Evidence that metabolism and chromosome copy number control mutually exclusive cell fates in *Bacillus subtilis*

Yunrong Chai\(^1\), Thomas Norman\(^1\), Roberto Kolter\(^2\) and Richard Losick\(^1,*\)

\(^1\)Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA and \(^2\)Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA, USA

*Bacillus subtilis* chooses between matrix production and spore formation, which are both controlled by the regulator Spo0A. We report that metabolism and chromosome copy number dictate which fate is adopted. Conditions that favour low Spo0A–P levels promote matrix production, whereas conditions favouring high levels trigger sporulation. Spo0A–P directs the synthesis of SinI, an antirepressor for the SinR repressor of matrix genes. The regulatory region of *sinI* contains an activator site that Spo0A–P binds strongly and operators that bind Spo0A–P weakly. Evidence shows that low Spo0A–P levels turn *sinI* ON and high levels turn *sinI* OFF and instead switch sporulation ON. Cells in which *sinI* and *sinR* were transplanted from their normal position near the chromosome replication terminus to positions near the origin and cells that harboured an extra copy of the genes were blocked in matrix production. Thus, matrix gene expression is sensitive to the number of copies of *sinI* and *sinR*. Because cells at the start of sporulation have two chromosomes and matrix-producing cells one, chromosome copy number could contribute to cell-fate determination.

The EMBO Journal (2011) 30, 1402–1413. doi:10.1038/emboj.2011.36; Published online 15 February 2011

Subject Categories: chromatin & transcription; differentiation & death; microbiology & pathogens

Keywords: *Bacillus subtilis*; Biofilm; chromosome copy number; metabolic control

Introduction

During development, genetically identical cells differentiate into a variety of specialized cell types. Elucidating the mechanisms that control the formation of alternative cell fates is a fundamental challenge in developmental biology. An attractive organism for addressing questions of how cells specialize is the bacterium *Bacillus subtilis*, which is able to adopt several distinct cell fates (Vlamakis *et al*., 2008; López *et al*., 2009). For example, in nutrient replete conditions, *B. subtilis* is chiefly in the form of single, motile cells. However, under conditions of mild nutrient depletion, cells differentiate into long chains of matrix-producing cells that assemble into multicellular communities known as biofilms (Vlamakis *et al*., 2008; Chai *et al*., 2010). Under more extreme conditions of nutrient limitation, cells enter an alternative developmental pathway in which they differentiate into dormant spores (Sonenshein, 2000). Spores also arise in biofilms as they age, with matrix-producing cells differentiating into spore-forming cells (Branda *et al*., 2001; Vlamakis *et al*., 2008).

Both spore formation and biofilm formation are governed by the same regulatory protein Spo0A (hereafter 0A–P) (Branda *et al*., 2001; Hamon and Lazazzera, 2001; Piggot and Hilbert, 2004; Fujita and Losick, 2005). 0A–P directly activates genes and operons (spolIA, spolIE, and spolIG) that govern entry into sporulation (Molle *et al*., 2003; Fujita *et al*., 2005). In addition, 0A–P indirectly triggers biofilm formation by turning ON the gene for SinI (Shafikhani *et al*., 2002; Fujita *et al*., 2005). SinI is an antirepressor that antagonizes the repressor SinR (Bai *et al*., 1993; Lewis *et al*., 1996; Kearns *et al*., 2005). SinR, in turn, binds to multiple tandem sites in the regulatory regions of the *epsA-O* operon (henceforth *eps*), which is responsible for the production of exopolysaccharide (Kearns *et al*., 2005), and the *yqX-M-sipW-tasA* operon (henceforth *yqXM*), which encodes an amyloid-like fibre component of the matrix (Brandt *et al*., 2006; Chu *et al*., 2006; Romero *et al*., 2010). SinR also binds to the gene for an additional regulatory protein SlrR (Chu *et al*., 2008; Kobayashi, 2008). Evidence indicates that SinR and SlrR constitute a self-reinforcing double-negative loop that locks cells in the matrix-producing state (Chai *et al*., 2010). Thus, 0A–P-directed synthesis of SinI sets in motion a regulatory circuit that derepresses matrix operons.

A key question is how the same regulator, 0A–P, can control the formation of two distinct cell types, matrix producers and sporulating cells. Part of the answer is that 0A–P-activated genes and operons for sporulation have weak binding sites for the phosphoprotein and are only activated under nutritional conditions in which 0A–P reaches high levels (Molle *et al*., 2003; Fujita *et al*., 2005). Biofilm formation, in contrast, seems to be favoured under conditions in which 0A–P reaches low-to-intermediate levels (Fujita *et al*., 2005; Chai *et al*., 2008; Aguilar *et al*., 2010). These findings still leave important gaps in our understanding. Why does biofilm formation not take place under conditions in which 0A–P reaches high levels and why do sporulating cells not produce matrix? Here, we report evidence for a mechanism tied to the metabolic state of cells that differentially favours biofilm formation or sporulation and a chromosome copy number mechanism that prevents cells that have entered the sporulation pathway from expressing matrix genes.
Results

*sinI* contains putative 0A–P binding sites upstream and downstream of the transcription start site

We observed that the promoter region for *sinI* contains five potential binding sites for 0A–P located upstream and downstream of the transcription start site. The consensus-binding site for 0A–P is the seven-basepair sequence `TTCGACA`, which is known as the 0A-box (Molle et al., 2003). Just upstream of the −35 region of the promoter is a perfect match to the 0A-box (Figure 1), which was previously shown by footprinting experiments to be a binding site for the response regulator (Shafikhani et al., 2002). In addition, genetic evidence presented here demonstrates that the box is required for *sinI* transcription. Thus, the upstream 0A-box is an activation site at which the binding of 0A–P turns ON the *sinI* gene.

