Science Progress (2006), 89(3|4), 243–266

Deciphering a complex genetic regulatory network: the *Bacillus subtilis* σ^{W} protein and intrinsic resistance to antimicrobial compounds

JOHN D. HELMANN

ABSTRACT

Bacillus subtilis*, a spore-forming soil bacterium, is the preeminent model system for the analysis of gene regulation in Gram-positive bacteria. Early genetic analyses established that this organism uses alternative sigma (σ) subunits to reprogram RNA polymerase to activate genes required for growth phase transitions, motility, general stress response, and sporulation. Unexpectedly, the genome sequence predicts the presence of an additional seven σ subunits: all members of the extracytoplasmic function (ECF) σ subfamily of regulators that typically respond to cell envelope stresses. Here, we review our current understanding of one of these σ factors, σ^{W}, with an emphasis on experimental strategies and approaches. Exposure to cell envelope active antibiotics and toxic peptides triggers a signaling cascade that releases σ^{W} from its cognate anti-σ thereby allowing transcription of ~60 σ^{W}-dependent genes. These genes encode proteins that inactivate, sequester, or eliminate toxic compounds from the cell.*

Keywords: *gene regulation, transcription, RNA polymerase, sigma, cell wall, antibiotic, bacteriocin, genomics, microarray*

John D. Helmann is a Professor of Microbiology at Cornell University in Ithaca, NY 14853-8101, USA. He received his PhD in 1987 from U.C. Berkeley where he studied *Bacillus subtilis* σ^D with M.J. Chamberlin. After post-doctoral studies with C.T. Walsh (Harvard Medical School) he joined the Cornell faculty in 1990. His laboratory currently studies gene regulation in *B. subtilis* in response to cell envelope, metal ion, and oxidative stress. He may be contacted at jdh9@cornell.edu

Scope and purpose

In this article, I will review how genetic approaches have enabled a detailed understanding of the role that the σ^W regulatory protein plays in protecting *Bacillus subtilis* against antimicrobial compounds made by other organisms in its environment. My emphasis will be on the scientific method, the development and refinement of hypotheses, and the logical progression of the experiments. Therefore, I will present results and ideas in approximately the order in which they were developed. This work is an example of how modern molecular genetic and genome-enabled approaches can be used in an organism with well developed genetics. Similar strategies are applicable to the molecular dissection of regulatory circuits in many microbial systems.

Microbiology in the age of genomics

The development of rapid DNA sequencing and assembly technologies has revolutionized microbiology. We now have access to the complete genome sequence of several hundred bacterial species including the best studied model organisms, numerous pathogens, and selected organisms of industrial or environmental interest. This wealth of sequence information has helped stimulate the burgeoning fields of bioinformatics and systems biology. On a practical level, the availability of complete genome sequences has greatly accelerated the ability of bacterial geneticists to define systems of interest using both forward and reverse genetics.

In a typical forward genetic screen, mutants are sought that are altered in a biological property or process, and the corresponding genes are then identified. With the complete genome sequence already in hand, it is comparatively easy to identify the corresponding genetic change. For example, the site of a transposon insertion can be identified by sequencing the transposon-chromosome boundary and the genetic locus identified directly. Locations

of spontaneous mutations can often be determined by isolating the relevant gene(s) on a complementing plasmid and then determining the corresponding region of the chromosome. While it is still necessary to use DNA sequencing to determine the sites of mutation, few among us miss the tedium of actually having to sequence and assemble gene sequences *de novo* as a first step in making inferences about the role of the corresponding gene product.

While it is frequently true that identifying the relevant genes provides insights into their biochemical functions or activities, this is often not the case. Approximately one-third to one-half of bacterial genes are of uncertain function, despite our ability to group most of them into families of related sequences spanning from a few to hundreds of organisms. Understanding the roles of these "hypothetical conserved" proteins is a major challenge for future work. The availability of large amounts of genome sequence information, coupled with the abundance of genes of unknown function, has also led to a large increase in the approach often known as "reverse genetics."

Experiments based on reverse genetics begin with a gene and search for a phenotype, rather than the converse. The rationale for using a reverse genetics approach depends on the system and, of course, the investigator. For example, an inability to disrupt a gene may be evidence that it is essential for viability and therefore a possible new drug target. Some scientists focus on unknown function genes that are universally (or near universally) conserved, reasoning that any insights will be applicable to a large number of organisms. Reverse genetic approaches are also useful for defining specific roles for proteins that can only be assigned a general function. For example, in our studies we seek to identify the physiological roles of genes encoding alternative sigma (σ) subunits of RNA polymerase. While these proteins almost certainly play a regulatory role by redirecting RNA polymerase to specific promoter sites, understanding the scope and purpose of their regulatory influence remains a formidable challenge.

Bacterial sigma factors

The ability of cells to adapt to a changing environment depends on the timely regulation of gene expression. In bacteria, gene expression is usually regulated at the level of transcription. This often involves activators and repressors that bind DNA and fine-tune the ability of RNA polymerase to initiate transcription from adjacent

promoter sequences. In addition, the promoter selectivity of RNA polymerase can be altered by the production of alternative σ factors[1].

The σ subunit binds to the core RNA polymerase (generating the holoenzyme) and enables site-specific DNA-binding at appropriate promoter sites (frequently associated with conserved sequences at the -10 and -35 regions relative to the start point of transcription). Most promoters in the cell are recognized by the major, vegetative σ factor (*E. coli* σ^{70} or *B. subtilis* σ^{A}). Production of a new σ can activate large numbers of unlinked genes, all controlled from promoter sites of characteristic sequence distinct from those recognized by the major σ factor[1].

The number of σ factors varies greatly between different bacterial species. *Escherichia coli* encodes ~ 7 σ factors, *Bacillus subtilis* encodes at least 18, while *Streptomycetes coelicolor* encodes >60. Prior to the sequencing of the *B. subtilis* genome (completed in 1997), the functional role of over half of the encoded σ factors had been determined by traditional genetic and biochemical analyses (Table 1). Since *B. subtilis* is the model organism for the Gram positive bacteria, and is the second most studied bacterial species (after *E. coli*), it was surprising to find that the genome sequence contained an additional seven genes encoding σ factors. Since σ factors often activate transcription from large numbers of genes, mutations in σ factors can have pleiotropic effects: yet none of these seven genes corresponded to known genetic loci. This suggests

Table 1 *The σ factors of Bacillus subtilis[33]: more FUN[8] ahead*

σ	Function(s)	Size of regulon
σ^{A}	Primary σ	~4000
σ^{B}	General stress response	~200
σ^{D}	Flagella, chemotaxis, autolysins	57
σ^{E}	Sporulation – early mother cell	60
σ^{F}	Sporulation – late mother cell	19
σ^{G}	Sporulation – late forespore	54
σ^{H}	Transition state regulation, antibiotic production	26
σ^{K}	Sporulation, late forespore	39
σ^{L}	(σ^{54}-type) levanase, amino acid catabolism	23
σ^{M}	FUN (function unknown), antibiotic stress (?)	?
σ^{V}	FUN	?
σ^{W}	Intrinsic antimicrobial resistance	~60
σ^{X}	FUN, controls some cell surface properties	15
σ^{Y}	FUN, possible bacteriocin synthesis and resistance	?
σ^{Z}	FUN	?
σ^{ylaC}	FUN	?

that they might regulate specialized subsets of genes active in processes that had not been well studied in the laboratory.

