cells already undergoing disulfide stress (68).

The 2-Cys OhrR class, exemplified by *X. campestris* OhrR, contains an N-terminal sensing cysteine, C22, and two additional C-terminal cysteines, C127 and C131 (86). Only C22 and C127 are essential for regulator function (85, 86). Initial oxidation of the sensing C22 to sulfenic acid is followed by rapid formation of an intermolecular disulfide bond with C127 of the other subunit in the dimer (Fig. 3B). It is the formation of this intermolecular disulfide bond, not the oxidation of C22 to sulfenic acid, which renders the protein incapable of DNA binding, since amino acid substitutions at C127 in OhrR result in a protein that binds DNA in both its reduced and oxidized forms (3, 86). Therefore, disulfide bond formation in 2-Cys OhrR is functionally analogous to mixed disulfide formation in 1-Cys OhrR. Disulfide bond formation or mixed disulfide formation likely also serves to protect OhrR from overoxidation, which permanently inactivates the repressor (3, 68). The presence of a reduction system for oxidized OhrR containing either intermolecular or mixed disulfide bonds has been suggested but has not been identified (68). Possible candidates include systems utilizing dihydrolipoamide which have recently been shown to be involved in the reductive recycling of the organic peroxide-specific peroxiredoxins Ohr and OsmC in *Xylella fastidiosa* (24), while in *B. subtilis* and other *Firmicutes*, the bacilliredoxins, made up of several putative thiol-disulfide reductases (17, 43), are an attractive possibility.

The crystal structures of reduced unbound *X. campestris* OhrR and reduced *B. subtilis* OhrR bound to its operator sequence are similar (Fig. 3A and B), and conserved residues that appear to perform similar functions in both proteins have been identified (47, 81). Oxidation and disulfide bond formation in *X. campestris* OhrR induce dramatic structural changes, resulting in a 28° rotation of the winged helix-turn-helix DNA binding domains to positions that are incompatible with DNA binding (81) (Fig. 3B). No crystal structure is available for a 1-Cys OhrR containing a mixed disulfide; however, some insight into possible structural changes can be gained from the known structures of the reduced and oxidized forms of SarZ. C13-SOH in oxidized SarZ retains several of the normal hydrogen bonding interactions found in the reduced form. Thus, oxidation to Cys sulfenic acid has only a small effect on the positions of the DNA binding helices. This is consistent with observations that the sulfenic acid derivatives of OhrR and SarZ retain DNA binding activity (83, 89). Mixed thiol formation (C13-S-S-R) in SarZ disrupts these interactions, resulting in a conformational shift that increases the spacing between the DNA binding helices and negates DNA binding (89). The structural rearrangement observed in the mixed thiol derivative of SarZ is

less pronounced than that seen in the oxidized disulfide-containing *X. campestris* OhrR, which displays a much greater degree of rotation between the subunits in the dimer (Fig. 3B). It will be interesting to see if the structural changes that occur upon mixed thiol formation in 1-Cys OhrR are similar to those in the dithiol derivative of the 2-Cys class or if they more closely resemble those seen in SarZ. In this regard, it is interesting to note that a 1-Cys *B. subtilis* OhrR can be converted to a functional 2-Cys OhrR by changing either of two Q residues (G120 or Q124), at positions corresponding to C127 and C131 in the *X. campestris* OhrR, to C so that the structural changes leading to derepression may be similar for the 1-Cys and 2-Cys OhrR proteins (101).

ROLE OF OxyR, PerR, AND OhrR IN VIRULENCE

ROS are an important component of the host immune response (62, 92), and it is becoming clear that peroxide is an important signal governing the expression of virulence genes. This is perhaps best illustrated by the peroxide-sensing regulators MgrA and SarZ in *S. aureus*, which have evolved to orchestrate the adaptive shift necessary to promote growth within a host in response to peroxide (5, 53). Although they are viewed primarily as modulators of the oxidative stress response, OxyR (64, 104), PerR (35, 48, 91, 117), and OhrR (3) have all been shown to be required for virulence in at least one experimental system. In the case of OxyR and PerR, their effect on virulence is not solely related to the regulation of oxidative stress defense genes. OxyR induces the expression of virulence factors that allow pathogens to evade host innate immunity (74, 90), while PerR has been shown to directly regulate an extracellular virulence factor (mitogen factor 3) in *Streptococcus pyogenes* (117). Other clinically relevant processes, such as biofilm formation (99, 119) and the oxidation state-independent regulation of several epigenetically controlled promoters determining phase variation (9, 113, 114), are also regulated by OxyR.

CONCLUSION

Much is known concerning the basic structural, mechanistic, and regulatory details of the peroxide-sensing regulators that are the subject of this review; however, many questions remain. Clearly, OxyR and its orthologs have evolved to sense a variety of oxidants in addition to H_2O_2 and have assigned a role for OxyR as a sensor of nitrosative stress during anaerobic growth. There are at least two distinct mechanisms of oxidant sensing, involving oxidative modification of a sensing cysteine alone (13, 98) or in concert with a second disulfide bonding partner (13), that are known in OxyR. There are also indications of the potential involvement of a third cysteine in peroxide sensing (4). The LysR-type regulator BenM, which controls genes involved in benzoate catabolism, is known to sense two coinducers (cis,cis-muconate and benzoate) that bind to distinct sites on the protein (23). Binding of each ligand singly activates transcription, while simultaneous binding of both coinducers results in a synergistic increase in target gene transcription (11) that is not observed for all BenM target promoters (21). It will be interesting to learn how oxidant-specific structural changes in OxyR affect promoter binding affinity and other OxyR-target promoter interactions as well as the ability of OxyR to interact with regulatory partners. Our knowledge of the full spectrum of oxidants that are sensed by these regulators is likely to be incomplete. For example, OhrR has only recently been shown to sense hypochlorite (17), while its natural organic peroxide inducers remain unidentified.

Finally, ROS are an important signal to pathogens, indicating contact with the host. There is a growing realization of the role of peroxide-sensing regulators in adapting to this environmental shift through the control of processes such as virulence factor expression (74, 90) and biofilm formation (99, 119). Furthermore, the lethality of some antibiotics is due, at least in part, to increased production of ROS (58), and it has been demonstrated that the levels of oxidative stress protective enzymes can affect antibiotic resistance levels (6). Therefore, a detailed understanding of the mechanisms and physiological functions of oxidative stress responsive regulators will further our understanding of bacterial adaptation to environmental changes in general and is likely to shed new light on aspects of many clinically important processes.

ACKNOWLEDGMENTS

We apologize to those investigators whose work we did not cite due to space limitations. We thank Piyapol Munpiyamit for help with figure preparation.

Research funding to S.M. and J.M.D. was provided by grants from the Chulabhorn Research Institute, Mahidol University National Science and Technology Development Agency (NSTDA), and the Center of Excellence for Environmental Health, Toxicology, and Management of Chemicals.

REFERENCES

1. **Antelmann H, Helmann JD.** 2011. Thiol-based redox switches and gene regulation. Antioxid. Redox Signal. **14**:1049–1063.
2. **Aslund F, Zheng M, Beckwith J, Storz G.** 1999. Regulation of the OxyR transcription factor by hydrogen peroxide and the cellular thiol-disulfide status. Proc. Natl. Acad. Sci. U. S. A. **96**:6161–6165.
3. **Atichartpongkul S, Fuangthong M, Vattanaviboon P, Mongkolsuk S.** 2010. Analyses of the regulatory mechanism and physiological roles of *Pseudomonas aeruginosa* OhrR, a transcription regulator and a sensor of organic hydroperoxides. J. Bacteriol. **192**:2093–2101.
4. **Bae HW, Cho YH.** 2012. Mutational analysis of *Pseudomonas aeruginosa* OxyR to define the regions required for peroxide resistance and acute virulence. Res. Microbiol. **163**:55–63.
5. **Ballal A, Ray B, Manna AC.** 2009. *sarZ*, a *sarA* family gene, is transcriptionally activated by MgrA and is involved in the regulation of genes encoding exoproteins in *Staphylococcus aureus*. J. Bacteriol. **191**:1656–1665.
6. **Bizzini A, Zhao C, Auffray Y, Hartke A.** 2009. The *Enterococcus faecalis* superoxide dismutase is essential for its tolerance to vancomycin and penicillin. J. Antimicrob. Chemother. **64**:1196–1202.
7. **Brenot A, King KY, Caparon MG.** 2005. The PerR regulon in peroxide resistance and virulence of *Streptococcus pyogenes*. Mol. Microbiol. **55**:221–234.
8. **Brenot A, Weston BF, Caparon MG.** 2007. A PerR-regulated metal transporter (PmtA) is an interface between oxidative stress and metal homeostasis in *Streptococcus pyogenes*. Mol. Microbiol. **63**:1185–1196.
9. **Broadbent SE, Davies MR, van der Woude MW.** 2010. Phase variation controls expression of *Salmonella* lipopolysaccharide modification genes by a DNA methylation-dependent mechanism. Mol. Microbiol. **77**:337–353.
10. **Bsat N, Herbig AF, Cassillas-Martinez L, Setlow P, Helmann JD.** 1998. *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. Mol. Microbiol. **29**:189–198.
11. **Bundy BM, Collier LS, Hoover TR, Neidle EL.** 2002. Synergistic transcriptional activation by one regulatory protein in response to two metabolites. Proc. Natl. Acad. Sci. U. S. A. **99**:7693–7698.
12. **Calhoun LN, Kwon YM.** 2011. Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in *Escherichia coli*: a review. J. Appl. Microbiol. **110**:375–386.
13. **Chen H, et al.** 2008. A novel OxyR sensor and regulator of hydrogen peroxide stress with one cysteine residue in *Deinococcus radiodurans*. PLoS One **3**:e1602. doi:10.1371/journal.pone.0001602.
14. **Chen L, Keramati L, Helmann JD.** 1995. Coordinate regulation of