Downstream of the activation site are four sites that are imperfect matches to an 0A-box (designated as O1 through O4), each being a five out of seven match to the consensus (Figure 1). The O1 site is located between the −35 (TCGTTT) and −10 regions (TATAAT) of the promoter (Shafikhani et al., 2002). The remaining sites are downstream of the start site. Underscoring the potential significance of these sites, orthologs of *sinI* from *Bacillus amyloliquefaciens* and *Bacillus licheniformis* similarly exhibit 0A-like boxes upstream and downstream of their inferred transcription start sites (Figure 1) (Rey et al., 2004; Veith et al., 2004; Chen et al., 2007).

The positioning of the four 0A-like boxes within the region of DNA that would be covered by promoter-bound RNA polymerase suggests that the sites are operators at which the binding of 0A–P represses *sinI*. However, because they are imperfect matches to the consensus, the four sites are likely weaker binding sites than the upstream activation site.

We therefore hypothesized that at low-to-intermediate levels of 0A–P, the phosphoprotein stimulates transcription of *sinI* whereas at higher concentrations 0A–P represses *sinI*.

*sinI* expression is sustained in biofilm but not sporulation medium

Cells behave differently when grown in Difco sporulation (DS) or in MSgg media. DS medium promotes high levels of sporulation but little multicellular communities (Chai et al., 2008). Conversely, MSgg medium promotes the formation of elaborate multicellular communities but relatively little sporulation, except after prolonged incubation (Brandt et al., 2001). We hypothesize that in DS medium, 0A–P rapidly reaches a high level at which continued *sinI* expression is curtailed. Conversely, in MSgg medium, 0A–P persists at intermediate levels that are inadequate for efficient sporulation (which requires high 0A–P levels) but would allow sustained expression of *sinI* and hence biofilm formation.

To test our hypothesis, we monitored *sinI* expression in real time using a luciferase reporter system. Because light emission ceases rapidly when transcription of the luciferase gene is turned off, the luciferase system reports on the switching OFF, as well as the switching ON, of gene expression (Stewart and Williams, 1992). Accordingly, a *P_{scaA}-lux* transcriptional fusion was integrated into the chromosome at *scaA*, and cells harbouring the fusion were grown in DS or MSgg media. Expression of the fusion rose quickly in DS medium, reaching a peak at a time corresponding to the start of sporulation, and then rapidly declining (Supplementary Figure S1A). Under the same conditions, other promoters that are under 0A–P control, but which require high levels of 0A–P to be active, reach a peak level of expression 1–2 h after the start of sporulation (A Chastanet, personal observations).

Figure 1 The regulatory region of *sinI* contains both the activator and operator sites for 0A–P. Shown are nucleotide sequences of the regulatory regions of *sinI* in *B. subtilis, B. amyloliquefaciens*, and *B. licheniformis*. In *B. subtilis*, the −35 and −10 motifs of the ε5-dependent promoter (shown in lactic letters), the transcription start of *sinI* (indicated with arrow), the consensus 0A-box, and four imperfect 0A-boxes (underlined and noted as 0A-box, O1, O2, O3, and O4) were annotated accordingly. In *B. amyloliquefaciens* and *B. licheniformis*, the predicted −35 and −10 motifs of the ε5-dependent promoter were also shown in lactic letters. A consensus 0A-box was identified immediately upstream of the predicted −35 region in both promoters whereas five and three imperfect 0A-boxes (underlined) respectively, were also identified in the *sinI* promoters in *B. amyloliquefaciens* and *B. licheniformis*. The matched sequences in the imperfect 0A-boxes to the consensus 0A-box (`TTCGACA`) were shown in capital letters.
In contrast, in MSgg medium, expression of \( P_{\text{sinI}} \) continued to rise and was maintained at a high level for many hours (Supplementary Figure S1B). These results are consistent with the view that in DS medium, \( 0A \rightarrow P \) rapidly reaches a high level at which \( \text{sinI} \) is turned OFF, whereas in MSgg, \( 0A \rightarrow P \) reaches, and remains at, intermediate levels at which expression of \( \text{sinI} \) persists.

Mutations in the putative operators of \( 0A \rightarrow P \) derepress expression of \( \text{sinI} \)

To test whether the \( 0A \)-like boxes in \( \text{sinI} \) are operators, we introduced point mutations into each of the four sites. We refer to the mutations as M1–M4, and their altered sequences are shown in Figure 2A. We then created mutants that contained various combinations of the mutations. We also created a 43-bp deletion that removed \( O2 \) through \( O4 \) (\( D_{2-4} \), Figure 2A) as well as a scrambled sequence mutation that did not change the length of the \( O2 \) through \( O4 \) region but altered its sequence (see Materials and methods). Finally, we created a mutant of the upstream \( 0A \)-box (\( 0A_{\text{mut}} \), Figure 2A).

The effects of these mutations were analysed with the use of \( \text{lacZ} \) transcriptional fusions to the wild-type and mutant promoter regions of \( \text{sinI} \) (\( P_{\text{sinI-lacZ}} \), which were integrated into the chromosome of strain 3610 at \( amyE \). The results show that mutation of the upstream \( 0A \)-box severely impaired expression of \( P_{\text{sinI-lacZ}} \) expression in MSgg medium in which the highest levels of \( P_{\text{sinI-lacZ}} \) expression are normally seen (Figure 2B, left panel). Thus, the upstream \( 0A \)-box is an activation site.