Strategies for defining the functions of ECF sub-family σ factors

One notable feature of these seven σ factors is that all are members of the extracytoplasmic function (ECF) subfamily[2]. ECF σ factors are typically regulated by an anti-σ factor, often a transmembrane protein, that sequesters the σ in an inactive state until a signal releases the active σ. The prototypical ECF σ factors are the *E. coli rpoE* (σ^E) and *S. coelicolor sigE* (σ^E) proteins. In these, and many related systems, the ECF σ factor activates transcription of its own gene which is encoded in an operon together with the anti-σ factor (Figure 1). We set out to test the hypothesis that this same model would apply to the ECF σ factors of *B. subtilis* and to determine their physiological roles. Here, I will focus on the best understood example, σ^W.

To define the biological role of a σ factor, the first step is to simply make a null mutant and evaluate the physiological consequences. This is the standard, "reverse genetics" approach to defining function. Unfortunately, this approach was not very informative: null mutations in each of the seven ECF σ factor genes are without obvious effects on cell growth. The null mutant strains grow as well as wild-type under a variety of conditions, they sporulate, are competent for genetic transformation, and do not have aberrant cell morphology or nutritional requirements. In the case of a *sigX* null mutant, there was a slight increase in sensitivity to heat and oxidative stress, but the significance of these effects was not immediately clear[3]. Even multiply mutant strains (lacking three or four of the σ factors) were without obvious defects. In retrospect, this result was probably to be expected. We now view the ECF σ regulators as contigency genes that are activated only under specialized conditions[2]. If cells are grown under non-activating conditions, the regulatory system is essentially silent and cells with and without the system will appear the same.

We next set out to define the physiological role of the ECF σ factors by addressing two key questions: What signals activate the expression and/or activity of the σ factor? and What genes are controlled by the σ? In other words, we need to define both the Input and the Output of the regulatory system. These are interdependent problems: to define the signals that activate the ECF σ factor we need a reporter of its activity (*i.e.* a promoter dependent

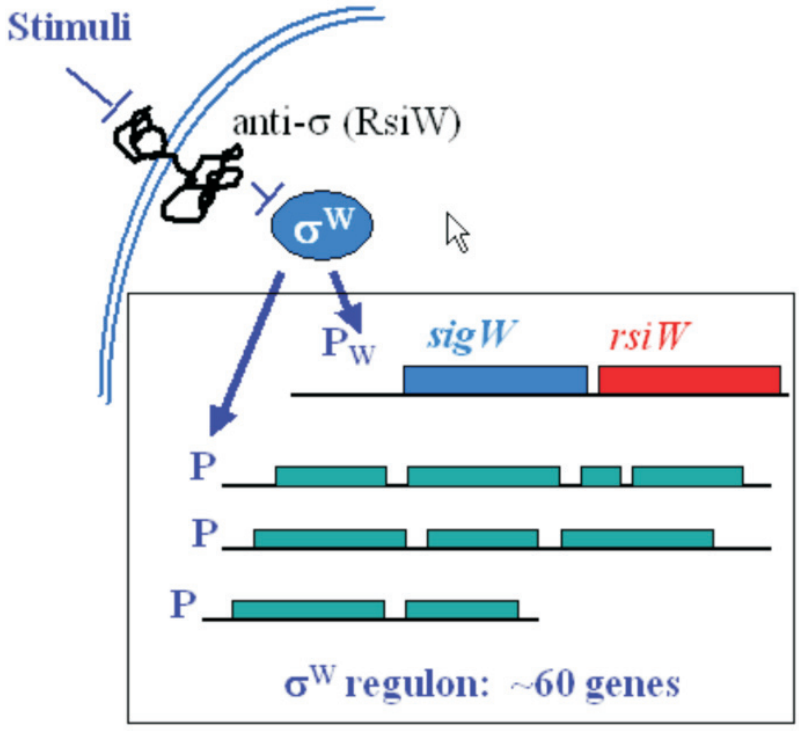

Fig. 1. *General scheme for gene regulation by* σ^W. *The general scheme presented her is typical for regulons controlled by members of the extracytoplasmic function (ECF)* σ *factor family. ECF* σ *factors are typically encoded in an operon with a gene encoding an anti-*σ *factor (RsiW = regulator of sigma-W). The latter protein is often, but not always, membrane-localized and is degraded (see Fig. 3), or otherwise inactivated, in response to specific stimuli (see Table 3). Once released from its anti-*σ*, the ECF* σ *factor binds RNA polymerase and transcriptionally activates both its own expression and the expression of its regulon (all regulated operons). In the case of* B. subtilis σ^W, *our current estimate is that this* σ *activates the expression of ~60 genes in 30 operons (Table 2).*

on the σ factor that we can fuse to a convenient reporter gene like *lacZ*) and, conversely, to identify the regulated promoters we need a way to turn the system on. The solution to this conundrum is the observation that many ECF σ factors are autoregulated (Figure 1). If we can identify the corresponding autoregulatory site upstream of the σ factor gene we can construct the required reporter system and may also gain insights into the particular sequences recognized by the cognate σ factor.

Autoregulation of σ^X and σ^W

We first applied this approach to the *sigX* regulon[4]. We found that the *sigX* gene is regulated by two promoters: one controlled by σ^A and the other dependent on σ^X (called P_X). To determine the precise nucleotides important for recognition by the σ^X-containing RNA polymerase, we isolated just P_X and generated a reporter fusion to *lacZ*. Using chemically synthesized oligonucleotides we incorporated random nucleotides throughout this region and assessed the effect on P_X activity. The results indicated that σ^X activity requires two regions of sequence analogous to the -35 and -10 elements recognized by σ^A. These sequences (tgtAAAC N_{16} CGwCww; where lower case is less important, N = any base, and w = A or T) were also found preceding several other operons which were subsequently shown to require σ^X for their activation. Naturally, these studies required that we first identify conditions under which σ^X was active. Fortunately, there is sufficient σ^X-dependent transcription in late-logarithmic phase cells that we were able to monitor activity. In addition, mutational analysis identified the gene downstream of *sigX* as encoding an anti-σ, RsiX. In an *rsiX* mutant strain, σ^X is constitutively active[3].