Bacillus subtilis peroxide stress genes by hydrogen peroxide and metal ions. Proc. Natl. Acad. Sci. U. S. A. **92**:8190–8194.
15. **Chen PR, et al.** 2006. An oxidation-sensing mechanism is used by the global regulator MgrA in *Staphylococcus aureus*. Nat. Chem. Biol. **2**:591–595.
16. **Chen PR, et al.** 2009. A new oxidative sensing and regulation pathway mediated by the MgrA homologue SarZ in *Staphylococcus aureus*. Mol. Microbiol. **71**:198–211.
17. **Chi BK, et al.** 2011. *S*-bacillithiolation protects against hypochlorite stress in *Bacillus subtilis* as revealed by transcriptomics and redox proteomics. Mol. Cell. Proteomics **10**:M111.009506.
18. **Choi H-J, et al.** 2001. Structural basis of the redox switch in the OxyR transcription factor. Cell **105**:103–113.
19. **Christman MF, Storz G, Ames BM.** 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. Proc. Natl. Acad. Sci. U. S. A. **86**:3484–3488.
20. **Chuchue T, et al.** 2006. *ohrR* and *ohr* are the primary sensor/regulator and protective genes against organic hydroperoxide stress in *Agrobacterium tumefaciens*. J. Bacteriol. **188**:842–851.
21. **Clark TJ, Momany C, Neidle EL.** 2002. The *benPK* operon, proposed to play a role in transport, is part of a regulon for benzoate catabolism in *Acinetobacter* sp. strain ADP1. Microbiology **148**:1213–1223.
22. **Cohen GM, d'Arcy Doherty M.** 1987. Free radical mediated cell toxicity by redox cycling chemicals. Br. J. Cancer Suppl. **8**:46–52.
23. **Craven SH, et al.** 2009. Inducer responses of BenM, a LysR-type transcriptional regulator from *Acinetobacter baylyi* ADP1. Mol. Microbiol. **72**:881–894.
24. **Cussiol JR, Alegria TGP, Szweda LI, Netto LES.** 2010. Ohr (organic hydroperoxide resistance protein) possesses a previously undescribed activity, lipoyl-dependent peroxidase. J. Biol. Chem. **285**:21943–21950.
25. **Eiamphungporn W, Soonsanga S, Lee J-W, Helmann JD.** 2009. Oxidation of a single active site suffices for the functional inactivation of the dimeric *Bacillus subtilis* OhrR repressor in vitro. Nucleic Acids Res. **37**:1174–1181.
26. **Faulkner MJ, Ma Z, Fuangthong M, Helmann JD.** 2012. Derepression of the *Bacillus subtilis* PerR peroxide stress response leads to iron deficiency. J. Bacteriol. **194**:1226–1235.
27. **Fuangthong M, Atichartpongkul S, Mongkolsuk S, Helmann JD.** 2001. OhrR is a repressor of *ohrA*, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. J. Bacteriol. **183**:4134–4141.
28. **Fuangthong M, Helmann JD.** 2002. The OhrR repressor senses organic hydroperoxides by reversible formation of a cysteine-sulfenic acid derivative. Proc. Natl. Acad. Sci. U. S. A. **99**:6690–6695.
29. **Fuangthong M, Helmann JD.** 2003. Recognition of DNA by three ferric uptake regulator (Fur) homologs in *Bacillus subtilis*. J. Bacteriol. **185**:6348–6357.
30. **Fuangthong M, Herbig AF, Bsat N, Helmann JD.** 2002. Regulation of the *Bacillus subtilis fur* and *perR* genes by PerR: not all members of the PerR regulon are peroxide inducible. J. Bacteriol. **184**:3276–3286.
31. **Gaballa A, Helmann JD.** 2002. A peroxide-induced zinc uptake system plays an important role in protection against oxidative stress in *Bacillus subtilis*. Mol. Microbiol. **45**:997–1005.
32. **Giro M, Carrillo N, Krapp AR.** 2006. Glucose-6-phosphate dehydrogenase and ferredoxin-NADP(H) reductase contribute to damage repair during the *soxRS* response of *Escherichia coli*. Microbiology **152**:1119–1128.
33. **Gonzalez-Flecha B, Demple B.** 1997. Homeostatic regulation of intracellular hydrogen peroxide concentration in aerobically growing *Escherichia coli*. J. Bacteriol. **179**:382–388.
34. **Gonzalez-Flecha B, Demple B.** 1999. Role for the *oxyS* gene in regulation of intracellular hydrogen peroxide in *Escherichia coli*. J. Bacteriol. **181**:3833.
35. **Gryllos I, et al.** 2008. PerR confers phagocytic killing resistance and allows pharyngeal colonization by group A *Streptococcus*. PLoS Pathog. **4**:e1000145. doi:10.1371/journal.ppat.1000145.
36. **Hahn J-S, Oh S-Y, Chater KF, Cho Y-H, Roe J-H.** 2000. H_2O_2-sensitive *fur*-like repressor CatR regulating the major catalase gene in *Streptomyces coelicolor*. J. Biol. Chem. **275**:38254–38260.
37. **Hahn J-S, Oh S-Y, Roe J-H.** 2002. Role of OxyR as a peroxide-sensing positive regulator in *Streptomyces coelicolor* A3(2). J. Bacteriol. **184**:5214–5222.
38. **Haridas V, et al.** 2005. Avicinylation (thioesterification): a protein modification that can regulate the response to oxidative and nitrosative stress. Proc. Natl. Acad. Sci. U. S. A. **102**:10088–10093.
39. **Harrison A, et al.** 2007. The OxyR regulon in nontypeable *Haemophilus influenzae*. J. Bacteriol. **189**:1004–1012.
40. **Hausladen A, Privalle CT, Keng T, De Angelo J, Stamler JS.** 1996. Nitrosative stress: activation of the transcription factor OxyR. Cell **86**:719–729.
41. **Hayashi K, Ohsawa T, Kobayashi K, Ogasawara N, Ogura M.** 2005. The H_2O_2 stress-responsive regulator PerR positively regulates *srfA* expression in *Bacillus subtilis*. J. Bacteriol. **187**:6659–6667.
42. **He X, et al.** 2010. Oral-derived bacterial flora defends its domain by recognizing and killing intruders—a molecular analysis using *Escherichia coli* as a model intestinal bacterium. Microb. Ecol. **60**:655–664.
43. **Helmann JD.** 2011. Bacillithiol, a new player in bacterial redox homeostasis. Antioxid. Redox Signal. **15**:123–133.
44. **Helmann JD.** 2002. OxyR: a molecular code for redox sensing. Sci. STKE **2002**(157):pe46. doi:10.1126/stke.20002.157.pe46.
45. **Heo Y-J, et al.** 2010. The major catalase gene (*katA*) of *Pseudomonas aeruginosa* PA14 is under both positive and negative control of the global transactivator OxyR in response to hydrogen peroxide. J. Bacteriol. **192**:381–390.
46. **Herbig AF, Helmann JD.** 2001. Roles of metal ions and hydrogen peroxide in modulating the interaction of the *Bacillus subtilis* PerR peroxide regulon repressor with operator DNA. Mol. Microbiol. **41**:849–859.
47. **Hong M, Fuangthong M, Helmann JD, Brennan RG.** 2005. Structure of an OhrR-*ohrA* operator complex reveals the DNA binding mechanism of the MarR family. Mol. Cell **20**:131–141.
48. **Horsburgh MJ, Clements MO, Crossley H, Ingham E, Foster SJ.** 2001. PerR controls oxidative stress resistance and iron storage proteins and is required for virulence in *Staphylococcus aureus*. Infect. Immun. **69**:3744–3754.
49. **Ieva R, et al.** 2008. OxyR tightly regulates catalase expression in *Neisseria meningitidis* through both repression and activation mechanisms. Mol. Microbiol. **70**:1152–1165.
50. **Imlay JA.** 2008. Cellular defenses against superoxide and hydrogen peroxide. Annu. Rev. Biochem. **77**:755–776.
51. **Imlay JA.** 2003. Pathways of oxidative damage. Annu. Rev. Microbiol. **57**:395–418.
52. **Jacquamet L, et al.** 2009. Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. Mol. Microbiol. **73**:20–31.
53. **Jonssona I-M, Lindholma C, Luongb TT, Leeb CY, Tarkowskia A.** 2008. *mgrA* regulates staphylococcal virulence important for induction and progression of septic arthritis and sepsis. Microbes Infect. **10**:1229–1235.
54. **Kaito C, Morishita D, Matsumoto Y, Kurokawa K, Sekimizu K.** 2006. Novel DNA binding protein SarZ contributes to virulence in *Staphylococcus aureus*. Mol. Microbiol. **62**:1601–1617.
55. **Keyer K, Imlay JA.** 1996. Superoxide accelerates DNA damage by elevating free-iron levels. Proc. Natl. Acad. Sci. U. S. A. **93**:13635–13640.
56. **Kim SO, et al.** 2002. OxyR: a molecular code for redox-related signaling. Cell **109**:383–396.
57. **Klomsiri C, Panmanee W, Dharmsthiti S, Vattanaviboon P, Mongkolsuk S.** 2005. Novel roles of *ohrR-ohr* in *Xanthomonas* sensing, metabolism, and physiological adaptive response to lipid hydroperoxide. J. Bacteriol. **187**:3277–3281.
58. **Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, Collins JJ.** 2007. A common mechanism of cellular death induced by bactericidal antibiotics. Cell **130**:797–810.
59. **Kona J, Brinck T.** 2006. A combined molecular dynamics simulation and quantum chemical study on the mechanism for activation of the OxyR transcription factor by hydrogen peroxide. Org. Biomol. Chem. **4**:3468–3478.
60. **Kullik I, Stevens J, Toledano MB, Storz G.** 1995. Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. J. Bacteriol. **177**:1285–1291.
61. **Kullik I, Toledano MB, Tartaglia LA, Storz G.** 1995. Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for oxidation and transcriptional activation. J. Bacteriol. **177**:1275–1284.
62. **Lamb C, Dixon RA.** 1997. The oxidative burst in plant disease resistance. Annu. Rev. Plant Physiol. Plant Mol. Biol. **48**:251–275.
63. **Lan L, Murray TS, Kazmierczak BI, He C.** 2010. *Pseudomonas aerugi-*

nosa OspR is an oxidative stress sensing regulator that affects pigment production, antibiotic resistance and dissemination during infection. Mol. Microbiol. **75**:76–91.
64. **Lau GW, Britigan BE, Hassett DJ.** 2005. *Pseudomonas aeruginosa* OxyR is required for full virulence in rodent and insect models of infection and for resistance to human neutrophils. Infect. Immun. **73**:2550–2553.
65. **Lee C, et al.** 2004. Redox regulation of OxyR requires specific disulfide bond formation involving a rapid kinetic reaction path. Nat. Struct. Mol. Biol. **11**:1179–1185.
66. **Lee J-W, Helmann JD.** 2006. Biochemical characterization of the structural Zn^{2+} site in the *Bacillus subtilis* peroxide sensor PerR. J. Biol. Chem. **281**:23567–23578.
67. **Lee J-W, Helmann JD.** 2006. The PerR transcription factor senses H_2O_2 by metal-catalysed histidine oxidation. Nature **440**:363–367.
68. **Lee J-W, Soonsanga S, Helmann JD.** 2007. A complex thiolate switch regulates the *Bacillus subtilis* organic peroxide sensor OhrR. Proc. Natl. Acad. Sci. U. S. A. **104**:8743–8748.
69. **Li H, Singh AK, McIntyre LM, Sherman LA.** 2004. Differential gene expression in response to hydrogen peroxide and the putative PerR regulon of *Synechocystis* sp. strain PCC 6803. J. Bacteriol. **186**:3331–3345.
70. **Liochev SI, Fridovich I.** 1992. Fumarase C, the stable fumarase of *Escherichia coli*, is controlled by the *soxRS* regulon. Proc. Natl. Acad. Sci. U. S. A. **89**:5892–5896.
71. **Liu Y, Bauer SC, Imlay JA.** 2011. The YaaA protein of the *Escherichia coli* OxyR regulon lessens hydrogen peroxide toxicity by diminishing the amount of intracellular unincorporated iron. J. Bacteriol. **193**:2186–2196.
72. **Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY.** 2006. Transcription profiling of the *mgrA* regulon in *Staphylococcus aureus*. J. Bacteriol. **188**:1899–1910.
73. **Luong TT, Newell SW, Lee CY.** 2003. *mgr*, a novel global regulator in *Staphylococcus aureus*. J. Bacteriol. **185**:3703–3710.
74. **Melstrom KAJ, et al.** 2007. Cytotoxicity of *Pseudomonas* secreted exotoxins requires OxyR expression. J. Surg. Res. **143**:50–57.
75. **Mongkolsuk S, Helmann JD.** 2002. Regulation of inducible peroxide stress responses. Mol. Microbiol. **45**:9–15.
76. **Mongkolsuk S, et al.** 2002. The repressor for an organic peroxide-inducible operon is uniquely regulated at multiple levels. Mol. Microbiol. **44**:793–802.
77. **Moore CM, Nakano MM, Wang T, Ye RW, Helmann JD.** 2004. Response of *Bacillus subtilis* to nitric oxide and the nitrosating agent sodium nitroprusside. J. Bacteriol. **186**:4655–4664.
78. **Morikawa K, et al.** 2006. Bacterial nucleoid dynamics: oxidative stress response in *Staphylococcus aureus*. Genes Cells **11**:409–423.
79. **Mukhopadhyay A, et al.** 2007. Cell-wide responses to low-oxygen exposure in *Desulfovibrio vulgaris* Hildenborough. J. Bacteriol. **189**:5996–6010.
80. **Mukhopadhyay P, Zheng M, Bedzyk LA, LaRossa RA, Storz G.** 2004. Prominent roles of the NorR and Fur regulators in the *Escherichia coli* transcriptional response to reactive nitrogen species. Proc. Natl. Acad. Sci. U. S. A. **101**:745–750.
81. **Newberry KJ, Fuangthong M, Panmanee W, Mongkolsuk S, Brennan RG.** 2007. Structural mechanism of organic hydroperoxide induction of the transcription regulator OhrR. Mol. Cell **28**:652–664.
82. **Ochsner UA, Vasil ML, Eyad A, Kislay P, Hassett DJ.** 2000. Role of the *Pseudomonas aeruginosa oxyR-recG* operon in oxidative stress defense and DNA repair: OxyR-dependent regulation of *katB-ankB*, *ahpB*, and *ahpC-ahpF*. J. Bacteriol. **182**:4533–4544.
83. **Oh S-Y, Shin J-H, Roe J-H.** 2007. Dual role of OhrR as a repressor and an activator in response to organic hydroperoxides in *Streptomyces coelicolor*. J. Bacteriol. **189**:6284–6292.
84. **Paget MS, Buttner MJ.** 2003. Thiol-based regulatory switches. Annu. Rev. Genet. **37**:91–121.
85. **Panmanee W, et al.** 2002. OhrR, a transcription repressor that senses and responds to changes in organic peroxide levels in *Xanthomonas campestris* pv. *phaseoli*. Mol. Microbiol. **45**:1647–1654.
86. **Panmanee W, Vattanaviboon P, Poole LB, Mongkolsuk S.** 2006. Novel organic hydroperoxide-sensing and responding mechanisms for OhrR, a major bacterial sensor and regulator of organic hydroperoxide stress. J. Bacteriol. **188**:1389–1395.
87. **Pomposiello PJ, Demple B.** 2002. Global adjustment of microbial physiology during free radical stress. Adv. Microb. Physiol. **46**:319–341.
88. **Poole LB.** 2005. Bacterial defenses against oxidants: mechanistic features of cysteine-based peroxidases and their flavoprotein reductases. Arch. Biochem. Biophys. **433**:240–254.
89. **Poor CB, Chen PR, Duguid EM, Rice PA, He C.** 2009. Crystal structures of the reduced, sulfenic acid, and mixed disulfide forms of SarZ, a redox active global regulator in *Staphylococcus aureus*. J. Biol. Chem. **284**:23517–23524.
90. **Ramsey MM, Whiteley M.** 2009. Polymicrobial interactions stimulate resistance to host innate immunity through metabolite perception. Proc. Natl. Acad. Sci. U. S. A. **106**:1578–1583.
91. **Rea R, Hill C, Gahan CGM.** 2005. *Listeria monocytogenes* PerR mutants display a small-colony phenotype, increased sensitivity to hydrogen peroxide, and significantly reduced murine virulence. Appl. Environ. Microbiol. **71**:8314–8322.
92. **Robinson JM.** 2009. Phagocytic leukocytes and reactive oxygen species. Histochem. Cell Biol. **131**:465–469.
93. **Schell MA.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. **47**:597–626.
94. **Seaver LC, Imlay JA.** 2001. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. J. Bacteriol. **183**:7173–7181.
95. **Seaver LC, Imlay JA.** 2001. Hydrogen peroxide fluxes and compartmentalization inside growing *Escherichia coli*. J. Bacteriol. **183**:7182–7189.
96. **Seib KL, et al.** 2007. Characterization of the OxyR regulon of *Neisseria gonorrhoeae*. Mol. Microbiol. **63**:54–68.
97. **Semchyshyn H, Bagnyukova T, Storey K, Storey V.** 2005. Hydrogen peroxide increases the activities of *soxRS* regulon enzymes and the levels of oxidized proteins and lipids in *Escherichia coli*. Cell Biol. Int. **29**:898–902.
98. **Seth D, Hausladen A, Wang Y-J, Stamler JS.** 2012. Endogenous protein *S*-nitrosylation in *E. coli*: regulation by OxyR. Science **336**:470–473.
99. **Shanks RM, et al.** 2007. A *Serratia marcescens* OxyR homolog mediates surface attachment and biofilm formation. J. Bacteriol. **189**:7262–7272.
100. **Sobota JM, Imlay JA.** 2011. Iron enzyme ribulose-5-phosphate 3-epimerase in *Escherichia coli* is rapidly damaged by hydrogen peroxide but can be protected by manganese. Proc. Natl. Acad. Sci. U. S. A. **108**:5402–5407.
101. **Soonsanga S, Lee J-W, Helmann JD.** 2008. Conversion of *Bacillus subtilis* OhrR from a 1-Cys to a 2-Cys peroxide sensor. J. Bacteriol. **190**:5738–5745.
102. **Soonsanga S, Lee J-W, Helmann JD.** 2008. Oxidant-dependent switching between reversible and sacrificial oxidation pathways for *Bacillus subtilis* OhrR. Mol. Microbiol. **68**:678–986.
103. **Storz G, Tartaglia LA, Ames BM.** 1990. Transcriptional regulator of oxidative stress inducible genes: direct activation by oxidation. Science **248**:189–194.
104. **Sund CJ, et al.** 2008. The *Bacteroides fragilis* transcriptome response to oxygen and H_2O_2: the role of OxyR and its effect on survival and virulence. Mol. Microbiol. **67**:129–142.
105. **Tao K.** 1999. In vivo oxidation-reduction kinetics of OxyR, the transcriptional activator for an oxidative stress-inducible regulon in *Escherichia coli*. FEBS Lett. **457**:90–92.
106. **Tao K, Zou C, Fujita N, Ishihama A.** 1995. Mapping of the OxyR protein contact site in the C-terminal region of RNA polymerase α subunit. J. Bacteriol. **177**:6740–6744.
107. **Tartaglia LA, Gimeno CJ, Storz G, Ames BN.** 1992. Multidegenerate DNA recognition by the OxyR transcriptional regulator. J. Biol. Chem. **267**:2038–2045.
108. **Toledano MB, et al.** 1994. Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. Cell **78**:897–909.
109. **Traoré DAK, et al.** 2006. Crystal structure of the apo-PerR-Zn protein from *Bacillus subtilis*. Mol. Microbiol. **61**:1211–1219.
110. **Traoré DAK, et al.** 2009. Structural and functional characterization of 2-oxo-histidine in oxidized PerR protein. Nat. Chem. Biol. **5**:53–59.
111. **Tseng H-J, McEwan AG, Apicella MA, Jennings MP.** 2003. OxyR acts as a repressor of catalase expression in *Neisseria gonorrhoeae*. Infect. Immun. **71**:550–556.
112. **van Vliet AHM, Baillon M-LA, Penn CW, Ketley JM.** 1999. *Campylobacter jejuni* contains two *fur* homologs: characterization of iron-responsive regulation of peroxide stress defense genes by the PerR repressor. J. Bacteriol. **181**:6371–6376.
113. **Wallecha A, Correnti J, Munster V, van der Woude M.** 2003. Phase