To analyse the effects of the putative operator mutations, we used DS medium, which, as we have seen, curtails expression of \( \text{sinI} \). Mutants of either \( O1 \) alone or \( O2 \) alone exhibited higher expression levels than that observed for the wild type, whereas mutants of \( O3 \) and \( O4 \) had no detectable effect (Figure 2B, middle panel). The highest levels of derepression were observed with the deletion mutation that removed \( O2 \) through \( O4 \) (\( D_{2-4} \); \( 10 \)-fold; open triangles in the right panel, Figure 2B) and with the scrambled sequence mutant (approximately six-fold; Figure 2B, stars in the right panel). In contrast to these results, neither the deletion mutation nor the scrambled sequence mutation increased \( P_{\text{sinI-lacZ}} \) expression strongly in MSgg medium (Supplementary Figure S5). In toto, these results are consistent with the view that the \( 0A \)-like boxes are operator sites that repress
sinI expression under conditions in which 0A-P reaches high levels but not under conditions in which 0A-P is maintained at intermediate levels.

**Operator mutations enhance pellicle formation and matrix gene expression in DS medium**

We wondered whether enhanced expression of sinI caused by the operator mutations would promote biofilm formation in DS medium. To answer this question, we introduced mutations into the operator sites of the intact sinI gene located at its native locus (see Materials and methods). In brief, we generated a DNA fragment that contained a kanamycin resistance gene (kan) flanked by P_sinI-sinI sinR on one side and the tasA gene on the other (Supplementary Figure S2A). Transformants in which the DNA had been introduced into the homologous region of the chromosome by double recombination were obtained by selection for drug resistance. This strategy allowed us to replace the P_sinI-sinI sinR tasA chromosomal region with DNA that had been manipulated in vitro. Using this strategy, we first generated a strain (YC285) in which the kan gene was inserted between sinR and tasA without altering the sinI promoter (Supplementary Figure S2A). YC285 is essentially indistinguishable from the wild type (3610) in its ability to form biofilms on MSgg medium (data not shown), indicating that kan insertion had little or no effect on biofilm formation.

Next, and using the same strategy, we created five mutant strains in which the native sinI gene was replaced with copies of the gene harbouring various operator mutations. We then compared pellicle formation by these mutant strains with that of the YC285 strain in DS medium. Mutant strains harbouring a deletion of O2 through O4 (Δ2–4, Supplementary Figure S2B) or base substitution mutations in all four operators (M1–4, Supplementary Figure S2B) formed much more robust pellicles than did the YC285 reference strain containing unaltered sinI (WT, Supplementary Figure S2B). Strains harbouring mutations in O1 or O2 exhibited a mild increase in their capacity to form pellicles (M1 and M2, Supplementary Figure S2B). Finally, no significant enhancement in pellicle formation was observed for a mutant strain containing mutations in both O3 and O4 (M3, 4, Supplementary Figure S2B). In contrast to the results obtained in DS medium, and consistent with expectations, none of the mutants exhibited a significantly increased capacity to form pellicles in MSgg medium (data not shown).

Finally, we asked whether the operator mutations were causing increased expression of the eps and yqxM matrix operons in DS medium. To address this question, we introduced transcriptional fusions to the promoters for the two matrix operons (P_espA-lacZ and P_yqxM-lacZ) into the amyE gene of some mutant strains described above as well as into the reference strain YC285. The results largely paralleled what we had observed for pellicle formation (Supplementary Figure S2C). That is, expression of both fusions was significantly higher in the deletion mutant (Δ2–4, open triangles), the M1 mutant (open circles), and the M2 mutant (stars) than in fusion-bearing derivatives of YC285 (filled diamonds). Meanwhile, little or no increase in expression was observed for the M3, 4 double-operator mutant (open squares in Supplementary Figure S2C).

**The DNA-binding domain of 0A adheres to the 0A-like boxes**

Finally, we asked whether the 0A-like boxes are binding sites for 0A. To address this question, we employed both electrophoretic mobility shift assays (EMSAs) and DNase I footprinting. Because full-length 0A requires phosphorylation to bind to DNA efficiently (Ireton et al., 1993; Jiang et al., 2000), we used a truncated form of 0A containing its C-terminal, DNA-binding domain (aa143–267, referred to as 0A-CTD), which is known to bind bona fide targets of 0A independently of phosphorylation (Molle et al., 2003). As demonstrated in Figure 3A, when a DNA fragment containing the wild-type sinI promoter was used as the probe and incubated with increasing concentrations of 0A-CTD, we observed multiple shifted bands. Among them was one that corresponded to a...
low molecular weight complex and appeared at low concentrations of 0A-CTD, and others that corresponded to higher molecular weight complexes and appeared at higher concentrations of 0A-CTD (Figure 3A, panel a). These results suggest that 0A-CTD binds to multiple sites in the sinl promoter region.

Next, we tested binding to DNA fragments harbouring mutations of the 0A-like boxes. When DNA containing a mutated upstream 0A-box (0A\textsuperscript{mut}) was used as probe, we still observed multiple shifted bands (Figure 3A, panel b). However, the band corresponding to the low molecular weight complex in panel a was much reduced (Figure 3A, panel b). This finding suggests that the low molecular weight complex was chiefly formed by the binding of 0A-CTD to the upstream activation site. It also indicates that 0A-CTD was binding to additional sites in the promoter region. We then used as a probe a doubly mutated DNA carrying a mutation of the upstream 0A-box and a deletion of O2 through O4 (0A\textsuperscript{mut} Δ2-4; Figure 3A, panel c). This time, no obvious shifted complex was observed even at the highest concentrations of 0A-CTD (Figure 3A, panel c). This finding suggests that the high molecular weight complexes in panels a and b were due to 0A-CTD binding to the sites within the O2 to O4 region. Further evidence in support of this conclusion came from a mixed probe experiment in which 0A-CTD was simultaneously challenged with both the singly mutated 0A\textsuperscript{mut} DNA and the doubly mutated 0A\textsuperscript{mut} Δ2–4 DNA. The results show that in the mixture, the mobility of the singly mutated DNA but not the doubly mutated DNA was shifted (the two DNAs could be distinguished by their size difference; Figure 3A, panel d). Finally, the modest shift in mobility caused by the binding of 0A-CTD to the activation site and the large shift observed for binding to the operators is consistent with the idea that the activation site consists of a single 0A-CTD binding site whereas the operator region consists of multiple tandem binding sites.