Several puzzling observations emerged during our characterization of the σ^X regulatory system and these motivated us to investigate the role of another ECF σ factor, σ^W, in more detail. First, it became apparent that some of the promoters that we had identified as possible targets for σ^X-directed transcription (based on the presence of the characteristic promoter sequence), were not, in fact, dependent on σ^X. Indeed, these promoter sites (as monitored by primer-extension analysis of the specific RNA transcript) were silent in wild-type cells and even when *rsiX* was absent[4]. Unexpectedly, these promoter sites were active in cells lacking σ^X. We hypothesized that (i) these sites might actually be the targets of another ECF σ factor and (ii) the activity of this other σ might be elevated in strains lacking σ^X. This turns out to be correct since we were able to show, in several cases, that the transcriptional activity detected in the *sigX* mutant strain was lost in a *sigXsigW* double mutant. By monitoring promoter activity *in vitro*, using reconstituted σ^X and σ^W holoenyzmes we demonstrated that some promoters are recognized only by σ^X (*e.g.* the autoregulatory site of the *sigXrsiX* operon; P_X), some only by σ^W (*e.g.* the *sigW* autoregulatory site, P_W), and some by both holoenzyme forms[5].

These initial studies led to the following model for the interactions between σ^X and σ^W. The σ^X regulon is normally activated during late-logarithmic growth phase[3]. In contrast, the σ^W regulon is normally activated (albeit to only a low level) in early stationary phase[6]. Both σ factors are autoregulated: *sigX* is partially dependent on σ^X for its expression (from P_X[3]) and *sigW* is entirely dependent on σ^W (from P_W[6]). In the absence of σ^X, a P_W reporter fusion turns on earlier than normal. Conversely, in a *sigW* mutant strain, there is elevated expression of P_X and other σ^X-dependent promoters. We can envision at least two mechanisms to account for this type of "cross-talk." First, this type of regulation may be due to physiological compensation: cells lacking *sigX* experience an undefined stress that activates P_W and, conversely, in the absence of *sigW* the activation of σ^X is accentuated. Second, this could be an example of promoter occlusion. We hypothesized that the σ^X holoenzyme binds to P_W (which has a quite similar promoter sequence to P_X) but is unable to initiate transcription. In effect, the σ^X holoenzyme transcriptionally represses P_W. Similarly, the σ^W holoenzyme might be able to transcriptionally occlude P_X and thereby repress this promoter (as well as, indirectly, other σ^X-dependent promoters). While we have not excluded this latter model, the physiological compensation model is currently favored, if only because there is more precedent for these types of effects.

The role of σ^W as assessed by identification of target genes

To define the function of σ^W, we next attempted to identify target genes that might depend on σ^W for expression. Analysis of P_W, the autoregulatory site of the *sigWrsiW* operon, identified consensus elements of TGAAAC(−35) and CGTA(−10)[6]. By searching the *B. subtilis* genome for the presence of these sequence motifs (with a 16 nucleotide spacer) we identified 15 additional candidate promoters upstream of genes (several sites detected within coding regions may represent false-positives). Remarkably, all 15 of these sites were used by σ^W *in vivo* as judged by analysis of *lacZ* fusions and primer extension mapping of transcript start sites[7]. As expected, transcription was dependent on σ^W and, as noted above, elevated in a *sigX* mutant strain.

We anticipated that defining the σ^W regulon (the complete set of genes activated by σ^W) would provide immediate insights into the physiological role of this σ. However, of the 30 or so genes

Table 2 *Overview of the σ^W regulon*

Functional group[a]	Operons
σ^W and its anti-σ	*sigW rsiW*
Negative regulator of σ^W	*ysdB*
Cell envelope synthesis	*pbpE racX*, *yuaFGI (?)*
Resistance to fosfomycin	*fosB*
Resistance to SdpC	*yfhLM*, *yknWXYZ*
Resistance to sublancin	*yqeZyqfAB*
Resistance to *B. amyloquifaciens*	*ydbST*
Predicted proteases	*yjoB*, *yteIJ*
Detoxification (?)	*ybfO*, *yceC*, *ydjP*, *ythPQ*
Small peptides (bacteriocins?)	*ydjO*, *yvlC*, *yxzE*, *yoaF*, *yoaG*
Unknown function (FUN)	Remaining genes

[a]The functional role is assigned where known. Possible functional roles, as inferred from sequence similarities, are indicated by (?) Approximately one-half of known σ^W-dependent genes are still in the unknown function category.

controlled by these promoter sites, only one had an annotated function: the *pbpE* gene encodes a minor penicillin-binding protein of uncertain function. The remaining genes were all assigned an arbitrary name as part of the genome project. In *B. subtilis*, unknown function genes are assigned names starting with "y" (with additional letters denoting their approximate chromosomal location). One mnemonic device is that the "y" stands for "why is this gene there?" Perhaps optimistically, such "*f*unction *un*known" genes have also been referred to as FUN genes[8]. In our case, however, we had merely linked an unknown function σ factor to the transcriptional control of several dozen unknown function genes. The corresponding proteins included potential transporters and several hydrolases (Table 2), and the majority are predicted to be located in the membrane[7]. Proposing a precise physiological role for σ^W on the basis of this initial list was difficult. One clue emerged from the finding that the transporters controlled by σ^W bore some resemblance to efflux systems involved in export of toxic peptides (bacteriocins). In addition, σ^W controls the expression of several small peptides that might be candidates for bacteriocins. Therefore, we speculated that σ^W might control the production of, and resistance to, antimicrobial compounds. However, an initial screen of commercially available antibiotics, as well as a variety of other toxic agents, failed to define any obvious sensitivities in the *sigW* mutant strain.

We hypothesized that this initial bioinformatic screen probably underestimated the true extent of the σ^W regulon. To develop a more comprehensive inventory of σ^W-regulated genes, we employed

three additional strategies[9]. First, it is well known that σ factors often tolerate some degeneracy in their target promoter sequences as well as variability in the spacer length. Therefore, we repeated our computer-aided search of the genome sequence using degenerate recognition elements and thereby identified additional candidate target genes. Second, we compared the mRNA profile (transcriptome) of wild-type and *sigW* mutant cells to identify genes expressed at a lower level in the mutant strain. This approach was moderately successful, but was limited by the fact that many *sigW* dependent promoters are weakly active even in wild-type cells. Moreover, it is not clear, even when effects are observed, whether these are a direct or an indirect effect of the *sigW* mutation. Therefore, we developed a third approach to more directly identify those promoter sites that can be directly read by the σ^W holoenzyme. For this experiment, σ^W was added to core RNAP to generate the σ^W holoenzyme. This enzyme was then used to transcribe total genomic DNA (digested into fragments using restriction enzymes). The resulting radiolabelled transcripts were then hybridized to a nylon membrane filter (macroarray) containing duplicate spots of DNA fragments representing each of the >4000 genes of the *B. subtilis* genome. Signals appear when a σ^W-activated transcript anneals to the corresponding gene (often only the first gene in the operon). This technique, referred to as ROMA (*r*un-*o*ff transcription/*ma*croarray *a*nalysis) led to the identification of several additional target sites. Together, the bioinformatic, transcriptome, and ROMA approaches defined a σ^W regulon of ~ 60 genes controlled by ~ 30 promoter sites[9] (including those listed in Table 2).