variation of Ag43 is independent of the oxidation state of OxyR. J. Bacteriol. **185**:2203–2209.
114. **Wallecha A, Munster V, Correnti J, Chan T, van der Woude M.** 2002. Dam- and OxyR-dependent phase variation of *agn43*: essential elements and evidence for a new role of DNA methylation. J. Bacteriol. **184**:3338–3347.
115. **Wang X, et al.** 2006. Mutational analysis to define an activating region on the redox-sensitive transcriptional regulator OxyR. J. Bacteriol. **188**:8335–8342.
116. **Wei Q, et al.** 2012. Global regulation of gene expression by OxyR in an important human opportunistic pathogen. Nucleic Acids Res. **40**:4320–4333.
117. **Wen Y-T, et al.** 2011. Differential secretomics of *Streptococcus pyogenes* reveals a novel peroxide regulator (PerR)-regulated extracellular virulence factor mitogen factor 3 (MF3). Mol. Cell. Proteomics **10**:M110.007013.
118. **Wu H-J, et al.** 2006. PerR controls Mn-dependent resistance to oxidative stress in *Neisseria gonorrhoeae*. Mol. Microbiol. **60**:401–416.
119. **Wu J, Lin X, Xie H, Burne R.** 2008. OxyR is involved in coordinate regulation of expression of *fimA* and *sod* genes in *Porphyromonas gingivalis*. FEMS Microbiol. Lett. **282**:188–195.
120. **Yin L, et al.** 2010. DRA0336, another OxyR homolog, involved in the antioxidation mechanisms in *Deinococcus radiodurans*. J. Microbiol. **48**:473–479.
121. **Zeller T, et al.** 2007. Regulation of hydrogen peroxide-dependent gene expression in *Rhodobacter sphaeroides*: regulatory functions of OxyR. J. Bacteriol. **189**:3784–3792.
122. **Zhang A, et al.** 1998. The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. EMBO J. **17**:6061–6068.
123. **Zheng M, Aslund F, Storz G.** 1998. Activation of the OxyR transcription factor by reversible disulfide bond formation. Science **279**:1718–1721.
124. **Zheng M, Doan B, Schneider TD, Storz G.** 1999. OxyR and SoxRS regulation of *fur*. J. Bacteriol. **181**:4639–4643.
125. **Zheng M, et al.** 2001. DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. J. Bacteriol. **183**:4562–4570.

JB Journals.ASM.org

MINIREVIEW

Many Means to a Common End: the Intricacies of (p)ppGpp Metabolism and Its Control of Bacterial Homeostasis

Anthony O. Gaca, Cristina Colomer-Winter, José A. Lemos

Center for Oral Biology and Department of Microbiology and Immunology, University of Rochester Medical Center, Rochester, New York, USA

In nearly all bacterial species examined so far, amino acid starvation triggers the rapid accumulation of the nucleotide second messenger (p)ppGpp, the effector of the stringent response. While for years the enzymes involved in (p)ppGpp metabolism and the significance of (p)ppGpp accumulation to stress survival were considered well defined, a recent surge of interest in the field has uncovered an unanticipated level of diversity in how bacteria metabolize and utilize (p)ppGpp to rapidly synchronize a variety of biological processes important for growth and stress survival. In addition to the classic activation of the stringent response, it has become evident that (p)ppGpp exerts differential effects on cell physiology in an incremental manner rather than simply acting as a biphasic switch that controls growth or stasis. Of particular interest is the intimate relationship of (p)ppGpp with persister cell formation and virulence, which has spurred the pursuit of (p)ppGpp inhibitors as a means to control recalcitrant infections. Here, we present an overview of the enzymes responsible for (p)ppGpp metabolism, elaborate on the intricacies that link basal production of (p)ppGpp to bacterial homeostasis, and discuss the implications of targeting (p)ppGpp synthesis as a means to disrupt long-term bacterial survival strategies.

While analyzing nucleotide extracts of *Escherichia coli*, Cashel and Gallant visualized two spots by thin-layer chromatography that could be implicated in the inhibition of stable RNA accumulation provoked by amino acid starvation, which they dubbed "magic spots" (1). These magic spots were later identified as the hyperphosphorylated guanosine derivatives ppGpp (GDP, 3′-diphosphate) and pppGpp (GTP, 3′-diphosphate), collectively referred to as (p)ppGpp, or alarmones (2, 3). Subsequent studies revealed that (p)ppGpp is responsible for activation of the stringent response (SR), a highly conserved stress response to nutrient starvation (4, 5). Generally speaking, accumulation of (p)ppGpp induces large-scale transcriptional alterations leading to general repression of genes required for rapid growth, such as rRNA genes, and concomitant activation of genes involved in amino acid biosynthesis, nutrient acquisition, and stress survival. In addition to transcriptional control, (p)ppGpp has been shown to directly inhibit the activity of several enzymes, including DNA primase, translation factors, and enzymes involved in GTP biosynthesis (6) (Fig. 1). Ultimately, the SR reallocates cellular resources toward adaptation to a semidormant state, facilitating survival under unfavorable conditions (5, 7). Although initially defined as a response to amino acid and carbon starvation, the term SR has since been expanded to include any regulatory effect exerted by cellular (p)ppGpp accumulation irrespective of the triggering mechanism (4).

The broad physiological alterations induced by (p)ppGpp accumulation rely heavily upon transcriptional alterations. In *Gammaproteobacteria*, such as the Gram-negative paradigmatic organism *E. coli*, transcriptional control during the SR is predominately accomplished through the direct interaction of (p)ppGpp with the interface of the β′ and ϖ subunits of RNA polymerase (RNAP) (8–10) and is greatly potentiated by the presence of DksA, a small protein that binds to the RNAP secondary channel (11). As a general rule, the discriminator region between the −10 sequence and the transcriptional start site dictates whether (p)ppGpp will function as a repressor (GC-rich region) or as an activator (AT-rich region) (12, 13). Moreover, (p)ppGpp indirectly regulates transcription by either stabilizing the binding of alternative σ factors or interfering with the activity of transcriptional regulators (14–17).

Despite commonalities in the general outcome, the mechanisms of transcriptional control exerted by (p)ppGpp in low-GC Gram-positive bacteria (*Firmicutes*) are fundamentally distinct from those described in *E. coli*. *Firmicutes* lack a DksA homolog and GC- or AT-rich discriminators and, perhaps most importantly, (p)ppGpp does not physically interact with RNAP (18). Rather, (p)ppGpp indirectly affects transcription in this bacterial group by regulating the concentration of the initiating nucleotide of transcription (iNTP, or position +1). For example, the *Bacillus subtilis rrn* operons use GTP as the iNTP, while the *ilv-leu* operon, which is responsible for branch chain amino acid (BCAA) biosynthesis, uses ATP (19, 20). It follows that in *B. subtilis*, and likely other *Firmicutes*, SR induction and (p)ppGpp accumulation correspond to a sharp drop in the GTP level that is accompanied by an increase in the ATP pool (21). In addition to changes in iNTP pools, *Firmicutes* evolved a second regulatory network based on the inverse relationship between (p)ppGpp and GTP. The DNA-binding capacity, and therefore transcriptional regulation, of the global metabolic regulator CodY is modulated by its interaction with GTP and BCAA, particularly isoleucine (22). The link between (p)ppGpp and CodY in *Firmicutes* has been recently reviewed (23) and will be only briefly discussed in this article. Col-

Accepted manuscript posted online 20 January 2015

Citation Gaca AO, Colomer-Winter C, Lemos JA. 2015. Many means to a common end: the intricacies of (p)ppGpp metabolism and its control of bacterial homeostasis. J Bacteriol 197:1146–1156. doi:10.1128/JB.02577-14.

Editor: W. Margolin

Address correspondence to José A. Lemos, jose_lemos@urmc.rochester.edu.

doi:10.1128/JB.02577-14

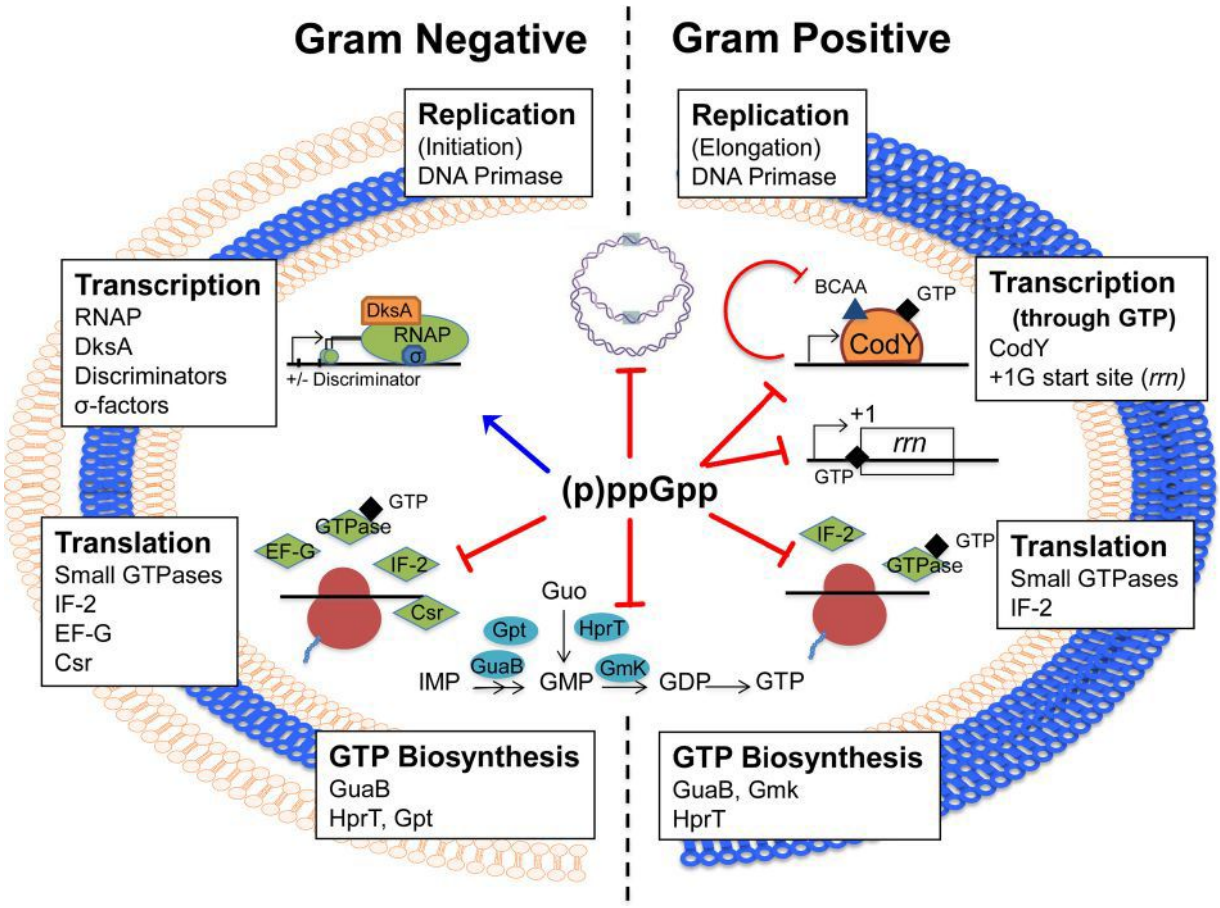


FIG 1 Main targets of (p)ppGpp in *Firmicutes* and *Gammaproteobacteria*. The regulatory nucleotides (p)ppGpp alter cellular metabolism in response to stress by directly binding to a variety of enzymes. For transcription in *Gammaproteobacteria*, (p)ppGpp binds directly to RNAP together with the DksA coeffector to modulate promoter selection (8, 9, 11). Repression or activation of target genes is dependent on discriminator sequences encoded in the promoter region (12). Furthermore, (p)ppGpp promotes alternative σ factor usage (64). In *Firmicutes*, (p)ppGpp is unable to bind to RNAP (8, 9). In this case, transcription is indirectly regulated through depletion of GTP. Low intracellular levels of GTP and BCAA alleviate repression of the global transcriptional regulator CodY, leading to activation of amino acid biosynthesis, nutrient transport, and virulence genes (22, 23, 67). In addition, GTP depletion represses transcription of genes, such as rRNA (*rrn*) operons, requiring GTP as their initiating nucleotide (+1G). For GTP biosynthesis in *Gammaproteobacteria*, (p)ppGpp lowers GTP production by inhibiting IMP dehydrogenase (GuaB), the first enzyme in the *de novo* guanine nucleotide synthesis pathway (107), as well as hypoxanthine phosphoribosyltransferase (HprT) and xanthine-guanine phosphoribosyltransferase (Gpt). In *Firmicutes*, (p)ppGpp rapidly depletes GTP levels through inhibition of GuaB, HprT, and guanylate kinase (Gmk), halting both *de novo* and salvage pathways (67). For translation in *Gammaproteobacteria*, (p)ppGpp directly alters translation initiation and elongation by inactivating initiation factor 2 (IF-2) and elongation factor G (EF-G) (108, 109). Moreover, ribosomal maturation and mRNA translation/stability are affected by inhibition of small GTPases and interaction with the Csr system, respectively (6, 110, 111). In *Firmicutes*, (p)ppGpp directly inhibits protein translation through inactivation of small GTPases and IF-2 (108, 112). For DNA replication, (p)ppGpp-specific inhibition of DNA primase is observed in both bacterial groups, whereas replication is blocked at the initiation step in *Gammaproteobacteria* and at the elongation step in *Firmicutes* (113–115).