Finally, we carried out DNase I footprinting to visualize the binding sites for 0A-CTD (Figure 3B). When we applied increasing amounts of 0A-CTD, we observed sites in the promoter region that became more resistant (indicated with arrows in Figure 3B) or more sensitive (indicated with stars in Figure 3B) to DNase I treatment. These sites were located in O1, O2, and in the promoter –10 region between them. Thus, O1 and O2 are indeed recognized by the DNA-binding domain of 0A.

**Proper sinl expression depends on read-through transcription from an upstream operon**

Just upstream of sinl and sinR is a two-gene operon, yqhHG in the same orientation as sinl and sinR (Figure 2C). We unexpectedly observed that insertions of a spec drug resistance gene at two sites upstream to, and in the opposite orientation as, sinl markedly impaired biofilm formation (Figure 2C; Supplementary Figure S3A). One insertion was located between the sinl promoter and the 3' end of the upstream operon. The other insertion was located at the immediate 5' end of the yqhH open-reading frame (Figure 2C). Finally, a third insertion in the yqhH open-reading frame but in the same orientation as sinl and sinR did not block biofilm formation (data not shown). The simplest interpretation of these results is that proper sinl expression depends on read-through as well as its own promoter.

Consistent with this possibility, no apparent terminator (i.e. a p-independent terminator) is present in the sequence between the operon and the promoter for sinl (Figure 1). Also, orthologs of yqhHG are found immediately upstream of sinl and sinR in the Bacillus species B. amylohydrogenificiens, B. licheniformis, and Bacillus pumilus (Rey et al., 2004; Veith et al., 2004; Chen et al., 2007; Gioia et al., 2007). Again, no apparent transcription terminator was found between the end of yqhHG and the sinl promoter in these species (Figure 1). The use of a construct in which lacZ was fused to the promoter region for yqhH (P\textsubscript{yqhH-lacZ}) indicated that the yqhHG operon is expressed at a low, constitutive level in MSgg medium (data not shown).

As a final test of the idea that proper expression of sinl depends on read-through, we integrated DNA containing either yqhHG sinl sinR or sinl sinR alone into the chromosome at the bkdB site of a deletion mutant (Δsinl sinR) lacking sinl and sinR. Supplementary Figure S3B shows that the bkdB::yqhHG sinl sinR construct complemented the deletion mutation but that the bkdB::sinl sinR construct did not. In toto, the findings indicate that yqhH yqhG sinl sinR is a four-gene cassette, which is transcribed from an upstream promoter P\textsubscript{yqhH} and two internal promoters, the Spot0A–P-activated promoter P\textsubscript{sinl}, and the constitutive promoter P\textsubscript{sinR} (Figure 2C).

**Proper sinl and sinR function depends on chromosome position**

We were surprised to discover that the capacity of the yqhHG sinl sinR cassette to complement the Δsinl sinR mutation was dependent upon its site of insertion into the chromosome. Insertions at sites distal to the origin of replication, bkdB(213') and ylnF(140'), complemented effectively, restoring biofilm formation to levels indistinguishable from those of the wild type (Figure 4A and B). In contrast, insertions at sites proximal to the origin, amyE(28') and ywrK(313') failed to complement (Figure 4A and B). Thus, in toto, sinl and sinR located at three origin-distal locations (the native location (218') and the bkdB and ylnF loci) supported biofilm formation, whereas insertions at two origin-proximal locations (amyE and ywrK) did not (Figure 4B).

We also measured expression of the yqxM matrix operon in a strain (YC687) in which sinl and sinR were at an origin-proximal position (amyE) using a P\textsubscript{yqxM-lacZ} reporter. The results showed that the expression of the reporter was drastically lower than that seen in a wild-type strain (YC686) in which sinl and sinR were at their normal position on the chromosome (Figure 5A).

Also, as judged by immunoblot analysis, moving sinl and sinR to an origin-proximal location (amyE) blocked the appearance of the regulatory protein SlrR (Figure 5B) whose synthesis is under the direct negative control of SinR and whose presence is diagnostic of biofilm formation (Chu et al., 2008; Kobayashi, 2008; Chai et al., 2010). Little or no SlrR could be detected in a lysate from cells of a strain in which sinl and sinR were at amyE (Figure 5B). In contrast, the regulatory protein was readily detected in a lysate from cells in which sinl and sinR were at their normal position (Figure 5B).
Because genes located near the origin are present in higher copy number than those near the terminus in cells undergoing DNA replication (Yoshikawa and Sueoka, 1963; Berkmen and Grossman, 2006; Lee and Grossman, 2006), we reasoned that the function of sinI and sinR is highly sensitive to gene dosage. To investigate this possibility, we introduced a second copy of the yqhHG sinI sinR fragment at the origin-distal site bkbB (213°) in an otherwise wild-type strain (i.e. a strain having a wild-type sinI and sinR at the native locus).