The characterization of additional, σ^W-dependent genes served to reinforce the notion that σ^W controls an ''antibiosis'' regulon likely involved in the production of, and resistance to, antimicrobial compounds[9]. Although many of the σ^W target genes encode unknown function proteins, one encoded a potential fosfomycin resistance gene. Fosfomycin is a small, cell permeable compound made by certain soil bacteria that inhibits the first committed step in peptidoglycan biosynthesis. Indeed, a *sigW* mutant strain is much more sensitive to fosfomycin than wild-type and this σ^W-dependent gene (renamed *fosB*) is both necessary and sufficient for resistance[10]. This finding strengthened our hypothesis that σ^W is involved in defense against antimicrobial agents. However, it is also clear that σ^W controls ~ 60 genes and this finding does not shed much light on the roles of the other genes.

General challenges in defining regulons

This work illustrates some of the common challenges that arise when attempting to define the regulon of a transcription factor. First, if one can obtain a detailed understanding of the sequence requirements for DNA-binding, it is often possible to identify a significant subset of the regulated target genes using simple bioinformatic searches (pattern recognition or searches using weight matrices or hidden-Markov models). However, in our experience (with several different regulatory proteins) this approach by itself often identifies only one-half to perhaps two-thirds of the actual target genes. With proteins that have highly degenerate or poorly understood binding specificity, the results are even worse. This approach is plagued by both false-positives (sites that appear to match the consensus but that are not involved in regulation, or that are inappopriately positioned relative to adjacent genes) and false-negatives (sites too diverged from the search pattern to be recognized by the computer algorithm; although the regulatory protein has no difficulty finding such sites *in vivo*!).

Second, microarray-based (transcriptome) approaches are very powerful but the results are much better if conditions that strongly activate the regulon can be defined first. Moreover, the contribution of one regulator to gene expression can be easily masked by others. For example, a target gene may be repressed by another protein, and therefore not expressed under the growth conditions used, or it may be expressed from multiple promoters or controlled by multiple regulatory proteins and this may drown out the contribution of the regulator under study. An even more serious problem with the transcriptome approach is that there are often indirect, propagated effects of activating a regulator. The regulator may control the expression of other transcription factors that then activate or repress their target genes. In other cases, activation of one regulon may create stresses on the cell that lead to activation of other responses.

Several approaches have been developed to deconvolute the complex transcriptional patterns observed in transcriptome studies and separate the direct from the indirect effects of the regulator under study. One of the most powerful, the so-called ChIP-to-chip approach, uses protein-DNA crosslinking to identify those DNA regions that are associated, *in vivo*, with a regulatory protein[11]. In our work we introduced a related technique, ROMA[9], to define those DNA regions where a given regulator (in our case, a particular holoenzyme) is active. This technique is complementary

to the *in vivo* approach, but it too has its limitations. For example, ROMA will only detect those promoter sites that are active *in vitro*, on linear DNA, in the absence of activator proteins. Our major conclusion, from our investigations into the scope of the σ^W regulon, is that a comprehensive inventory requires the integration of multiple experimental approaches.[9]

The role of σ^W as assessed by identification of inducing signals

In parallel with our efforts to define the σ^W regulon, we also sought to identify those growth conditions that might activate expression of this system. There are two general strategies for the identification of inducing signals. In the first, one exposes cells to various physical or chemical stresses and monitors the activity of a suitable reporter (*e.g.* a P_W-*lacZ* fusion). In the second, one generates a library of random mutations to identify genes that affect activity of the reporter fusion. These genes could include direct regulators of σ^W expression or activity, or participate in pathways that, when perturbed, generate a stress signal that activates σ^W.

As an example of the first approach, we tested the effects of a large number of physical and chemical stresses to identify factors that might activate P_W and thereby up-regulate the whole regulon. Since the initial clues from the target gene mapping suggested a possible role in antibiotic resistance, we included in our screen a large number of antimicrobial compounds[12]. This work led to the finding that inhibitors of cell wall biosynthesis and certain membrane-active compounds strongly activate σ^W (Table 3). Examples of inducing

Table 3 *Chemical and physical conditions that induce the σ^W regulon*

Class	Examples of inducing compounds/conditions
Environmental stress	Alkali shock, SPP1 phage infection
Cell wall antibiotics	Vancomycin, cephalosporin, D-cycloserine, (and weakly by bacitracin, tunicamycin, and fosfomycin)
Toxic peptides / bacteriocins	LL-37 (human cathelicidin), SdpC, sublancin
Membrane-active agents	Triton-X-100, nigericin