lectively, these results indicate that (p)ppGpp indirectly controls the action of RNAP in *Firmicutes* through modulation of intracellular purine concentrations (19, 24, 25).

After its discovery in the late 1960s, a series of contemporaneous investigations defined many of the currently accepted features of (p)ppGpp and the SR. These features include the identification and characterization of the major enzymes responsible for the metabolism of (p)ppGpp, the first glimpses into the mode of action of (p)ppGpp as a regulatory nucleotide, and determination of the pleiotropic effects it exerts on cell physiology. Despite steady progress in (p)ppGpp research since its discovery, the arrival of the genomic era in the late 1990s and the recurrent association of (p)ppGpp with virulence expression and antibiotic persistence reignited the field in recent years. As a result, the number of investigations linking (p)ppGpp to a plethora of disparate processes, such as growth rate control, motility, sporulation, biofilm formation, competence, stress tolerance, persistence, and virulence, have dramatically increased. As part of this new surge in "alarmone" research, new enzymatic sources of (p)ppGpp synthesis and degradation were discovered, and a better appreciation for the biological relevance of (p)ppGpp beyond SR activation has been developed. In the next pages, we focus on recent insights into the diversity of enzymes involved in (p)ppGpp metabolism and elaborate on the intricacies that link incremental production of (p)ppGpp to bacterial homeostasis. Recent developments on antimicrobial therapies that target alarmone production are also discussed. For a complete survey of the field, we direct the reader to other recent reviews (5, 7, 18, 26, 27).

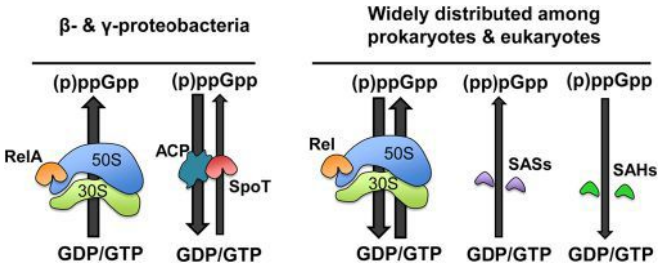


FIG 2 RelA-SpoT homolog (RSH) family enzymes controlling (pp)pGpp metabolism. In *Betaproteobacteria* and *Gammaproteobacteria*, the synthesis and hydrolysis of (p)ppGpp is catalyzed by RelA and SpoT. RelA functions as the primary (p)ppGpp synthetase responsible for induction of the SR. SpoT is bifunctional, acting as the primary (p)ppGpp hydrolase but also capable of weak (p)ppGpp synthetase activity in response to nutritional signals not sensed by RelA (5, 116, 117). Outside of *Betaproteobacteria* and *Gammaproteobacteria*, the bifunctional Rel is the primary enzyme responsible for (p)ppGpp metabolism in most bacterial species (39). Like RelA, Rel is responsible for the induction of the SR during amino acid starvation and, like SpoT, responsible for the hydrolysis of (p)ppGpp (118). In addition, many bacteria can have one or more stand-alone SAS enzymes that appear to have weak but constitutive activity (39, 43, 44, 46, 47). Small alarmone hydrolases (SAHs) have so far only been characterized in higher eukaryotes, but they appear to have potent hydrolase activity comparable to SpoT (56).

CHANGING PARADIGMS: THE DISCOVERY OF "SHORT" RelA/SpoT HOMOLOGS

Even before the discovery of (p)ppGpp, mutations abolishing stringent control in *E. coli* had been mapped to the RNA control (RC) locus. The RC locus was later identified as the site of the *relA* gene and named for the "relaxed" phenotype of these mutants that no longer exhibited an inverse relationship between amino acid availability and stable RNA accumulation (28). RelA is found in *Betaproteobacteria* and *Gammaproteobacteria* and is chiefly responsible for the accumulation of (p)ppGpp during amino acid starvation in these organisms (Fig. 2) (5). Initially, the association of RelA with stalled ribosomes containing uncharged tRNAs at the acceptor site was shown to trigger (p)ppGpp synthesis. Some 30 years later, the enzyme was predicted to "hop" from one stalled ribosome to another, such that alarmone production accurately reflected the number of starved ribosomes (29). Recently, single-molecule *in vivo* investigations have indicated that RelA is tightly bound to the ribosome during active translation but rapidly dissociates when starvation is induced and deacylated tRNA accumulates (30). Once off the ribosome, RelA performs multiple rounds of catalysis in what has been termed the "extended hopping" mechanism (30). Interestingly, heat shock promotes a similar but more transient dissociation, suggesting a mechanism by which other non-starvation-based stresses induce (p)ppGpp synthesis (30). Furthermore, RelA activity is induced by ppGpp, a positive allosteric feedback mechanism that contributes to the rapid activation of the SR (31).

The second enzyme regulating (p)ppGpp accumulation in *Betaproteobacteria* and *Gammaproteobacteria* is the RelA homolog SpoT (Fig. 2). SpoT possesses both synthetic and hydrolytic activities, albeit the synthetase activity is weak compared to RelA. Interestingly, (p)ppGpp synthesis by SpoT is induced by unique signals not sensed by RelA, which include carbon, fatty acid, and iron starvation (32–34). In addition, SpoT interacts with the GTP-binding protein CgtA (35, 36), which has been proposed to promote hydrolase activity, thereby maintaining low (p)ppGpp levels under nutrient-rich conditions (37). However, the effects of CgtA on basal (p)ppGpp pools are quantitatively minor and deserve further verification.

For some time, RelA and SpoT were considered the enzymatic paradigms for (p)ppGpp metabolism. However, over the past 2 decades, our understanding of the enzymes controlling the synthesis and degradation of (p)ppGpp has undergone several adjustments. The first modification came with the characterization of a RelA/SpoT homolog (RSH) protein from *Streptococcus dysgalactiae* subsp. *equisimilis*, named Rel_{Seq}, containing functional characteristics of both RelA and SpoT, i.e., strong synthetase activity like RelA and hydrolase activity like SpoT (38) (Fig. 2). Recently, phylogenetic studies indicated that the combined strong synthetase and hydrolase activities characteristic of Rel enzymes are ancestral to RelA and SpoT and more widely distributed in prokaryotes, including *Actinobacteria*, *Firmicutes*, and *Alpha-*, *Delta-* and *Epsilonproteobacteria* among many others (39). *In vitro* studies with the *Mycobacterium tuberculosis* Rel (Rel_{Mtb}) revealed that, like RelA, synthetase activity is stimulated by a complex of ribosomes, uncharged tRNA, and mRNA (40). Finally, biochemical and structural studies revealed that synthetase and hydrolase activities of Rel_{Seq} are reciprocally regulated through two mutually exclusive conformations adapted by the catalytic N-terminal half of the full-length protein (41, 42). When Rel is in a hydrolase-ON state, the active configuration of the hydrolase domain is structurally different from that of the inactive synthetase domain. When the synthetase domain is ligand bound, it becomes active, placing the enzyme in a synthetase-ON and hydrolase-OFF state. It should be noted that a certain level of ambiguity has and still exists around the nomenclature of this enzyme. Rel has been commonly referred to as both RelA, because of its essentiality to SR activation, and Rsh, due to its RelA/SpoT hybrid characteristic. Recently, Atkinson and colleagues proposed that the enzymes involved in (p)ppGpp metabolism be divided into two categories: "long," multidomain RSHs, which include the bifunctional Rel and SpoT and the monofunctional RelA, and the more recently discovered "short" single-domain RSHs described below (Fig. 2) (39). In the interest of a unified nomenclature, we will follow this naming convention here.

More recently, the genomes of different Gram-positive organisms were found to encode single-domain (p)ppGpp synthetases, known as small alarmone synthetases (SASs), which lack both the C-terminal regulatory domain and the Mn^{2+}-dependent hydrolase domain present in long RSHs (43). These enzymes are ubiquitous in *Firmicutes* and have also been found in *Actinobacteria* and *Vibrio* species (39, 44–47).

To date, SASs have been best characterized in *Firmicutes* that encode the RelP and RelQ enzymes, albeit RelP is absent in certain species (39, 43). Several lines of evidence indicate that SAS activity is important for maintaining low basal levels of (p)ppGpp under balanced growth conditions, as their inactivation reverted the slow growth phenotype of Δ*rel* strains, a finding that was attributed to high basal levels of (p)ppGpp (43–45). Moreover, both *relP* and *relQ* were shown to contribute to the essentiality of hydrolase activity of the Rel enzyme in *Staphylococcus aureus*, and spontaneously generated mutations on both *relP* (*yjbM*) and *relQ* (*ywaC*) were found to suppress the growth defect caused by *rel* inactivation in *B. subtilis* (47, 48). In *M. tuberculosis*, inactivation of the SAS, Rv1366, did not provide a noticeable phenotype, even when introduced in the rel_{Mtb} background, suggesting Rv1366

lacks appreciable synthetase activity *in vivo* (49). This is consistent with recent studies by Bag and colleagues, who demonstrated that Rv1366 is unable to synthesize ppGpp *in vitro* (50). In contrast, the *Mycobacterium smegmatis* SAS, MSM_5849, exhibited weak (p)ppGpp synthetase activity as well as Mn(II)-dependent RNase HII activity (51). Of note, this SAS-RNase HII fusion appears to occur with some frequency among species of mycobacteria. Finally, one SAS, termed RelV, has been characterized in *V. cholerae* and found to be highly conserved among *Vibrio* species (46). The presence of RelV was unexpected, as *Vibrio* species encode the canonical *relA*/*spoT* genes of other *Gammaproteobacteria* and this is the only known example of a SAS-containing enzyme in this bacterial group (46). RelV was shown to efficiently synthesize (p)ppGpp both *in vitro* and *in vivo* (46, 52).

The seemingly stable coexistence of long RSHs with SASs suggests some evolutionary advantage to this apparent enzymatic redundancy. The contribution of a SAS to stress tolerance and cell homeostasis, particularly in the presence of a fully active Rel enzyme, is not entirely clear, albeit recent studies have shed some light onto their biological roles. In *S. aureus*, simultaneous inactivation of *relP* and *relQ* significantly enhanced susceptibility to cell wall-targeting antibiotics (47). Interestingly, competence and basal (p)ppGpp production by RelP appears to be interconnected in *Streptococcus mutans*. The transcription of the *S. mutans relP* is under the control of and cotranscribed with the RelRS two-component system. A gene cluster located upstream of *relP* encoding two ABC transporters (RcrPQ) and a DNA-binding transcriptional regulator (RcrR) were found to be critical for competence development and maintenance of basal (p)ppGpp through activation of *relRSP* (53). It has been hypothesized that the Rcr (Rel competence-related) operon secretes a quorum-sensing molecule sensed by RelRS that modulates cell growth and competence in response to specific signals (53). As detailed below, low levels of (p)ppGpp produced by RelQ, the sole SAS found in enterococci, are sufficient for persistence and full virulence in *Enterococcus faecalis* (45, 54). However, this linkage of low basal (p)ppGpp levels, rather than the SR, to virulence is not entirely unexpected. In several Gram-negative pathogens, virulence is only abolished or attenuated in *relA spoT* double-knockout strains and not in *relA* single mutants, which produce low levels of (p)ppGpp due to weak SpoT synthetase activity (7, 26). Based on these pioneering studies, it appears that bacteria could utilize SASs to fine-tune cell physiology in a species-specific manner.

During the biochemical characterization of *E. faecalis* RelQ, we recently found that this enzyme was able to efficiently utilize GMP as a pyrophosphate acceptor to synthesize pGpp (GMP, 3′-diphosphate). We followed this observation by showing that pGpp has specific regulatory effects on (p)ppGpp-controlled processes, including strong inhibition of enzymes involved in GTP biosynthesis (A. O. Gaca, C. Colomer-Winter, P. Kudrin, J. Beljantseva, K. Liu, B. Anderson, J. D. Wang, D. Rejman, K. Potrykus, M. Cashel, V. Hauryliuk, and J. A. Lemos, submitted for publication). Interestingly, pGpp was previously shown to accumulate in *B. subtilis* during amino acid limitation, but its source, either direct synthesis from GMP or nonspecific degradation of (p)ppGpp, was not determined (55).