Consistent with our gene dosage hypothesis, the resulting strain (YC562) was severely impaired in biofilm formation (Figure 4B, right-most panel) as well as in the matrix gene expression (Figure 5A). Immunoblot analysis confirmed that SlrR was absent in the lysate prepared from YC562 cells (Figure 5B).

Finally, we wondered whether relief from metabolic repression (achieved by deleting the 0A~P operators in the regulatory region of sinI) would reverse the block in biofilm formation observed in the experiment of Figure 4B. We found that introducing a deleted-mutated copy of yqhHG sinI sinR in which sinI lacked the 0A~P operators (Δ2–4) to an origin-proximal location (amyE) in place of sinI and sinR at their normal location or introducing a second copy of yqhHG sinI sinR at an origin-distal location such that both copies of sinI lacked the 0A~P operators resulted in strains that were severely impaired in biofilm formation (although slightly less so than the corresponding strains with intact sinI; data not shown). Thus, the block in biofilm formation caused by mispositioned or excess copies of sinI and sinR could not be significantly relieved by eliminating 0A~P-mediated repression.

The binding of SinR to the promoter for a matrix operon demonstrates cooperativity

Why does doubling the dosage of sinI and sinR result in a block in biofilm formation? Because SinR binds to tandem sites in target promoters, we hypothesized that SinR binding to DNA is cooperative. In contrast, as single SinI molecules bind to SinR, we hypothesized that the interaction of SinI with SinR is not cooperative. Thus, doubling the copy number of the genes (and presumably the concentration of the proteins) would be expected to have a disproportionate effect on repression versus antirepression.

To investigate this possibility, we carried out electrophoretic gel mobility assays with the promoter region for the eps operon as a probe. The results show that the probe was substantially shifted to slower migrating species over a relatively narrow range of SinR concentrations (Figure 5C). That is, increasing the concentration three-fold (from 30 to 100 nM) substantially depleted the amount of free probe. Next, we investigated the effect of increasing concentrations of SinI on the amount of unbound DNA when the concentration of SinR was held constant at either 100 or 200 nM. The results show that even at the lowest concentration of SinI...
Cells entering sporulation are blocked in matrix gene expression

In light of our discoveries that high levels of Spo0A∼P represses sinl expression and that two copies of the sinl sinR genes prevent derepression of matrix genes, we predicted that cells entering sporulation would be blocked in matrix gene expression. Entry into sporulation occurs at high levels of Spo0A∼P, which as we have seen represses sinl and cells entering sporulation have two complete copies of the genome. Only after the stage of asymmetric division, when the sporangium is divided into large and small compartments, are chromosomes segregated from each other by translocation of one of the chromosomes across the septum (Errington, 2001). Also, because sinl and sinR are located near the terminus, both sets of the genes persist in the large compartment until almost the end of the segregation process. Thus, either high Spo0A∼P levels or doubling the gene copy number or both could be sufficient to block matrix gene expression upon entry into sporulation.

To determine whether matrix gene expression is blocked at the start of sporulation, we did two experiments, one designed to monitor gene expression in the cell population and the other in individual cells. In the first experiment, we constructed a strain that contains both a luciferase gene reporter fused to the yqxM matrix operon (P<yqxM-lux> and a green fluorescent protein (gfp) reporter fused to an early sporulation gene (P<yqxM-gfp>). The use of luciferase enabled us to measure real-time changes in the matrix gene expression (shown in Figure 7A; filled diamonds) when the cells started to enter sporulation (indicated by expression of the P<spor>reporter in Figure 7B). The results show that the P<spor>reporter was off at 7.5 and 8.5 h after inoculation, but on at 9.5 h (vertical arrows in Figure 7A and columns in Figure 7B). In contrast, the activity of the P<yqxM-lux>reporter peaked at 8.5 h then started to decline (Figure 7A). These results are consistent with the idea that matrix genes switch OFF as cells start to enter sporulation.

In the second experiment, we used a construct in which a fluorescent reporter encoding a proteolytically unstable gfp-LCN was fused to the promoter for the yqxM matrix operon (P<yqxM-gfp-LCN> (Pan and Losick, 2003). Cells harbouring the fusion (strain YCS85) were grown to
Figure 6 Block of matrix gene expression by transplanting sinI sinR to an origin-proximal locus in a strain engineered to induce sinI during growth. (A) Fluorescent microscopic analysis of cells harbouring both a $P_{yqXM}\text{-}yfp$ reporter and an IPTG-inducible copy of $spo0A\text{-}sad67$. Panels on the left represent cells with sinI and sinR genes at the native locus (origin distal, YC716) and those on the right are cells with transplanted $yqhHG\text{ sinI sinR}$ at the amyE locus (origin proximal, YC717). The operators for $0A\text{-}P$ were also deleted ($\Delta 2–4$) from the sinI regulatory region in both strains to avoid repression by high levels of $0A\text{-}Sad67$. (B) Flow cytometric analysis of the same cell samples from (A). In panel b, upon treatment of IPTG (10 $\mu$M) for 60 min, about half of the YC716 cells were strongly turned on the matrix reporter, whereas under the same condition, the YC717 cells were largely blocked in expression of the matrix reporter (panel d).