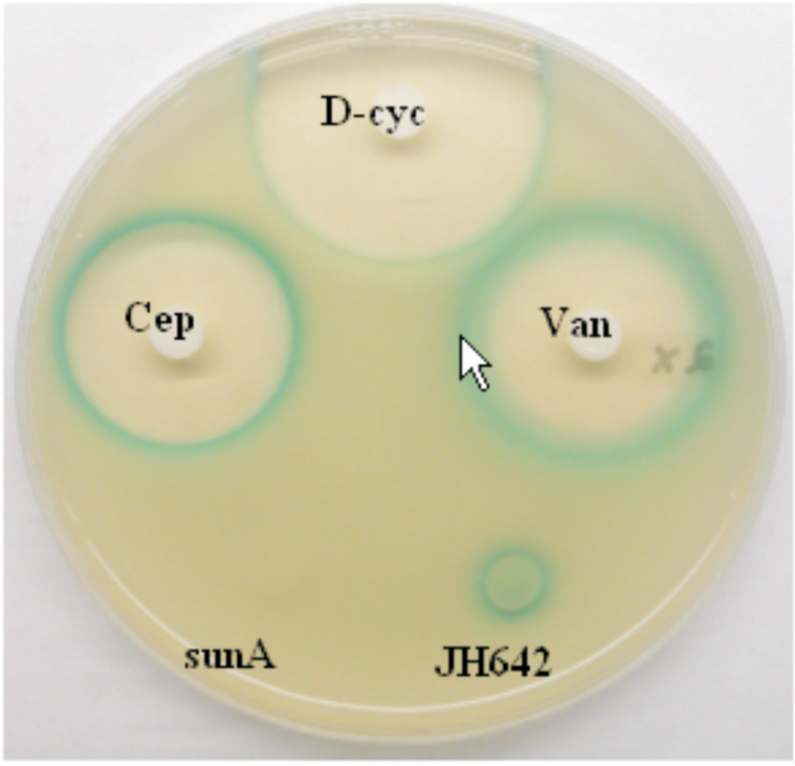

Fig. 2. *Induction of the σ^W regulon by antibiotics. Antibiotic sensitivity and gene induction can be simultaneously monitored using a "zone-of-inhibition" assay. A petri plate is inoculated with a* B. subtilis *strain containing a P_W-lacZ reporter fusion. Activation of the lacZ gene generates β-galactosidase which is visualized by the blue color of the cleavage product of the X-gal indicator (incorporated into the plate). Filter disks containing antibiotics (Cep = cephalosporin, C-cyc = D-cycloserine, and Van = vancomycin) are laid on the plate and diffusion of the antibiotic generates a zone of growth inhibition surrounding each disk. The size of the zone is an indicator of antibiotic sensitivity, and the development of a blue ring is indicative of activation of P_W. The two spots at the bottom of the plate illustrate the principle of the "spot-on-lawn" experiment. In this case the plate is spotted with several microliters of a culture of bacteria. A blue ring (and small zone of inhibition; not visible in this example) results from the production of the bacteriocin sublancin from* B. subtilis *strain JH642. The isogenic strain lacking the sublancin structural gene (sunA) fails to induce P_W.*

compounds include cephalosporin, vancomycin, D-cycloserine, nigericin, and the detergent Triton-X-100 with weak induction noted by fosfomycin, bacitracin, and tunicamycin (Figure 2). Using DNA microarray technology, we confirmed that vancomycin induced the majority of the previously defined σ^W target

genes, as expected[12]. These findings immediately suggested that some or perhaps many of the σ^W controlled genes might confer resistance to these antibiotics. It was therefore frustrating to observe that, apart from fosfomycin, the *sigW* mutant was not detectably more sensitive to any of these inducers. These antibiotics are, in effect, gratuitous inducers: although they turn on the σ^W regulon, the cell does not gain any growth advantage. A similar phenomenon was discovered by Wiegert and Schumann during an investigation of the alkali stress response: transcriptome analyses revealed a strong induction of the σ^W regulon during alkali shock[13]. However, a *sigW* mutant was not more sensitive to alkali stress suggesting that this, too, was due to gratuitous induction.

As an example of the second approach, we screened a library of transposon insertion mutations for induction of P_W[14]. For these studies, we used a P_W-*cat*-*lacZ* operon fusion in which activation of P_W confers chloramphenicol resistance (increased *cat* expression) and a blue color on X-gal plates (increased *lacZ* expression). The strongest induction was observed in strains containing a transposon insertion in the *yvbA* gene encoding a putative DNA-binding regulatory protein (see below for an explanation of this effect). In addition, we recovered one insertion in a known σ^W-dependent gene, *ysdB*, suggesting that this gene either directly or indirectly down-regulates the σ^W-dependent stress response. Finally, we recovered insertions in various genes possibly involved in antibiotic synthesis or resistance including predicted multidrug efflux pumps[14]. While this screen was successsful, it was likely not saturating. We did not recover insertions in *rsiW*, the anti-σ factor for *sigW*.

We also tested for effects of genes previously linked to the regulation of antibiotic resistance. In *B. subtilis*, nutrient limitation and high cell density activates the sporulation pathway controlled by the Spo0A protein. In addition to activating the sporulation pathway, Spo0A represses transcription of *abrB*, which encodes a repressor of "antibiotic resistance" (originally measured as resistance to polymyxin B). As a result, *spo0A* mutants are both sporulation defective and antibiotic sensitive. In contrast, a *spo0AabrB* double mutant is still sporulation defective, but regains antibiotic resistant. The *sigWrsiW* operon is directly repressed by the AbrB protein, which contributes to the fact that this operon is induced as cells enter stationary phase[15]. In addition, at least some σ^W target genes are also repressed by AbrB.

A genetic link between *sigW* and the antimicrobial peptide SdpC

While the developing evidence suggested a strong link between σ^W and resistance to antimicrobial compounds, there was a frustrating lack of correspondence between the nature of the known inducers, and the role of σ^W-controlled target genes. For example, fosfomycin is a weak inducer, even though σ^W is critical for fosfomycin resistance[10]. Conversely, σ^W and its regulon is induced by vancomycin, cephalosporins, and alkali stress: but a mutant is not more sensitive to these stresses[12]. Thus, it is hard to argue that this is the *raison d'etre* for σ^W.

An important advance in our understanding of the σ^W regulon emerged from the characterization of the *sdpABC* operon[16]. Losick and colleagues compiled a complete inventory of those genes controlled by the master regulator of sporulation (Spo0A), using many of the same approaches described above for σ^W. Typically, Spo0A-dependent genes encode functions needed for sporulation and mutants have sporulation defects. Thus, it was something of a surprise when it was found that mutations in the Spo0A-dependent *sdpABC* operon actually sporulated better than wild-type, with both faster spore formation and a higher sporulation efficiency (hence the name; *s*porulation *d*elaying *p*roteins). This was attributed to the production of a toxic peptide, later shown to be the product of the *sdpC* gene[17,18]. Production of SdpC results in lysis of those cells in the population that have not yet activated the Spo0A transcription factor and the released nutrients presumably allow the producer cells to delay their commitment to sporulation. Importantly, cells making SdpC are also resistant to its toxic effects because SdpC activates the expression of a convergent operon (*sdpRI*) encoding a regulatory protein (SdpR) and an immunity protein (SdpI)[18]. Intriguingly, the *sdpR* gene is the same as the *yvbA* gene that was identified in our genetic selection for mutations that up-regulate σ^W.

These converging lines of investigation led to the following model. In cells containing a transposon insertion in the *yvbA* gene there is no expression of either *yvbA(sdpR)* or *sdpI* and these cells are consequently sensitive to the SdpC peptide. When cells begin making SdpC in the absence of its specific immunity peptide (SdpI) there is induction of σ^W. This model was confirmed by demonstrating that induction of σ^W was observed in an *sdpI* mutant, but not in cells that are also mutant for *sdpC*.

To extend these studies we used a simple, but powerful assay known as "spot-on-lawn." In this technique, a petri plate is seeded with a *B. subtilis* strain (*e.g.* wild-type or a *sigW* mutant) at low cell density and then an antibiotic-producing strain is spotted at high cell density in the center of the plate. Since antibiotic production typically commences in the post-exponential growth phase, the cells spotted at high density will reach saturation and begin producing antibiotics while the lawn cells are still growing and therefore highly sensitive to antibiotic inhibition. When cells producing SdpC are spotted onto a lawn of cells containing a P_W-*lacZ* reporter fusion (and lacking the immunity peptide SdpI), there is a notable zone of growth inhibition surrounded by a region of P_W induction (a blue ring). This indicates that SdpC released from one population of cells can activate *sigW* in a sensitive population of neighboring cells[17].