The most recent major discovery in (p)ppGpp metabolism was the identification of stand-alone small alarmone hydrolases (SAHs) (Fig. 2) (56). These enzymes were first discovered in metazoans and were named Mesh1. Deletion of Mesh1 in *Drosophila melanogaster* resulted in slow body growth and impaired starvation resistance, features reminiscent of the bacterial SR. In fact, heterologous expression of Mesh1 in *E. coli* can substitute for the absence of SpoT-mediated hydrolase activity and abolish RelA-induced cell growth delays (56). The fact that (p)ppGpp has not been detected in eukaryotes besides plant chloroplasts suggests that Mesh1 might function to degrade related polyphosphorylated nucleotides with a similar function in metazoans. Putative SAH homologs have since been identified in a diverse range of organisms, including prokaryotes, but functional studies are still missing (39). In addition to SAHs, nonspecific hydrolases, including Nudix and the phosphohydrolase MazG, also appear to be important for maintaining intracellular (p)ppGpp at levels compatible with balanced growth (57, 58). The presence of nonspecific hydrolases may explain why deletion of the primary (p)ppGpp hydrolase (i.e., Rel) is not lethal in some SAS-encoding bacterial species.

(p)ppGpp BEYOND THE SR: BASAL (p)ppGpp POOLS AND CELL HOMEOSTASIS

Most research regarding the protective and regulatory aspects of (p)ppGpp has focused on SR induction. Here, we have defined the SR as the rapid and dramatic accumulation of (p)ppGpp that occurs during stress, resulting in the strong repression of macromolecular biosynthesis and activation of stress survival pathways. This is not to say that the contributions of subtle fluctuations in (p)ppGpp pools to the control of core cellular processes and growth rate have been completely overlooked (59–62). However, more recently, the notion that (p)ppGpp exerts important regulatory effects at concentrations below those needed to activate the SR has been more thoroughly revisited (10, 54, 63, 64). In an elegant study using cells that were progressively starved, the Conway lab identified two distinct regulatory cascades activated by (p)ppGpp in *E. coli* (64). First, at slightly elevated (p)ppGpp concentrations, the leucine-responsive protein (Lrp) regulon, a global regulator of genes involved in amino acid biosynthesis and transport, is activated (65). At higher (p)ppGpp concentrations, the RpoS regulon, controlled by the general stress sigma factor (σ^S), is induced. Thus, *E. coli* cells have discretely calibrated responses to a gradient of (p)ppGpp (64). In this biphasic stress response, activation of Lrp represents an initial attempt to restore intracellular amino acid pools at the onset of starvation. However, if this homeostatic mechanism fails to meet cellular amino acid demands, the RpoS regulon is activated to ensure survival (Fig. 3A). Of note, Lrp is restricted to *Betaproteobaceria* and *Gammaproteobacteria* and, although more widely distributed than Lrp, σ^S is absent in several bacteria, including some major pathogenic species. Thus, the concentration-dependent regulatory networks of (p)ppGpp in other bacteria have yet to be investigated.

We have explored in some detail the importance of (p)ppGpp outside the SR in the opportunistic pathogen *E. faecalis*. During phenotypic characterization of Δ*rel* and Δ*rel* Δ*relQ* [(p)ppGpp^0] strains, both lacking the ability to mount the SR, we observed important differences between the two strains that could be clearly attributed to either the presence or absence of RelQ. In the Δ*rel* strain, basal levels of (p)ppGpp synthesized by RelQ supported wild-type levels of antibiotic persistence and virulence (45, 54, 66). Only the deletion of both *rel* and *relQ*, completely abolishing (p)ppGpp synthesis, resulted in virulence attenuation and increased sensitivity to antibiotics (45, 54, 66). These phenotypes can be explained, in part, by the metabolic dysregulation observed

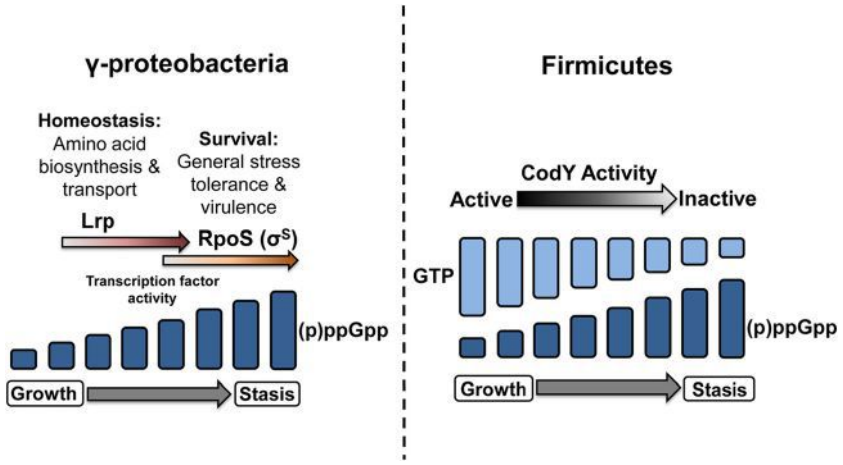


FIG 3 Simplified view of the concentration-dependent effects of (p)ppGpp on growth and survival in the most studied bacterial groups. In *Gammaproteobacteria*, Lrp and RpoS (σ^S) function as discretely calibrated regulatory units triggered by increasing concentrations of (p)ppGpp (64). Moderate accumulation of (p)ppGpp due to mild nutrient limitation activates expression of the Lrp regulon, which is responsible for inducing pathways needed to restore metabolic homeostasis. Severe nutrient limitation causes a more dramatic increase in (p)ppGpp pools responsible for activation of the RpoS (σ^S) regulon, inducing the expression of stress survival genes. In *Firmicutes*, an inverse relationship between (p)ppGpp and GTP dictates whether cells maintain a physiological program compatible with growth or switch to a state best suited for survival under slow growth or stasis (23, 71). Under nutrient-rich conditions compatible with rapid growth, CodY is active, controlling gene expression to direct pyruvate into fermentation and secretion of by-products like lactate, acetate, and ethanol. Nutrient starvation triggers the accumulation of (p)ppGpp and a subsequent drop in GTP. Alleviation of CodY regulation due to BCAA and GTP depletion activates transcription of genes involved in nutrient transport, amino acid biosynthesis, and virulence. The inability to maintain GTP and (p)ppGpp homeostasis leads to extreme fluctuations in GTP which can be highly inhibitory and, possibly, lethal (67). This cartoon is not inclusive of all pathways affected by alteration in (p)ppGpp or GTP levels but highlights the unique pathways used by several well-studied bacterial groups to propagate incremental alterations in these nucleotide pools.

in the Δ*rel* Δ*relQ* strain (54). In the absence of (p)ppGpp, *E. faecalis* undergoes large-scale transcriptional alterations in secondary carbon metabolism, which result in a shift from a homolactic to a heterofermentative metabolism with a concomitant increase in H_2O_2 production (54). These results suggest that *E. faecalis*, and likely other *Firmicutes*, depends on small amounts of (p)ppGpp to control the pace and direction of carbon flow and to accurately respond to external and internal metabolic cues. In addition, complete loss of (p)ppGpp also leads to dysregulation of GTP homeostasis (see below), which can severely impair cell fitness (54, 67).

In both *B. subtilis* and *E. faecalis*, (p)ppGpp directly and strongly inhibits GTP biosynthesis by targeting HprT (hypoxanthine-guanine phosphoribosyltransferase) and Gmk (GMP kinase). The 50% inhibitory concentrations of (p)ppGpp for HprT and Gmk are extremely low, ranging from 11 to 80 μM, indicating that basal (p)ppGpp levels are important for GTP homeostasis (54, 68). This is evidenced by the observation that the addition of exogenous guanosine to (p)ppGpp^0 strains, which is converted to GTP via the salvage pathway, significantly increased GTP levels, whereas GTP levels remained constant in (p)ppGpp$^+$ strains (54, 68). The accumulation of GTP in (p)ppGpp^0 strains was highly inhibitory and could even induce cell death ("death by GTP") in *B. subtilis* (67). In contrast, in *E. coli* uptake of purine as well as pyrimidines is blocked by ppGpp, and yet the presence of both bases or nucleosides in the medium does not give lethal elevations of the corresponding nucleotides in (p)ppGpp^0 strains (69). In addition to the association with rRNA gene transcription and CodY activation (see below), GTP activates a wide variety of anabolic processes, including nucleic acid synthesis and all three steps of translation (70). Based on the central role of GTP in cell homeostasis, it is not surprising that a tight inverse correlation between GTP pools and bacterial survival has been observed (67). The reduction of intracellular GTP pools in *B. subtilis*, either through mutation of enzymes in the *de novo* GTP biosynthesis pathway or by treatment with decoyinine, a selective inhibitor of GMP synthetase, enhanced resistance to amino acid limitation (67). Collectively, these studies indicate that one of the primary protective mechanisms of (p)ppGpp is to regulate intracellular GTP pools (Fig. 3B). As discussed below, the inverse relationship of GTP and (p)ppGpp may function as a metabolic switch that controls growth and survival.

(p)ppGpp AND CodY

As stated before, *Firmicutes* also integrate the inverse relationship between (p)ppGpp and GTP levels to control activity of the transcriptional regulator CodY (23). Specifically, reduction in cellular pools of the coeffector GTP by (p)ppGpp results in less stable CodY-DNA interactions, thereby alleviating CodY regulation (71). Importantly, the (p)ppGpp/CodY association is critical for nutrient stress tolerance and virulence in several bacterial pathogens (72–77). Yet, the relative contributions of GTP- and CodY-dependent mechanisms for stress tolerance and virulence controlled by (p)ppGpp are not entirely clear, as GTP depletion can also affect stress survival in a CodY-independent manner (67, 68). In addition, CodY in both *Lactococcus* and *Streptococcus* species appears to be insensitive to GTP (73, 78, 79). Parsing the contributions of GTP pool fluctuations and CodY to (p)ppGpp-controlled phenomena in *Firmicutes* will be an important step in fully understanding the protective effects of (p)ppGpp. Additionally, it will be important to determine if more subtle changes in GTP levels due to maintenance of basal (p)ppGpp pools are sufficient to influence CodY activity.

THE CENTRAL ROLE OF (p)ppGpp IN BACTERIAL PERSISTENCE

Bacterial persisters are phenotypic variants that enter a slow-growing or dormant state and transiently become multidrug tol-

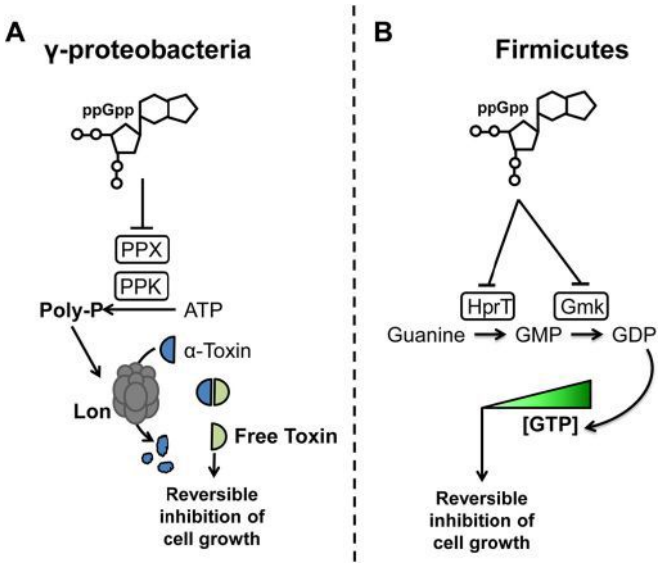


FIG 4 (p)ppGpp-triggered systems controlling persister cell formation. (A) In *E. coli*, (p)ppGpp sits atop a sequential biochemical signaling network for the induction of persisters. The accumulation of (p)ppGpp inhibits the activity of exopolyphosphatase (PPX), allowing poly(P) to accumulate. Poly(P) then activates Lon protease to target and degrade antitoxins of type II toxin-antitoxin (TA) pairs. Free toxins go on to inhibit processes essential for active cell growth (89). (B) The nucleotides (p)ppGpp and GTP may act as a metabolic switch in *Firmicutes*. The accumulation of (p)ppGpp leads to a dramatic reduction in GTP levels by directly consuming GTP and GDP but also inhibits the activity of Gmk and HprT by blocking both *de novo* and salvage pathways of GTP biosynthesis (67). GTP is an essential cofactor in numerous anabolic process needed for cell growth. This reduction in GTP slows bacterial growth and leads to alleviation of CodY regulation (see Fig. 3). An inverse correlation between GTP and cell fitness has been observed, with a reduction in GTP pools having a general protective effect (67, 119).

erant (80). Because persisters are able to grow when the antibiotic is removed, bacterial persistence has been implicated in chronic and recurrent infections, particularly those of biofilm origin. Although (p)ppGpp accumulation has been linked to antibiotic tolerance for a while (81–84), only recently have the underlying mechanisms by which (p)ppGpp mediates bacterial persistence begun to be elucidated (Fig. 4). In particular, different laboratories have systematically shown that increased (p)ppGpp levels parallel observed increases in persistence (85–89). By coupling microfluidics with fluorescent reporter gene fusions in *E. coli*, the Gerdes group demonstrated that (p)ppGpp levels vary stochastically in exponentially growing cultures, and these investigators confirmed that persistence and (p)ppGpp levels are positively correlated at the single-cell level (89). In the current *E. coli*-based model, (p)ppGpp triggers persistence by activation of toxin-antitoxin (TA) loci through a regulatory cascade that involves (i) allosteric inhibition by (p)ppGpp of the exopolyphosphatase (Ppx) enzyme responsible for inorganic polyphosphate (PolyP) degradation, (ii) activation of the Lon protease by poly(P), (iii) Lon-dependent degradation of antitoxins from different TA modules, and (iv) inhibition of transcription and translation by the free toxins of the TA modules (86). Interestingly, the Hip (high-persistence) TA system, the first genetic factor implicated in persister cell formation (90), was shown to impair glutamyl tRNA synthase (GltX) activity via direct phosphorylation of GltX by the HipA toxin (91, 92). This phosphorylation event was shown to induce accumulation of uncharged tRNAs, thereby triggering substantial amounts of (p)ppGpp production through RelA. This persister-activation cascade is further supported by studies with *Salmonella enterica* serovar Typhimurium that have shown that (p)ppGpp, Lon, and TA modules are also required for *Salmonella* persistence, either within macrophages or during antibiotic treatment (93–95). In addition to stochastic activation of TA modules by (p)ppGpp, Nguyen and colleagues proposed that the SR protected *E. coli* and *Pseudomonas aeruginosa* from the lethal effects of antibiotics through an active mechanism, which included the induction of antioxidant defenses (96, 97).