Figure 7 Matrix genes were shut off at the start of sporulation. (A) Luciferase activities (blue line with filled diamonds) of a strain (YC734) that harbours both a $P_{yqXM}\text{-lux}$ luciferase reporter and a $P_{spoIIA}\text{-gfp}$ reporter. Red line with filled squares represents growth curve of YC734 cultured in MSGg medium at 37°C. At 8.5 h after inoculation, activities of the luciferase reporter peaked and then started to decline. (B) Fluorescent microscopic analysis of YC734 cells at 7.5 h (panel a), 8.5 h (panel b), and 9.5 h (panel c) after inoculation. The $P_{spoIIA}\text{-gfp}$ reporter in YC734 was turned on at 9.5 h in a large proportion of cells, indicating that cells started to enter sporulation pathway.
late-stationary phase when a small number of sporulating cells started to appear. Supplementary Figure S4 and Table I show that whereas many (but not all; Chai et al., 2008) non-sporulating cells expressed the reporter fewer sporulating cells (as identified by the presence of a polar septum and as indicated with arrows in Supplementary Figure S4) did so. In contrast, a high proportion of both sporulating and non-sporulating cells exhibited fluorescence when, as a control, we used a strain (YC161 with Pspank-gfp) that produced GFP constitutively.

Finally, we asked whether metabolic control (high Spo0A~P levels repressing sinI) alone was sufficient to account for the block in the matrix gene expression during sporulation. To do so, we repeated the experiment of Supplementary Figure S4 using a strain (YC586) harbouring a deletion-mutated sinI gene that lacked the 0A~P operators (Δ2–4). The results obtained with YC586 were similar to those obtained with YCS85 (Supplementary Figure S4; Table I). We conclude that the shut off of matrix gene expression during sporulation is not accounted for, at least not to a substantial extent, by Spo0A~P-mediated repression of sinI.

Discussion

Sporulation and matrix production are mutually exclusive cell fates (Vlamakis et al., 2008; López et al., 2009). Certain metabolic conditions (growth in DS medium) favour spore formation but not the production of matrix, whereas other conditions (growth in MSGg medium) favour matrix production. Yet, both developmental processes depend on the activity of the master regulator 0A (Branda et al., 2001; Hamon and Lazazzera, 2001; Fujita and Losick, 2005). It was hypothesized that sinI is expressed at lower levels of 0A~P than are genes that govern entry into sporulation (e.g. spoIIE and the spoIIA and spoIIG operons). However, the nature of the mechanism (or mechanisms) that would prevent sinI from being expressed at high 0A~P levels and thereby block the expression of matrix genes was unclear. Here, we have discovered two mechanisms that render sporulation and matrix production mutually exclusive, which we refer to as metabolic control and chromosome copy number control.

Metabolic control is mediated by the presence of both activator and operator sites for 0A~P in the regulatory region of sinI. Low levels of 0A~P activate sinI transcription from the high-affinity activator site but high levels of the phosphoprotein repress transcription by binding to the low-affinity operator sites. Thus, in medium in which 0A~P rapidly rises to high levels, transcription of sinI is quickly curtailed and cells directly enter sporulation instead of forming a biofilm. Conversely, in medium in which 0A~P remains at intermediate levels for prolonged periods, sinI is expressed at a high and persistent level, promoting biofilm formation. Meanwhile, few cells in the population attain high enough levels of 0A~P to sporulate until the population has aged for an extended period of time.

Eventually, some cells in the population do start to sporulate in biofilm-inducing medium. Conceivably, the high levels of 0A~P attained by these cells might have prevented them from expressing sinI and hence switched off the expression of matrix genes. However, this seemed unlikely for the following reason. In earlier work, we showed that SinI sets in motion a self-reinforcing double-negative feedback loop that locks cells in a matrix-producing state (Chai et al., 2010). Thus, a matrix-producing cell that attains high levels of 0A~P and hence starts to sporulate might have been expected to continue expressing matrix genes even though synthesis of SinI had been shut off. What then accounts for the fact, as documented here, that sporulating cells do not express matrix genes, even when the operator sites for 0A~P have been deleted from sinI?

We propose that the doubling of chromosome copy number and hence of the copy number of sinI and sinR at the start of sporulation represents a second, cell-fate determining regulatory mechanism that rapidly shuts down matrix gene expression in cells that have entered the pathway to sporulate (Figure 4C). Under conditions of slow growth, and even under conditions of rapid growth when multifork replication is occurring, cells have two copies of the terminal region of the chromosome, where sinI and sinR are located, only transiently before undergoing division. In contrast, a unique feature of sporulation is the presence of two complete copies of the chromosome at the start of the process, these chromosomes forming a single elongated structure known as the axial filament. Only after the next stage of sporulation when the cell undergoes asymmetric division are the two chromosomes segregated into separate cells (Errington, 2001). We posit that the presence of two copies of the chromosome and hence of sinI and sinR rapidly shuts off transcription of matrix genes.

What is the molecular mechanism by which this doubling of gene copy number causes matrix gene expression to shut down? Matrix gene expression is controlled by a competition between antirepression by SinI and repression by SinR. We propose that the key to chromosome copy number control lies in the different sensitivities of antirepression and repression to gene dosage. Because the binding of SinI to SinR is not cooperative, doubling the copy number would be expected to increase the degree of antirepression monotonically. In contrast, the binding of SinR to DNA is, as we have shown, cooperative. This means that small changes in SinR concentration can lead to large changes in the degree of repression. In particular, cooperativity exhibits critical concentrations near which large changes in promoter activity can occur. If doubling copy number raises the concentration of SinR above this critical point, then the strength of repression can more than double.