In the course of these studies it became apparent that σ^W also functions to protect cells against SdpC toxicity. Cells lacking the immunity protein SdpI are impaired in growth when SdpC is expressed, but the *sdpIsigW* double mutant is much more severely affected[17,19]. We thus hypothesized that σ^W may control genes that provide "intrinsic" resistance to antimicrobial compounds distinct from that conferred by their specific immunity genes. As a species, *B. subtilis* can synthesize perhaps two dozen or more antibacterial compounds[20]. In most cases, the genes for biosynthesis and specific immunity are closely linked and co-regulated (*e.g.* the convergent *sdpABC spdRI* operons). Therefore, any cells expressing a toxic antimicrobial will also express the needed resistance functions. However, the genes for synthesis and immunity are often located on mobile genetic elements and are present in some strains (as "antibiosis" islands), but not in closely related strains[20]. We therefore reasoned that σ^W may play a role in providing a broad-based, intrinsic immunity to *B. subtilis* against a range of compounds likely to be made by closely related organisms including other Bacilli, and perhaps some of the many agents produced by *Streptomyces* spp.

Role of σ^W in providing intrinsic immunity to antimicrobial compounds

To further explore the link between the σ^W regulon and antimicrobial resistance we next focused our attention on compounds known to be made by various *Bacillus* species using the spot-on-lawn assay. Remarkably, in nearly every case where we could

discern noticable growth inhibition, the inhibitory effects were greater on lawns of the *sigW* mutant (and/or reduced in the *rsiW* mutant). We investigated two cases in detail: an unknown antimicrobial compound produced by *B. amyloquifaciens* and a potent bacteriocin (sublancin) encoded on the *B. subtilis* SPβ prophage[17]. For both of these examples, exposure to the antimicrobial agent induces the σ^W regulon (see Figure 2) and a *sigW* mutant is significantly more sensitive than wild-type. In the case of sublancin, the prophage encodes both bacteriocin production and resistance. However, non-lysogens (such as our wild-type strain) lack this specific resistance. In this case, σ^W controls the most significant resistance determinants.

We next sought to identify which, of the $\sim$60 or so genes controlled by σ^W, were important for resistance to these various antimicrobial agents. By screening a panel of mutant strains, each individually deleted for a σ^W-controlled operon, we were able, in each case, to link resistance to one or two specific operons[17]. Resistance to SdpC is dependent on an ABC-transporter system that likely functions to export SdpC from cells and also on a membrane-protein, YfhL (Table 2). Interestingly, YfhL is a homolog of the SdpI specific immunity protein[18], suggesting that these two membrane proteins work by a similar mechanism: sequestration of the toxic peptide in the membrane. YfhL is most important in protecting cells from SdpC when it is present in the environment (produced by neighboring cells). In contrast, the efflux system is most important in protecting cells that are making SdpC (but lack SdpI) from its toxic effects[17]. Using a similar strategy we identified operons controlled by σ^W that confer resistance to sublancin and the unidentified compound from *B. amyloquifaciens*. Together with the previous identification of *fosB* as a fosfomycin resistance gene, this allows us to assign clear functions to at least six of the $\sim$30 operons controlled by σ^W (Table 2). Sequence inspection suggests that many of the others may also provide intrinsic resistance against antimicrobials, particularly toxic peptides, although the corresponding compounds are not yet identified.

A major physiological role of σ^W is therefore to provide resistance to antimicrobials and thereby allow *B. subtilis* to better compete in the soil microenvironment. In experiments in which various strains are co-inoculated on solid medium, the nature of this competition can be easily visualized. Typically, in a spot-on-lawn assay the antibiotic-producing strain remains contained within the original zone of inoculation. In contrast, if the lawn is composed of *sigW* mutant cells, the antibiotic producing strain can

lyse the cells in the lawn and overspread the plate, gradually displacing the original lawn of *B. subtilis*. From these types of observations we conclude that σ^W, by providing resistant to antimicrobial agents, contributes to the ability of *B. subtilis* cells to maintain their niche in the competitive microenvironment of the soil[17].

Although we have made significant headway in our understanding of σ^W and its role, many questions remain. There are some hints that σ^W may control the production of one or more antimicrobial compounds. Thus, exposure of cells to antimicrobial compounds would induce both defensive measures as well as offensive ones. The identity of the σ^W-regulated bacteriocin(s) is not yet clear. In addition, many other σ^W-controlled proteins could, in principle, play a role in resistance against toxic compounds or peptide antibiotics but the corresponding compounds have not yet been identified. These include, for example, membrane-localized peptidases that might degrade toxic peptides that must accumulate within the membrane to exert their toxic effects. Finally, as reviewed in the next section, we are only just beginning to understand the pathways by which the cell perceives the presence of toxic compounds and activates the σ^W stress response.

How are inducing signals interpreted and integrated?

Numerous stress conditions are known to activate the σ^W regulon including cell wall active antibiotics (vancomycin, cephalosporin, fosfomycin, D-cycloserine), membrane active compounds (nigericin, Triton-X-100), alkali shock, and toxic peptides and bacteriocins (SdpC, sublancin)[12,13,17,18,21]. What is not clear is how these signals are perceived and how this triggers activation of σ^W. The model that has emerged, to date, is reminiscent of that developed in detail in earlier studies of the *E. coli* ECF σ factor, σ^E.[22]

Activation of *E. coli* σ^E is triggered by unfolded proteins in the periplasm in a process that has been well characterized both genetically and biochemically[22]. In this case, σ^E is held in an inactive complex with a membrane-localized anti-σ factor, RseA. This protein, in turn, is degraded by a proteolytic cascade in response to inducing signals. The initial activating event is the cleavage of the transmembrane anti-σ on the periplasmic side of the membrane by a protease (a site I protease). This cleaved

anti-σ is then cleaved within the membrane by a site II protease (an example of regulated intramembrane proteolysis or RIP). This releases a soluble fragment of the anti-σ together with the σ factor into the cytosol. Further degradation of the remaining anti-σ fragment finally releases active σ^E protein which then binds RNA polymerase core enzyme to generate the functional holoenzyme[22].

The first component of the σ^W regulatory cascade to be defined was the membrane-localized anti-σ, RsiW. This protein, encoded by the gene immediately downstream of σ^W, is a negative regulator of σ^W activity[10,13]. Analysis of *B. subtilis* proteases with similarity to the known members of the *E. coli* cascade identified the relevant site II protease (YluC) and also defined a role for the ClpXP protease in degradation of the truncated RsiW : σ^W complex to release active σ^W.[23,24] However, the *B. subtilis* genome does not encode an obvious homolog of the site I protease that initiates the activation cascade. The identity of this protein emerged instead, from two different genetic screens.