Although reasonable progress has been made in understanding the underlying mechanisms by which (p)ppGpp mediates persistence in *E. coli* and related *Gammaproteobacteria*, homologous molecular pathways leading to persistence in other bacterial groups remain poorly understood. Yet, recent studies support the involvement of (p)ppGpp in persister cell formation in members of the *Firmicutes* phylum. For example, whole-genome sequencing analysis of bacterial isolates from a patient with recurrent *S. aureus* infections identified a single-nucleotide substitution in the *rel* gene that affected hydrolase activity and caused accumulation of (p)ppGpp (98). Similar observations have been made with laboratory strains, as increased (p)ppGpp production due to spontaneous point mutations in *rel* were observed in *S. aureus* populations that survived lethal doses of methicillin (87, 88).

In agreement with the generalized role of (p)ppGpp in persister activation, complete lack of (p)ppGpp in *E. faecalis* (Δ*rel* Δ*relQ* strain) dramatically reduced the number of persisters (45, 54, 66). However, loss of Rel did not lower persistence rates and, depending on the drug target, the Δ*rel* strain, which has ~4-fold-higher basal levels of ppGpp due to constitutive alarmone synthesis by RelQ (54), produced a significantly higher number of persisters than the parent strain (45). This finding is a departure from the general concept that the SR mediates bacterial persistence, as activation of the SR in *E. faecalis* and *Firmicutes* in general is dependent on the Rel enzyme (43, 45, 66, 74, 99–101). Resolution of this notion will require systematic studies on basal (p)ppGpp elevation in other related species. Taking into consideration that RelP and RelQ mediate tolerance against cell wall-active antibiotics in S. *aureus*, it becomes clear that the association of (p)ppGpp with persistence in *Firmicutes* does not relate directly to the SR but rather to basal (p)ppGpp pools. Based on the critical role of (p)ppGpp in GTP homeostasis (54, 67), it is tempting to speculate that (p)ppGpp mediates persistence in *Firmicutes* via GTP regulation. Thus, while (p)ppGpp may indeed function as a universal mediator for persister formation, the underlying molecular mechanisms acting upstream and downstream of (p)ppGpp signaling may vary among bacteria. Clearly, additional investigations are necessary to address this possibility, particularly studies that draw a direct comparison between the mechanisms of persistence in phylogenetically diverse bacteria.

(p)ppGpp AS A TARGET FOR ANTIMICROBIAL DRUG DEVELOPMENT

The intimate association of (p)ppGpp regulation in bacterial persistence and virulence makes (p)ppGpp signaling interference a promising target for drug development. In fact, two recent studies have confirmed the potential usefulness of antimicrobial approaches that interfere with (p)ppGpp accumulation. In the first study, Relacin, a (p)ppGpp analog, was shown to specifically in-

hibit synthesis activity of RelA and Rel enzymes (102). Relacin was also shown to reduce *in vitro* survival of *B. subtilis* and *Streptococcus pyogenes* and hinder biofilm formation and sporulation processes in *B. subtilis* (102). In a more recent study, a broad-spectrum peptide (peptide 1018) was found to specifically bind and promote (p)ppGpp degradation (103). Interestingly, peptide 1018 showed stronger activity toward biofilms causing cell death at concentrations that did not affect planktonic cells. In a follow-up report, those authors demonstrated that low doses of peptide 1018 act synergistically with conventional antibiotics to kill a variety of drug-resistant pathogens (104). Based on these two promising examples, it seems reasonable to envision the usefulness of specific (p)ppGpp inhibitors as antipersister drugs in combination therapies.

CONCLUDING REMARKS

About 45 years ago, the discovery of two hyperphosphorylated guanosine nucleotides that were capable of reprogramming cell physiology provided one of the first clues into the complexity of how bacteria utilize alarmone production to regulate a multilayered network controlling bacterial growth and survival. Although much has been learned in the interceding years, several outstanding questions remain. For example, we have just started to understand the biological significance of basal (p)ppGpp under balanced growth conditions. Despite decades of research, many aspects of the mechanism by which RSH enzymes catalyze the conversion of ATP and GTP/GDP to pppGpp/ppGpp are still unknown. Moreover, basic questions as to how and when (pp)pGpp is produced by SASs and how (p)ppGpp mediates persistence in organisms with very different lifestyles are also poorly understood.

As discussed in this review, the ubiquity and critical role of RSH enzymes in bacteria, combined with their absence in mammalian cells, mark the enzymes involved in (p)ppGpp metabolism as potential targets for the development of new antimicrobial strategies. However, an important aspect that should not be overlooked in the development of such antimicrobials is the prevalence of SASs in important bacterial pathogens. Given the essential role of basal (p)ppGpp pools in bacterial virulence and persistence, it will be critical to identify compounds that target both "long" RSHs and SASs, thereby effectively eliminating basal (p)ppGpp production. Another interesting possibility is a reverse approach, i.e., identifying compounds that target (p)ppGpp hydrolase activity, thus locking the bifunctional enzymes (SpoT or Rel) in a synthetic mode that, conceivably, would lead to toxic accumulation of (p)ppGpp. A similar approach that relied on target activation rather than inactivation was recently shown to aid in the elimination of bacterial persisters. Specifically, the acyldepsipeptide (ADEP4) antibiotic has been shown to activate the ClpP protease, leading to uncontrolled proteolysis and ultimately cell death (105). Although *clpP* null mutants arose at a high frequency, combinatorial treatment with ADEP4 and rifampin eradicated biofilm persisters in both *in vitro* and *in vivo* models (106). Regardless of the direction taken to interfere with (p)ppGpp metabolism, a better structural and biochemical understanding of the mechanisms of action of the enzymes responsible for synthesis and degradation of (p)ppGpp can be an invaluable resource in the development of clinically effective antibiotics.

ACKNOWLEDGMENTS

We thank Mike Cashel for helpful discussions and critical reading of the manuscript. We also thank James H. Miller for comments and suggestions to improve the text.

A.O.G. was supported by an NIDCR Training Grant in Oral Science (T90 DE021985).

REFERENCES

1. **Cashel M, Gallant J.** 1969. Two compounds implicated in the function of the RC gene of *Escherichia coli*. Nature **221:**838–841. http://dx.doi.org/10.1038/221838a0.
2. **Cashel M, Kalbacher B.** 1970. The control of ribonucleic acid synthesis in *Escherichia coli*. V. Characterization of a nucleotide associated with the stringent response. J Biol Chem **245:**2309–2318.
3. **Haseltine WA, Block R, Gilbert W, Weber K.** 1972. MSI and MSII made on ribosome in idling step of protein synthesis. Nature **238:**381–384. http://dx.doi.org/10.1038/238381a0.
4. **Cashel M, Gentry DR, Hernandez VJ, Vinella D.** 1996. The stringent response, p 1458–1496. *In* Neidhardt FC, Curtiss R, III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed, vol 1. ASM Press, Washington, DC.
5. **Potrykus K, Cashel M.** 2008. (p) ppGpp: still magical? Annu Rev Microbiol **62:**35–51. http://dx.doi.org/10.1146/annurev.micro.62.081307.162903.
6. **Kanjee U, Ogata K, Houry WA.** 2012. Direct binding targets of the stringent response alarmone (p)ppGpp. Mol Microbiol **85:**1029–1043. http://dx.doi.org/10.1111/j.1365-2958.2012.08177.x.
7. **Dalebroux ZD, Swanson MS.** 2012. ppGpp: magic beyond RNA polymerase. Nat Rev Microbiol **10:**203–212. http://dx.doi.org/10.1038/nrmicro2720.
8. **Ross W, Vrentas CE, Sanchez-Vazquez P, Gaal T, Gourse RL.** 2013. The magic spot: a ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. Mol Cell **50:**420–429. http://dx.doi.org/10.1016/j.molcel.2013.03.021.
9. **Zuo Y, Wang Y, Steitz TA.** 2013. The mechanism of *E. coli* RNA polymerase regulation by ppGpp is suggested by the structure of their complex. Mol Cell **50:**430–436. http://dx.doi.org/10.1016/j.molcel.2013.03.020.
10. **Mechold U, Potrykus K, Murphy H, Murakami KS, Cashel M.** 2013. Differential regulation by ppGpp versus pppGpp in *Escherichia coli*. Nucleic Acids Res **41:**6175–6189. http://dx.doi.org/10.1093/nar/gkt302.
11. **Paul BJ, Berkmen MB, Gourse RL.** 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. Proc Natl Acad Sci U S A **102:**7823–7828. http://dx.doi.org/10.1073/pnas.0501170102.
12. **Wagner R.** 2002. Regulation of ribosomal RNA synthesis in *E. coli*: effects of the global regulator guanosine tetraphosphate (ppGpp). J Mol Microbiol Biotechnol **4:**331–340.
13. **Travers AA.** 1980. Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. J Bacteriol **141:**973–976.
14. **Jishage M, Kvint K, Shingler V, Nystrom T.** 2002. Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev **16:**1260–1270. http://dx.doi.org/10.1101/gad.227902.
15. **Bernardo LM, Johansson LU, Solera D, Skarfstad E, Shingler V.** 2006. The guanosine tetraphosphate (ppGpp) alarmone, DksA and promoter affinity for RNA polymerase in regulation of sigma-dependent transcription. Mol Microbiol **60:**749–764. http://dx.doi.org/10.1111/j.1365-2958.2006.05129.x.
16. **Costanzo A, Ades SE.** 2006. Growth phase-dependent regulation of the extracytoplasmic stress factor, σ^E, by guanosine 3′,5′-bispyrophosphate (ppGpp). J Bacteriol **188:**4627–4634. http://dx.doi.org/10.1128/JB.01981-05.
17. **Sze CC, Shingler V.** 1999. The alarmone (p)ppGpp mediates physiological-responsive control at the σ^{54}-dependent Po promoter. Mol Microbiol **31:**1217–1228. http://dx.doi.org/10.1046/j.1365-2958.1999.01264.x.
18. **Wolz C, Geiger T, Goerke C.** 2010. The synthesis and function of the alarmone (p)ppGpp in Firmicutes. Int J Med Microbiol **300:**142–147. http://dx.doi.org/10.1016/j.ijmm.2009.08.017.
19. **Krasny L, Gourse RL.** 2004. An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. EMBO J **23:**4473–4483. http://dx.doi.org/10.1038/sj.emboj.7600423.
20. **Krasny L, Tiserova H, Jonak J, Rejman D, Sanderova H.** 2008. The identity