This proposed relationship between copy number control and the relative sensitivity of the two mechanisms can be formalized mathematically. In the Supplementary data, we show that there is a simple rule-of-thumb for approximating the response of a SinR-regulated promoter to a doubling of SinI and SinR concentration:

\[
\text{fold change in expression} \approx 1 + \frac{H_R}{H_I} \left(\frac{1}{H_I} - 1\right)
\]
The $H$ terms here are logarithmic sensitivities, which are relative measures of change. $|H_{[\text{S}]}/H_{[\text{R}]})|$ is sensitivity of matrix gene expression to SinR concentration: if SinR concentration goes up $1\%$, matrix gene expression will decrease $|H_{[\text{S}]}/H_{[\text{R}]})|\%$. ($H_{[\text{R}]})$ is negative since SinR reduces expression; the absolute value signs in the formula are included to take this into account.) $|H_{[\text{S}]})|$ measures the sensitivity of the titration mechanism to Sinl concentration: if Sinl concentration goes up $1\%$, the amount of free SinR (which is what is assumed to bind SinR-regulated promoters) goes down $|H_{[\text{S}]})|\%$.

This formula explains what is necessary for copy number control by Sinl and SinR to work. First, to see a decrease in expression, it must be the case that $|H_{[\text{S}]})|$ is less than one (so that $H_{[\text{S}]})\>1$ is negative). If the titration of SinR as a function of Sinl can be well modelled by a Hill function, a simple way to ensure this condition is to have non-cooperative binding of Sinl to SinR, which is precisely what happens.

In this case, $1−|H_{[\text{S}]})|$ turns out to be the fraction of free SinR (labelled $f$), so that the fold change is $1−f|H_{[\text{S}]})|$. This recasting of the relationship exposes the conflict between mechanisms. At near saturating levels of Sinl, as expected in stationary phase, the fraction of free SinR will be small. Thus, according to our model, cooperativity of SinR binding can be understood as necessary for overcoming this block to allow a large decrease in matrix gene expression to occur.

Finally, our discovery that a mere doubling in gene copy number can have a profound effect on gene expression may provide insight into chromosome-counting mechanisms involved in sex determination in nematodes and X chromosome inactivation in mammals (Meyer, 2000; Payer and Lee, 2008; Chow and Heard, 2009). Sex determination in Caenorhabditis elegans is mediated by xol-1, which promotes the male fate when it is ON and the hermaphrodite fate when it is OFF (Rhind et al., 1995). Whether xol-1 is ON or OFF is determined by the number of X chromosomes. Repression of xol-1 is mediated by multiple genes on the X chromosome (X signal elements), including sex-1, whose product is a repressor for xol-1 (Carmi et al., 1998; Carmi and Meyer, 1999; Gladden et al., 2007). The xol-1 gene is OFF when X signal element genes are present in two copies and ON when present in one copy (Gladden et al., 2007). An appealing possibility for part of the basis for chromosome counting is that cooperative binding of Sex-1 molecules to xol-1 or synergistic interactions among X signal element proteins or both amplify the effect of doubling chromosome copy number to cause a disproportionate repressive effect on xol-1 expression. Likewise, in X chromosome inactivation in mammals, it is thought that a series of transcription factors including Oct4 that are present at low concentrations at the time of X chromosome inactivation localize preferentially to one of the two X chromosomes, the future active one, as a result of cooperativity in binding to DNA (Chao et al., 2002; Donohoe et al., 2009). This asymmetric distribution of the factors, in turn, stimulates expression of Tsix, which mediates silencing of the future, inactive X chromosome, and results in the loss of Tsix expression on the future, active X chromosome (Lee and Lu, 1999; Lee et al., 1999; Donohoe et al., 2009). As more is learned about the molecular mechanisms of sex determination and X chromosome inactivation, it will be interesting to see whether they share a common underlying logic with each other and with the bacterial example we have presented here.

Materials and methods

**Strains, media, and regents**

*B. subtilis* strains PY79, 3610, and their derivatives were grown in Luria-Bertani (LB) medium (10 g tryptone, 5 g yeast extract, and 5 g NaCl per liter broth) at 37°C for general purposes. The media used for assays of biofilm formation include M5gg and DS media. The recipe for M5gg medium was described previously (Branda et al., 2001). The DS medium was prepared as following: 8 g Bacto nutrient broth (Difco), 1 g KCl, 0.12 g MgSO₄·7H₂O, and 0.06 g NaOH per liter broth. Just before use, the following sterile solutions were added: 1 ml of 1 M Ca(NO₃)₂, 1 ml of 10 mM MnCl₂, and 1 ml of 1 mM FeSO₄. *Escherichia coli* strains were grown in LB medium at 37°C for general purposes and at 30°C for protein expression. A list of strains used in this work was provided in Supplementary Table S1. Plasmids and primers were summarized in Supplementary Table S2. Antibiotics were added to the media at the following concentrations: 10 μg/ml of tetracycline, 100 μg/ml of spectinomycin, 10 μg/ml of kanamycin, 5 μg/ml of chloramphenicol, and 1 μg/ml of erythromycin for *B. subtilis* strains, and 100 μg/ml of ampicillin for E. coli cells.

**Site-directed mutagenesis of the 0A–P binding sites**

To introduce point mutations to the consensus 0A-box in the sinl promoter (creating 0Amut in Figure 2A), we applied overlapping PCR mutagenesis. Briefly, we amplified two overlapping (by 27 bp) DNA fragments by PCR. The first DNA fragment spanned from a region upstream to the 0A-box to a region immediately downstream of the 0A-box, and it was amplified by using 3610 chromosomal DNA as the template and primers Pₐaf1 and Pₐar1. The second DNA fragment spanned from a region immediately upstream to the 0A-box to the start of sinl and it was amplified by using primers Pₐaf1 and Pₐar1. These two PCR products were gel purified and used as overlapping primers in the second round of polymerization reaction, which generated about 500 bp sequence of the regulatory region of sinl with nucleotide substitutions in the 0A-box.