The site I protease responsible for regulated σ^W activity was identified by Ellermeier and Losick during an analysis of factors contributing to SdpC resistance[19]. To identify factors that might be involved in SdpC resistance, these authors selected for spontaneous mutants that had enhanced growth in cells lacking SdpI. This screen led to the identification of null mutations in both RsiW and YsdB, two negative regulators of σ^W (*ysdB* had been identified using the transposon-based screen mentioned previously[14]). In addition, another class of dominant (gain-of-function) mutations were recovered with changes in a gene encoding a predicted, membrane-localized protein YpdC. These mutants, *ypdC**, generate constitutively active forms of a site I protease that cleaves RsiW exterior to the cell membrane, thereby initiating the proteolytic cascade needed for σ^W release. This gene was therefore renamed *prsW* (*p*rotease that *r*egulates *sigW*)[19].

This same gene was independently discovered by Heinrich and Wiegert by screening for transposon insertion (loss of function) mutations that increased the stability of RsiW[25]. In this case, they used an engineered RsiW fusion to the green fluorescent protein (GFP) to monitor levels of RsiW in whole cells by fluorescence. Among the transposon insertions with the strongest effect on stabilizing RsiW they recovered several with insertions in *ypdC*. A model summarizing our current understanding of the proteolytic

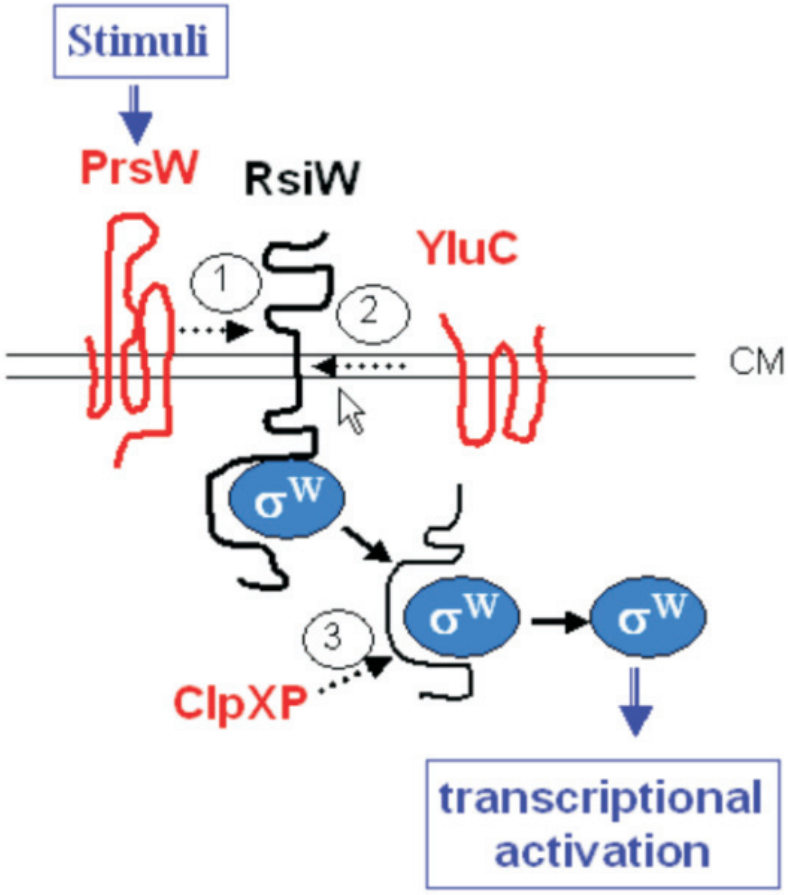

Fig. 3. *Scheme for the regulated intramembrane proteolysis of RsiW. Detection of appropriate inducing stimuli activates the σ^W regulon by a proteolytic cascade (steps 1, 2 and 3) that targets the RsiW anti-σ factor (black line). Induction is initiated by the PrsW protease which cleaves RsiW on the exterior face of the cytoplasmic membrane (CM). Once cleaved by PrsW, the initial cleavage product is further processed by the intramembrane protease YluC, and then the resulting cytosolic fragment is degraded by the ClpXP protease system. Release of σ^W allows binding to RNA polymerase core enzyme and transcriptional activation of the σ^W regulon. The processes that allow PrsW to detect cell envelope stress are not yet understood.*

cascade is presented in Figure 3. While this clearly represents a significant and satisfying advance, the actual mechanisms of signal perception remain elusive. How does the cell sense the presence of antibiotics that perturb cell wall or membrane function? Does PrsW interact directly with antimicrobial peptides or does it sense

an intermediate in cell wall synthesis? These and related questions are the topic of ongoing studies.

Conclusions and perspective

The work reviewed here began with a seemingly simple question: what is the function of the *sigW* gene? Since a null mutant had, in initial screens, no obvious phenotype we focused our attention on (i) the identification of target genes, and (ii) the identification of inducing signals. As these two parallel lines of investigation developed, it began to be clear that σ^W functions, in large part, to control the intrinsic resistance to a variety of antimicrobial compounds. This explains, to a large extent, why *sigW* mutants do not have an obvious phenotype under most laboratory growth conditions. In general, much of our knowledge of *B. subtilis* physiology has been gained by the exploration of processes that occur in monoculture (sporulation, competence, and numerous adaptive and stress responses). In contrast, the major role of the σ^W regulon is only apparent in mixed cultures that more closely mimic the situation in nature. In the complex microbial community found in the soil, the ability to resist the numerous antimicrobial compounds made by other bacteria (including even closely related organisms) is critical for maintaining one's niche.

The roles of the other six ECF σ factors encoded in the *B. subtilis* genome are still poorly understood. The σ^X factor controls at least two operons known to modulate cell surface charge and thereby contribute to antibiotic resistance[26]. The σ^M regulon, like σ^W, is activated by cell wall active antibiotics[12,21] and is responsible for the inducible synthesis of at least one antibiotic resistance gene[27,28]. The σ^Y regulon appears to contain a gene for a toxic peptide and another gene that encodes a potential immunity protein (similar to SdpI and YfhL)[29]. The other three ECF σ factors (σ^Z, σ^V, and σ^{ylaC}) are still poorly characterized, although they are the subject of ongoing studies[30–32].

Acknowledgements

Work in the Helmann laboratory on ECF σ factors and cell envelope stress responses is supported by the National Insitutes of Health (GM-047446).