of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in *Bacillus subtilis*. Mol Microbiol **69:**42–54. http://dx.doi.org/10.1111/j.1365-2958.2008.06256.x.
21. **Lopez JM, Dromerick A, Freese E.** 1981. Response of guanosine 5′-triphosphate concentration to nutritional changes and its significance for *Bacillus subtilis* sporulation. J Bacteriol **146:**605–613.
22. **Sonenshein AL.** 2005. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr Opin Microbiol **8:**203–207. http://dx.doi.org/10.1016/j.mib.2005.01.001.
23. **Geiger T, Wolz C.** 2014. Intersection of the stringent response and the CodY regulon in low GC Gram-positive bacteria. Int J Med Microbiol **304:**150–155. http://dx.doi.org/10.1016/j.ijmm.2013.11.013.
24. **Inaoka T, Ochi K.** 2002. RelA protein is involved in induction of genetic competence in certain *Bacillus subtilis* strains by moderating the level of intracellular GTP. J Bacteriol **184:**3923–3930. http://dx.doi.org/10.1128/JB.184.14.3923-3930.2002.
25. **Natori Y, Tagami K, Murakami K, Yoshida S, Tanigawa O, Moh Y, Masuda K, Wada T, Suzuki S, Nanamiya H, Tozawa Y, Kawamura F.** 2009. Transcription activity of individual *rrn* operons in *Bacillus subtilis* mutants deficient in (p)ppGpp synthetase genes *relA*, *yjbM*, and *ywaC*. J Bacteriol **191:**4555–4561. http://dx.doi.org/10.1128/JB.00263-09.
26. **Dalebroux ZD, Svensson SL, Gaynor EC, Swanson MS.** 2010. ppGpp conjures bacterial virulence. Microbiol Mol Biol Rev **74:**171–199. http://dx.doi.org/10.1128/MMBR.00046-09.
27. **Boutte CC, Crosson S.** 2013. Bacterial lifestyle shapes stringent response activation. Trends Microbiol **21:**174–180. http://dx.doi.org/10.1016/j.tim.2013.01.002.
28. **Alfoldi L, Stent GS, Clowes RC.** 1962. The chromosomal site of the RNA control (RC) locus in *Escherichia coli*. J Mol Biol **5:**348–355. http://dx.doi.org/10.1016/S0022-2836(62)80077-1.
29. **Wendrich TM, Blaha G, Wilson DN, Marahiel MA, Nierhaus KH.** 2002. Dissection of the mechanism for the stringent factor RelA. Mol Cell **10:**779–788. http://dx.doi.org/10.1016/S1097-2765(02)00656-1.
30. **English BP, Hauryliuk V, Sanamrad A, Tankov S, Dekker NH, Elf J.** 2011. Single-molecule investigations of the stringent response machinery in living bacterial cells. Proc Natl Acad Sci U S A **108:**E365–E373. http://dx.doi.org/10.1073/pnas.1102255108.
31. **Shyp V, Tankov S, Ermakov A, Kudrin P, English BP, Ehrenberg M, Tenson T, Elf J, Hauryliuk V.** 2012. Positive allosteric feedback regulation of the stringent response enzyme RelA by its product. EMBO Rep **13:**835–839. http://dx.doi.org/10.1038/embor.2012.106.
32. **Seyfzadeh M, Keener J, Nomura M.** 1993. spoT-dependent accumulation of guanosine tetraphosphate in response to fatty acid starvation in *Escherichia coli*. Proc Natl Acad Sci U S A **90:**11004–11008. http://dx.doi.org/10.1073/pnas.90.23.11004.
33. **Battesti A, Bouveret E.** 2006. Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. Mol Microbiol **62:**1048–1063. http://dx.doi.org/10.1111/j.1365-2958.2006.05442.x.
34. **Vinella D, Albrecht C, Cashel M, D'Ari R.** 2005. Iron limitation induces SpoT-dependent accumulation of ppGpp in *Escherichia coli*. Mol Microbiol **56:**958–970. http://dx.doi.org/10.1111/j.1365-2958.2005.04601.x.
35. **Jiang M, Sullivan SM, Wout PK, Maddock JR.** 2007. G-protein control of the ribosome-associated stress response protein SpoT. J Bacteriol **189:**6140–6147. http://dx.doi.org/10.1128/JB.00315-07.
36. **Wout P, Pu K, Sullivan SM, Reese V, Zhou S, Lin B, Maddock JR.** 2004. The *Escherichia coli* GTPase CgtAE cofractionates with the 50S ribosomal subunit and interacts with SpoT, a ppGpp synthetase/hydrolase. J Bacteriol **186:**5249–5257. http://dx.doi.org/10.1128/JB.186.16.5249-5257.2004.
37. **Raskin DM, Judson N, Mekalanos JJ.** 2007. Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in *Vibrio cholerae*. Proc Natl Acad Sci U S A **104:**4636–4641. http://dx.doi.org/10.1073/pnas.0611650104.
38. **Mechold U, Cashel M, Steiner K, Gentry D, Malke H.** 1996. Functional analysis of a relA/spoT gene homolog from *Streptococcus equisimilis*J Bacteriol **178:**1401–1411.
39. **Atkinson GC, Tenson T, Hauryliuk V.** 2011. The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. PLoS One **6:**e23479. http://dx.doi.org/10.1371/journal.pone.0023479.
40. **Avarbock D, Avarbock A, Rubin H.** 2000. Differential regulation of opposing RelMtb activities by the aminoacylation state of a tRNA.ribosome.mRNA.RelMtb complex. Biochemistry **39:**11640–11648. http://dx.doi.org/10.1021/bi001256k.
41. **Mechold U, Murphy H, Brown L, Cashel M.** 2002. Intramolecular regulation of the opposing (p)ppGpp catalytic activities of Rel(Seq), the Rel/Spo enzyme from *Streptococcus equisimilis*. J Bacteriol **184:**2878–2888. http://dx.doi.org/10.1128/JB.184.11.2878-2888.2002.
42. **Hogg T, Mechold U, Malke H, Cashel M, Hilgenfeld R.** 2004. Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response. Cell **117:**57–68. http://dx.doi.org/10.1016/S0092-8674(04)00260-0.
43. **Lemos JA, Lin VK, Nascimento MM, Abranches J, Burne RA.** 2007. Three gene products govern (p)ppGpp production by *Streptococcus mutans*. Mol Microbiol **65:**1568–1581. http://dx.doi.org/10.1111/j.1365-2958.2007.05897.x.
44. **Nanamiya H, Kasai K, Nozawa A, Yun CS, Narisawa T, Murakami K, Natori Y, Kawamura F, Tozawa Y.** 2008. Identification and functional analysis of novel (p)ppGpp synthetase genes in *Bacillus subtilis*. Mol Microbiol **67:**291–304. http://dx.doi.org/10.1111/j.1365-2958.2007.06018.x.
45. **Abranches J, Martinez AR, Kajfasz JK, Chavez V, Garsin DA, Lemos JA.** 2009. The molecular alarmone (p) ppGpp mediates stress responses, vancomycin tolerance, and virulence in *Enterococcus faecalis*. J Bacteriol **191:**2248–2256. http://dx.doi.org/10.1128/JB.01726-08.
46. **Das B, Pal RR, Bag S, Bhadra RK.** 2009. Stringent response in *Vibrio cholerae*: genetic analysis of *spoT* gene function and identification of a novel (p)ppGpp synthetase gene. Mol Microbiol **72:**380–398. http://dx.doi.org/10.1111/j.1365-2958.2009.06653.x.
47. **Geiger T, Kastle B, Gratani FL, Goerke C, Wolz C.** 2014. Two small (p)ppGpp synthases in *Staphylococcus aureus* mediate tolerance against cell envelope stress conditions. J Bacteriol **196:**894–902. http://dx.doi.org/10.1128/JB.01201-13.
48. **Srivatsan A, Han Y, Peng J, Tehranchi AK, Gibbs R, Wang JD, Chen R.** 2008. High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. PLoS Genet **4:**e1000139. http://dx.doi.org/10.1371/journal.pgen.1000139.
49. **Weiss LA, Stallings CL.** 2013. Essential roles for *Mycobacterium tuberculosis* Rel beyond the production of (p)ppGpp. J Bacteriol **195:**5629–5638. http://dx.doi.org/10.1128/JB.00759-13.
50. **Bag S, Das B, Dasgupta S, Bhadra RK.** 2014. Mutational analysis of the (p)ppGpp synthetase activity of the Rel enzyme of Mycobacterium tuberculosis. Arch Microbiol **196:**575–588. http://dx.doi.org/10.1007/s00203-014-0996-9.
51. **Murdeshwar MS, Chatterji D.** 2012. MS_RHII-RSD, a dual-function RNase HII-(p)ppGpp synthetase from *Mycobacterium smegmatis*. J Bacteriol **194:**4003–4014. http://dx.doi.org/10.1128/JB.00258-12.
52. **Dasgupta S, Basu P, Pal RR, Bag S, Bhadra RK.** 2014. Genetic and mutational characterization of the small alarmone synthetase gene *relV* of *Vibrio cholerae*. Microbiology **160:**1855–1866. http://dx.doi.org/10.1099/mic.0.079319-0.
53. **Seaton K, Ahn SJ, Sagstetter AM, Burne RA.** 2011. A transcriptional regulator and ABC transporters link stress tolerance, (p)ppGpp, and genetic competence in *Streptococcus mutans*. J Bacteriol **193:**862–874. http://dx.doi.org/10.1128/JB.01257-10.
54. **Gaca AO, Kajfasz JK, Miller JH, Liu K, Wang JD, Abranches J, Lemos JA.** 2013. Basal levels of (p)ppGpp in *Enterococcus faecalis*: the magic beyond the stringent response. mBio **4**(5):e00646-13. http://dx.doi.org/10.1128/mBio.00646-13.
55. **Nishino T, Gallant J, Shalit P, Palmer L, Wehr T.** 1979. Regulatory nucleotides involved in the Rel function of *Bacillus subtilis*. J Bacteriol **140:**671–679.
56. **Sun D, Lee G, Lee JH, Kim HY, Rhee HW, Park SY, Kim KJ, Kim Y, Kim BY, Hong JI, Park C, Choy HE, Kim JH, Jeon YH, Chung J.** 2010. A metazoan ortholog of SpoT hydrolyzes ppGpp and functions in starvation responses. Nat Struct Mol Biol **17:**1188–1194. http://dx.doi.org/10.1038/nsmb.1906.
57. **Gross M, Marianovsky I, Glaser G.** 2006. MazG: a regulator of programmed cell death in *Escherichia coli*. Mol Microbiol **59:**590–601. http://dx.doi.org/10.1111/j.1365-2958.2005.04956.x.
58. **Ooga T, Ohashi Y, Kuramitsu S, Koyama Y, Tomita M, Soga T, Masui R.** 2009. Degradation of ppGpp by nudix pyrophosphatase modulates the transition of growth phase in the bacterium *Thermus thermophilus*. J Biol Chem **284:**15549–15556. http://dx.doi.org/10.1074/jbc.M900582200.
59. **Sokawa Y, Sokawa J, Kaziro Y.** 1975. Regulation of stable RNA synthesis

and ppGpp levels in growing cells of *Escherichia coli*. Cell **5:**69–74. http://dx.doi.org/10.1016/0092-8674(75)90093-8.
60. **Sarubbi E, Rudd KE, Cashel M.** 1988. Basal ppGpp level adjustment shown by new *spoT* mutants affect steady state growth rates and *rrnA* ribosomal promoter regulation in *Escherichia coli*. Mol Gen Genet **213:**214–222. http://dx.doi.org/10.1007/BF00339584.
61. **Lazzarini RA, Cashel M, Gallant J.** 1971. On the regulation of guanosine tetraphosphate levels in stringent and relaxed strains of *Escherichia coli*. J Biol Chem **246:**4381–4385.
62. **Ryals J, Little R, Bremer H.** 1982. Control of rRNA and tRNA syntheses in *Escherichia coli* by guanosine tetraphosphate. J Bacteriol **151:**1261–1268.
63. **Potrykus K, Murphy H, Philippe N, Cashel M.** 2011. ppGpp is the major source of growth rate control in *E. coli*. Environ Microbiol **13:**563–575. http://dx.doi.org/10.1111/j.1462-2920.2010.02357.x.
64. **Traxler MF, Zacharia VM, Marquardt S, Summers SM, Nguyen HT, Stark SE, Conway T.** 2011. Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in *Escherichia coli*. Mol Microbiol **79:**830–845. http://dx.doi.org/10.1111/j.1365-2958.2010.07498.x.
65. **Cho BK, Barrett CL, Knight EM, Park YS, Palsson BO.** 2008. Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. Proc Natl Acad Sci U S A **105:**19462–19467. http://dx.doi.org/10.1073/pnas.0807227105.
66. **Gaca AO, Abranches J, Kajfasz JK, Lemos JA.** 2012. Global transcriptional analysis of the stringent response in *Enterococcus faecalis*. Microbiology **158:**1994–2004. http://dx.doi.org/10.1099/mic.0.060236-0.
67. **Kriel A, Bittner AN, Kim SH, Liu K, Tehranchi AK, Zou WY, Rendon S, Chen R, Tu BP, Wang JD.** 2012. Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance. Mol Cell **48:**231–241. http://dx.doi.org/10.1016/j.molcel.2012.08.009.
68. **Kriel A, Brinsmade SR, Tse JL, Tehranchi AK, Bittner AN, Sonenshein AL, Wang JD.** 2014. GTP dysregulation in *Bacillus subtilis* cells lacking (p)ppGpp results in phenotypic amino acid auxotrophy and failure to adapt to nutrient downshift and regulate biosynthesis genes. J Bacteriol **196:**189–201. http://dx.doi.org/10.1128/JB.00918-13.
69. **Hochstadt-Ozer J, Cashel M.** 1972. The regulation of purine utilization in bacteria. V. Inhibition of purine phosphoribosyltransferase activities and purine uptake in isolated membrane vesicles by guanosine tetraphosphate. J Biol Chem **247:**7067–7072.
70. **Pall ML.** 1985. GTP: a central regulator of cellular anabolism. Curr Top Cell Regul **25:**1–20. http://dx.doi.org/10.1016/B978-0-12-152825-6.50005-9.
71. **Sonenshein AL.** 2007. Control of key metabolic intersections in *Bacillus subtilis*. Nat Rev Microbiol **5:**917–927. http://dx.doi.org/10.1038/nrmicro1772.
72. **Bennett HJ, Pearce DM, Glenn S, Taylor CM, Kuhn M, Sonenshein AL, Andrew PW, Roberts IS.** 2007. Characterization of *relA* and *codY* mutants of *Listeria monocytogenes*: identification of the CodY regulon and its role in virulence. Mol Microbiol **63:**1453–1467. http://dx.doi.org/10.1111/j.1365-2958.2007.05597.x.
73. **Lemos JA, Nascimento MM, Lin VK, Abranches J, Burne RA.** 2008. Global regulation by (p)ppGpp and CodY in *Streptococcus mutans*. J Bacteriol **190:**5291–5299. http://dx.doi.org/10.1128/JB.00288-08.
74. **Geiger T, Goerke C, Fritz M, Schafer T, Ohlsen K, Liebeke M, Lalk M, Wolz C.** 2010. Role of the (p)ppGpp synthase RSH, a RelA/SpoT homolog, in stringent response and virulence of *Staphylococcus aureus*. Infect Immun **78:**1873–1883. http://dx.doi.org/10.1128/IAI.01439-09.
75. **Dineen SS, McBride SM, Sonenshein AL.** 2010. Integration of metabolism and virulence by *Clostridium difficile* CodY. J Bacteriol **192:**5350–5362. http://dx.doi.org/10.1128/JB.00341-10.
76. **Montgomery CP, Boyle-Vavra S, Roux A, Ebine K, Sonenshein AL, Daum RS.** 2012. CodY deletion enhances in vivo virulence of community-associated methicillin-resistant *Staphylococcus aureus* clone USA300. Infect Immun **80:**2382–2389. http://dx.doi.org/10.1128/IAI.06172-11.
77. **Chateau A, van Schaik W, Joseph P, Handke LD, McBride SM, Smeets FM, Sonenshein AL, Fouet A.** 2013. Identification of CodY targets in *Bacillus anthracis* by genome-wide in vitro binding analysis. J Bacteriol **195:**1204–1213. http://dx.doi.org/10.1128/JB.02041-12.
78. **Petranovic D, Guedon E, Sperandio B, Delorme C, Ehrlich D, Renault P.** 2004. Intracellular effectors regulating the activity of the *Lactococcus lactis* CodY pleiotropic transcription regulator. Mol Microbiol **53:**613–621. http://dx.doi.org/10.1111/j.1365-2958.2004.04136.x.
79. **Hendriksen WT, Bootsma HJ, Estevao S, Hoogenboezem T, de Jong A, de Groot R, Kuipers OP, Hermans PW.** 2008. CodY of *Streptococcus pneumoniae*: link between nutritional gene regulation and colonization. J Bacteriol **190:**590–601. http://dx.doi.org/10.1128/JB.00917-07.
80. **Lewis K.** 2007. Persister cells, dormancy and infectious disease. Nat Rev Microbiol **5:**48–56. http://dx.doi.org/10.1038/nrmicro1557.
81. **Kusser W, Ishiguro EE.** 1987. Suppression of mutations conferring penicillin tolerance by interference with the stringent control mechanism of *Escherichia coli*. J Bacteriol **169:**4396–4398.
82. **Rodionov DG, Ishiguro EE.** 1995. Direct correlation between overproduction of guanosine 3′,5′-bispyrophosphate (ppGpp) and penicillin tolerance in *Escherichia coli*. J Bacteriol **177:**4224–4229.
83. **Joseleau-Petit D, Thevenet D, D'Ari R.** 1994. ppGpp concentration, growth without PBP2 activity, and growth-rate control in *Escherichia coli*. Mol Microbiol **13:**911–917. http://dx.doi.org/10.1111/j.1365-2958.1994.tb00482.x.
84. **Tuomanen E, Cozens R, Tosch W, Zak O, Tomasz A.** 1986. The rate of killing of *Escherichia coli* by beta-lactam antibiotics is strictly proportional to the rate of bacterial growth. J Gen Microbiol **132:**1297–1304.
85. **Amato SM, Orman MA, Brynildsen MP.** 2013. Metabolic control of persister formation in *Escherichia coli*. Mol Cell **50:**475–487. http://dx.doi.org/10.1016/j.molcel.2013.04.002.
86. **Maisonneuve E, Gerdes K.** 2014. Molecular mechanisms underlying bacterial persisters. Cell **157:**539–548. http://dx.doi.org/10.1016/j.cell.2014.02.050.
87. **Mwangi MM, Kim C, Chung M, Tsai J, Vijayadamodar G, Benitez M, Jarvie TP, Du L, Tomasz A.** 2013. Whole-genome sequencing reveals a link between beta-lactam resistance and synthetases of the alarmone (p)ppGpp in *Staphylococcus aureus*. Microb Drug Resist **19:**153–159. http://dx.doi.org/10.1089/mdr.2013.0053.
88. **Kim C, Mwangi M, Chung M, Milheirico C, de Lencastre H, Tomasz A.** 2013. The mechanism of heterogeneous beta-lactam resistance in MRSA: key role of the stringent stress response. PLoS One **8:**e82814. http://dx.doi.org/10.1371/journal.pone.0082814.
89. **Maisonneuve E, Castro-Camargo M, Gerdes K.** 2013. (p) ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. Cell **154:**1140–1150. http://dx.doi.org/10.1016/j.cell.2013.07.048.
90. **Moyed HS, Bertrand KP.** 1983. *hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. J Bacteriol **155:**768–775.
91. **Kaspy I, Rotem E, Weiss N, Ronin I, Balaban NQ, Glaser G.** 2013. HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. Nat Commun **4:**3001. http://dx.doi.org/10.1038/ncomms4001.
92. **Germain E, Castro-Roa D, Zenkin N, Gerdes K.** 2013. Molecular mechanism of bacterial persistence by HipA. Mol Cell **52:**248–254. http://dx.doi.org/10.1016/j.molcel.2013.08.045.
93. **Slattery A, Victorsen AH, Brown A, Hillman K, Phillips GJ.** 2013. Isolation of highly persistent mutants of *Salmonella enterica* serovar Typhimurium reveals a new toxin-antitoxin module. J Bacteriol **195:**647–657. http://dx.doi.org/10.1128/JB.01397-12.
94. **Claudi B, Sprote P, Chirkova A, Personnic N, Zankl J, Schurmann N, Schmidt A, Bumann D.** 2014. Phenotypic variation of Salmonella in host tissues delays eradication by antimicrobial chemotherapy. Cell **158:**722–733. http://dx.doi.org/10.1016/j.cell.2014.06.045.
95. **Helaine S, Cheverton AM, Watson KG, Faure LM, Matthews SA, Holden DW.** 2014. Internalization of Salmonella by macrophages induces formation of nonreplicating persisters. Science **343:**204–208. http://dx.doi.org/10.1126/science.1244705.
96. **Nguyen D, Joshi-Datar A, Lepine F, Bauerle E, Olakanmi O, Beer K, McKay G, Siehnel R, Schafhauser J, Wang Y, Britigan BE, Singh PK.** 2011. Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. Science **334:**982–986. http://dx.doi.org/10.1126/science.1211037.
97. **Khakimova M, Ahlgren HG, Harrison JJ, English AM, Nguyen D.** 2013. The stringent response controls catalases in *Pseudomonas aeruginosa* and is required for hydrogen peroxide and antibiotic tolerance. J Bacteriol **195:**2011–2020. http://dx.doi.org/10.1128/JB.02061-12.
98. **Gao W, Chua K, Davies JK, Newton HJ, Seemann T, Harrison PF, Holmes NE, Rhee HW, Hong JI, Hartland EL, Stinear TP, Howden BP.** 2010. Two novel point mutations in clinical *Staphylococcus aureus* reduce linezolid susceptibility and switch on the stringent response to