Point mutations in the 0A–P operators were created similarly, except that different pairs of primers were used. The pairs of primers and corresponding operator mutations were described as follows: Pₐaf1-MF1 and Pₐar1-MR1 (for M1); Pₐaf2-MF2 and Pₐar2-MR2 (for M2); Pₐaf3-MF3 and Pₐar3-MR3 (for M3); Pₐaf4-MF4 and Pₐar4-MR4 (for M4); Pₐaf1-AR1 and Pₐar-AR1 (for Δ2–4); Pₐaf-RF1 and Pₐar-RS1-RF1 (for scrambled sequence of O2–O4).

Construction of plasmids and strains was described in Supplementary data.

**Assays for luciferase and β-galactosidase activities**

YC734 cells were grown in LB broth to exponential phase. Cells were then diluted into M5gg broth and were grown to mid-exponential phase. Cells were diluted again to 30 ml of fresh M5gg broth and were grown at 37°C with vigorous shaking. At various time points, cell samples were taken out, diluted appropriately, and measured for cell density at OD₆₀₀ and for luminescence at OD₅₃₂ by using a Multi-Detection Microplate Reader (BioTek). Luciferase activities are presented in arbitrary units and were normalized to cell density. Assays for β-galactosidase activity have been described previously (Chai et al., 2009).

As assays for pellicle and colony formation were described previously (Chai et al., 2010).

**Protein expression and purification**

Protein expression and purification of Spo0A-CTD (Molle et al., 2003) and of Sinl and SinR (Chai et al., 2009) were described previously.

**EMSAs and DNase I footprinting**

For assays of 0A-CTD binding to the sinl promoter, DNA probes used in the assays were generated by PCR using either plasmid or chromosomal DNA containing operator mutations or deletions as the templates, and using primers Pₐaf1-FP-F1 and Pₐaf1-FP-R1. Each PCR product was gel purified, digested with restriction enzyme EcoRI, and radio labelled by fill-in using the Klenow(exo-) fragment of T4 DNA polymerase in the presence of dioxy-nucleotides, dTTP and [α-32P]-dATP. DNA-binding reaction was incubated in 10 μl of binding buffer (10 mM Tris–HCl, 50 mM NaCl, 1 mM EDTA,
were incubated for 60 more minutes with shaking. A measure of 1 ml cells were harvested and washed twice with PBS buffer. Cells were then diluted in PBS and measured on a BD LSR II flow cytometer (BD Biosciences) with a solid-state laser at 488 nm. Fluorescent signals were collected by a 505-longpass filter and a 530/30-bandpass filter. Data were captured using FACS Diva software (BD Biosciences), and further analysed using FlowJo 7.5.5 software (http://www.flowjo.com). Figures were prepared using FlowJo 7.5.5.

Supplementary data
Supplementary data are available at The EMBO Journal Online (http://www.embojournal.org).

Conflict of interest
The authors declare that they have no conflict of interest.

References
Berkmen MB, Grossman AD (2006) Spatial and temporal organiza-
tion of the Bacillus subtilis replication cycle. Mol Microbiol 62: 57–71
Fujita M, Losick R (2005) Evidence that entry into sporulation in Bacillus subtilis is governed by a gradual increase in the level and activity of the master regulator Spo0A. Genes Dev 19: 2236–2244

5% glycerol, 1 mM DTT, 10 μg/ml BSA). Various concentrations of OA-CTD proteins were added to ~100 nM radiolabelled DNA probe and incubated on ice for 15 min. Reaction mixture was size fractionated on a 6% polyacrylamide gel (in 1 × TBE buffer) at 250 V.

For assays of SinR binding to the eps promoter, it was conducted similarly as described above except that SinI, if added, was incubated with SinR at room temperature for 20 min before both proteins were incubated with radiolabelled DNA probe. The ~200-bp DNA probe containing the eps promoter was amplified by PCR using primers FepsF1 and RepsR1, and radiolabelled similarly as described above.

DNase I footprinting assays followed a protocol described previously (Chai et al., 2010). The DNA probes containing the wild-type promoter region for the strn gene was amplified and end-labelled as described above.

Fluorescent microscopic analysis
Cells were grown in M5gg medium. A measure of 1 ml of cells were collected and washed twice with cold PBS buffer. Cells were suspended in 50 μl of cold PBS buffer. A measure of 3 μl of suspended cells was dropped to the centre of an agar-coated microscopy slide. Cover slides were pretreated with poly-L-lysine (Sigma). Samples were analysed using ×100 oil merged lens on an Olympus workstation BX61. Images were taken using an automated microscopy slide. Cover slides were pretreated with poly-L-lysine (Sigma). Samples were analysed using ×100 oil merged lens on an Olympus workstation BX61. Images were taken using an automated microscope (Universal Imaging Corporation) and Image J (http://rsweb.nih.gov/ij/).

Flow cytometry
For flow cytometric analysis, YC718 and YC719 cells were grown in M5gg medium to early exponential phase (OD600 = 0.3). IPTG was added to the cultures at a final concentration of 10 μM and cells were incubated for 60 more minutes with shaking. A measure of 1 ml cells were harvested and washed twice with PBS buffer. Cells were then diluted in PBS and measured on a BD LSR II flow cytometer (BD Biosciences) with a solid-state laser at 488 nm. Fluorescent signals were collected by a 505-longpass filter and a 530/30-bandpass filter. Data were captured using FACS Diva software (BD Biosciences), and further analysed using FlowJo 7.5.5 software (http://www.flowjo.com). Figures were prepared using FlowJo 7.5.5.