promote persistent infection. PLoS Pathog **6:**e1000944. http://dx.doi.org/10.1371/journal.ppat.1000944.
99. **Nascimento MM, Lemos JA, Abranches J, Lin VK, Burne RA.** 2008. Role of RelA of *Streptococcus mutans* in global control of gene expression. J Bacteriol **190:**28–36. http://dx.doi.org/10.1128/JB.01395-07.
100. **Kazmierczak KM, Wayne KJ, Rechtsteiner A, Winkler ME.** 2009. Roles of rel(Spn) in stringent response, global regulation and virulence of serotype 2 *Streptococcus pneumoniae* D39. Mol Microbiol **72:**590–611. http://dx.doi.org/10.1111/j.1365-2958.2009.06669.x.
101. **Eymann C, Homuth G, Scharf C, Hecker M.** 2002. *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. J Bacteriol **184:**2500–2520. http://dx.doi.org/10.1128/JB.184.9.2500-2520.2002.
102. **Wexselblatt E, Oppenheimer-Shaanan Y, Kaspy I, London N, Schueler-Furman O, Yavin E, Glaser G, Katzhendler J, Ben-Yehuda S.** 2012. Relacin, a novel antibacterial agent targeting the stringent response. PLoS Pathog **8:**e1002925. http://dx.doi.org/10.1371/journal.ppat.1002925.
103. **de la Fuente-Nunez C, Reffuveille F, Haney EF, Straus SK, Hancock RE.** 2014. Broad-spectrum anti-biofilm peptide that targets a cellular stress response. PLoS Pathog **10:**e1004152. http://dx.doi.org/10.1371/journal.ppat.1004152.
104. **Reffuveille F, de la Fuente-Nunez C, Mansour S, Hancock RE.** 2014. A broad-spectrum antibiofilm peptide enhances antibiotic action against bacterial biofilms. Antimicrob Agents Chemother **58:**5363–5371. http://dx.doi.org/10.1128/AAC.03163-14.
105. **Brotz-Oesterhelt H, Beyer D, Kroll HP, Endermann R, Ladel C, Schroeder W, Hinzen B, Raddatz S, Paulsen H, Henninger K, Bandow JE, Sahl HG, Labischinski H.** 2005. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. Nat Med **11:**1082–1087. http://dx.doi.org/10.1038/nm1306.
106. **Conlon BP, Nakayasu ES, Fleck LE, LaFleur MD, Isabella VM, Coleman K, Leonard SN, Smith RD, Adkins JN, Lewis K.** 2013. Activated ClpP kills persisters and eradicates a chronic biofilm infection. Nature **503:**365–370. http://dx.doi.org/10.1038/nature12790.
107. **Stayton MM, Fromm HJ.** 1979. Guanosine 5′-diphosphate-3′-diphosphate inhibition of adenylosuccinate synthetase. J Biol Chem **254:**2579–2581.
108. **Milon P, Tischenko E, Tomsic J, Caserta E, Folkers G, La Teana A, Rodnina MV, Pon CL, Boelens R, Gualerzi CO.** 2006. The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. Proc Natl Acad Sci U S A **103:**13962–13967. http://dx.doi.org/10.1073/pnas.0606384103.
109. **Mitkevich VA, Ermakov A, Kulikova AA, Tankov S, Shyp V, Soosaar A, Tenson T, Makarov AA, Ehrenberg M, Hauryliuk V.** 2010. Thermodynamic characterization of ppGpp binding to EF-G or IF2 and of initiator tRNA binding to free IF2 in the presence of GDP, GTP, or ppGpp. J Mol Biol **402:**838–846. http://dx.doi.org/10.1016/j.jmb.2010.08.016.
110. **Feng B, Mandava CS, Guo Q, Wang J, Cao W, Li N, Zhang Y, Zhang Y, Wang Z, Wu J, Sanyal S, Lei J, Gao N.** 2014. Structural and functional insights into the mode of action of a universally conserved Obg GTPase. PLoS Biol **12:**e1001866. http://dx.doi.org/10.1371/journal.pbio.1001866.
111. **Edwards AN, Patterson-Fortin LM, Vakulskas CA, Mercante JW, Potrykus K, Vinella D, Camacho MI, Fields JA, Thompson SA, Georgellis D, Cashel M, Babitzke P, Romeo T.** 2011. Circuitry linking the Csr and stringent response global regulatory systems. Mol Microbiol **80:**1561–1580. http://dx.doi.org/10.1111/j.1365-2958.2011.07663.x.
112. **Buglino J, Shen V, Hakimian P, Lima CD.** 2002. Structural and biochemical analysis of the Obg GTP binding protein. Structure **10:**1581–1592. http://dx.doi.org/10.1016/S0969-2126(02)00882-1.
113. **Denapoli J, Tehranchi AK, Wang JD.** 2013. Dose-dependent reduction of replication elongation rate by (p)ppGpp in *Escherichia coli* and *Bacillus subtilis*. Mol Microbiol **88:**93–104. http://dx.doi.org/10.1111/mmi.12172.
114. **Maciag M, Kochanowska M, Lyzen R, Wegrzyn G, Szalewska-Palasz A.** 2010. ppGpp inhibits the activity of *Escherichia coli* DnaG primase. Plasmid **63:**61–67. http://dx.doi.org/10.1016/j.plasmid.2009.11.002.
115. **Wang JD, Sanders GM, Grossman AD.** 2007. Nutritional control of elongation of DNA replication by (p)ppGpp. Cell **128:**865–875. http://dx.doi.org/10.1016/j.cell.2006.12.043.
116. **An G, Justesen J, Watson RJ, Friesen JD.** 1979. Cloning the *spoT* gene of *Escherichia coli*: identification of the *spoT* gene product. J Bacteriol **137:**1100–1110.
117. **Laffler T, Gallant J.** 1974. *spoT*, a new genetic locus involved in the stringent response in *E. coli*. Cell **1:**27–30. http://dx.doi.org/10.1016/0092-8674(74)90151-2.
118. **Mechold U, Malke H.** 1997. Characterization of the stringent and relaxed responses of *Streptococcus equisimilis*. J Bacteriol **179:**2658–2667.
119. **Bittner AN, Kriel A, Wang JD.** 2014. Lowering GTP level increases survival of amino acid starvation but slows growth rate for *Bacillus subtilis* cells lacking (p)ppGpp. J Bacteriol **196:**2067–2076. http://dx.doi.org/10.1128/JB.01471-14.

Anthony O. Gaca was raised in Grand Rapids, MI, and completed his B.S. in Microbiology at Michigan State University in 2008. While at Michigan State, he contributed to research characterizing metabolic pathways and nutritional requirements for the bioremediation of phenolic compounds and heavy metals. In 2013, Anthony completed his Ph.D. at the University of Rochester under the guidance of José Lemos. His doctoral research focused on defining the physiological contributions and transcriptional hierarchy under the control of (p)ppGpp in the opportunistic pathogen *Enterococcus faecalis*. Currently, he is characterizing the enzymatic function of small (p)ppGpp synthetases in *E. faecalis* and other *Firmicutes*. He began working in the area of bacterial stress tolerance to blend his bioremediation experience with his interest in bacterial pathogenesis.

Cristina Colomer-Winter is a Ph.D. student at the University of Rochester. She was raised in Barcelona, Spain. In 2009, she received a B.S. in Biology at the Universitat Autònoma de Barcelona, where she completed several internships at Boehringer Ingelheim (Germany) and at the Universität Würzburg. Later on, she completed her M.Sc. degree in Microbiology at the Universitat de Barcelona under the mentorship of Jaume Jané at B. Braun Medical (Rubí, Spain), characterizing the heat resistance of *Bacillus* spp. spores present in the bioburden. After graduating, she continued working for 3 years at the R&D Department of B. Braun, developing antibacterial medical devices and novel applications for tissue regeneration. She joined the lab of José Lemos in the summer of 2014 to start work on the interaction of (p)ppGpp, GTP, and CodY in *Enterococcus faecalis* pathophysiology.

Continued next page

José A. Lemos was raised in Rio de Janeiro, Brazil, and obtained his Ph.D. in Microbiology and Immunology from the Federal University of Rio de Janeiro in 2000, working with Angela Castro. He did postdoctoral research with Robert Burne at the University of Rochester and later at the University of Florida, where he initiated his studies on the stress response mechanisms of *Firmicutes*. Around this time, he became interested in the (p)ppGpp-mediated bacterial stringent response while investigating the regulation of toxin-antitoxin modules in *Streptococcus mutans*. In 2002, he became a research assistant professor in the Department of Oral Biology at the University of Florida. In 2007, he joined the faculty of the University of Rochester in the Center for Oral Biology, where he is currently an associate professor of microbiology and immunology. His laboratory investigates the stress survival mechanisms of streptococci and enterococci.