References

1. Paget, M. S. and Helmann, J. D. (2003) The sigma70 family of sigma factors. *Genome Biol.*, **4**, 203.
2. Helmann, J. D. (2002) The extracytoplasmic function (ECF) sigma factors. *Adv. Microbiol. Physiol.*, **46**, 47–110.
3. Huang, X., Decatur, A., Sorokin, A. and Helmann, J. D. (1997) The *Bacillus subtilis* sigma(X) protein is an extracytoplasmic function sigma factor contributing to survival at high temperature. *J. Bacteriol.*, **179**, 2915–2921.
4. Huang, X. and Helmann, J. D. (1998) Identification of target promoters for the Bacillus subtilis sigma X factor using a consensus-directed search. *J. Molec. Biol.*, **279**, 165–173.
5. Qiu, J. and Helmann, J. D. (2001)The -10 region is a key promoter specificity determinant for the Bacillus subtilis extracytoplasmic-function sigma factors sigma(X) and sigma(W). *J. Bacteriol.*, **183**, 1921–1927.
6. Huang, X., Fredrick, K. L. and Helmann, J. D. (1998) Promoter recognition by Bacillus subtilis sigmaW: autoregulation and partial overlap with the sigmaX regulon. *J. Bacteriol.*, **180**, 3765–3770.
7. Huang, X., Gaballa, A., Cao, M. and Helmann, J. D. (1999) Identification of target promoters for the Bacillus subtilis extracytoplasmic function sigma factor, sigma W. *Molec. Microbiol.*, **31**, 361–371.
8. Hinton, J. C. (1997) The Escherichia coli genome sequence: the end of an era or the start of the FUN? *Molec. Microbiol.*, **26**, 417–422.
9. Cao, M., Kobel, P. A., Morshedi, M. M., Wu, M. F., Paddon, C. and Helmann, J. D. (2002) Defining the Bacillus subtilis sigma(W) regulon: a comparative analysis of promoter consensus search, run-off transcription/macroarray analysis (ROMA), and transcriptional profiling approaches. *J. Molec. Biol.*, **316**, 443–457.
10. Cao, M., Bernat, B. A., Wang, Z., Armstrong, R. N. and Helmann, J. D. (2001) FosB, a cysteine-dependent fosfomycin resistance protein under the control of sigma(W), an extracytoplasmic-function sigma factor in Bacillus subtilis. *J. Bacteriol.*, **183**, 2380–2383.
11. Sikder, D. and Kodadek, T. (2005) Genomic studies of transcription factor-DNA interactions. *Curr. Opin. Chem. Biol.*, **9**, 38–45.
12. Cao, M., Wang, T., Ye, R. and Helmann, J. D. (2002) Antibiotics that inhibit cell wall biosynthesis induce expression of the Bacillus subtilis sigma(W) and sigma(M) regulons. *Molec. Microbiol.*, **45**, 1267–1276.
13. Wiegert, T., Homuth, G., Versteeg, S. and Schumann, W. (2001) Alkaline shock induces the Bacillus subtilis sigma(W) regulon. *Molec. Microbiol.*, **41**, 59–71.
14. Turner, M. S. and Helmann, J. D. (2000) Mutations in multidrug efflux homologs, sugar isomerases, and antimicrobial biosynthesis genes differentially elevate activity of the sigma(X) and sigma(W) factors in Bacillus subtilis. *J. Bacteriol.*, **182**, 5202–5210.
15. Qian, Q., Lee, C. Y., Helmann, J. D. and Strauch, M. A. (2002) AbrB is a regulator of the sigma(W) regulon in Bacillus subtilis. *FEMS Microbiol. Lett.*, **211**, 219–223.

16. Gonzalez-Pastor, J. E., Hobbs, E. C. and Losick, R. (2003) Cannibalism by sporulating bacteria. *Science*, **301**, 510–513.
17. Butcher, B. G. and Helmann, J. D. (2006) Identification of Bacillus subtilis sigma-dependent genes that provide intrinsic resistance to antimicrobial compounds produced by Bacilli. *Molec. Microbiol.*, **60**, 765–782.
18. Ellermeier, C. D., Hobbs, E. C., Gonzalez-Pastor, J. E. and Losick, R. (2006) A three-protein signaling pathway governing immunity to a bacterial cannibalism toxin. *Cell*, **124**, 549–559.
19. Ellermeier, C. D. and Losick, R. (2006) Evidence for a novel protease governing regulated intramembrane proteolysis and resistance to antimicrobial peptides in Bacillus subtilis. *Genes Dev.*, **20**, 1911–1922.
20. Stein, T. (2005) Bacillus subtilis antibiotics: structures, syntheses and specific functions. *Molec. Microbiol.*, **56**, 845–857.
21. Pietiainen, M., Gardemeister, M., Mecklin, M., Leskela, S., Sarvas, M. and Kontinen, V. P. (2005) Cationic antimicrobial peptides elicit a complex stress response in Bacillus subtilis that involves ECF-type sigma factors and two-component signal transduction systems. *Microbiology*, **151**, 1577–1592.
22. Rowley, G., Spector, M., Kormanec, J. and Roberts, M. (2006) Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nat. Rev. Microbiol.*, **4**, 383–394.
23. Schobel, S., Zellmeier, S., Schumann, W. and Wiegert, T. (2004) The Bacillus subtilis sigmaW anti-sigma factor RsiW is degraded by intramembrane proteolysis through YluC. *Molec. Microbiol.*, **52**, 1091–105.
24. Zellmeier, S., Schumann, W. and Wiegert, T. (2006) Involvement of Clp protease activity in modulating the Bacillus subtilis sigmaw stress response. *Molec. Microbiol.*, **61**, 1569–1582.
25. Heinrich, J. and Wiegert, T. (2006) YpdC determines site-1 degradation in regulated intramembrane proteolysis of the RsiW anti-sigma factor of Bacillus subtilis. *Molec. Microbiol.*, **62**, 566–579.
26. Cao, M. and Helmann, J. D. (2004) The Bacillus subtilis extracytoplasmic-function sigmaX factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides. *J. Bacteriol.*, **186**, 1136–1146.
27. Cao, M., Moore, C. M. and Helmann, J. D. (2005) Bacillus subtilis paraquat resistance is directed by sigmaM, an extracytoplasmic function sigma factor, and is conferred by YqjL and BcrC. *J. Bacteriol.*, **187**, 2948–2956.
28. Cao, M. and Helmann, J. D. (2002) Regulation of the Bacillus subtilis bcrC bacitracin resistance gene by two extracytoplasmic function sigma factors. *J. Bacteriol.*, **184**, 6123–6129.
29. Cao, M., Salzberg, L., Tsai, C. S., Mascher, T., Bonilla, C., Wang, T., Ye, R. W., Marquez-Magana, L. and Helmann, J. D. (2003) Regulation of the Bacillus subtilis extracytoplasmic function protein sigma(Y) and its target promoters. *J. Bacteriol.*, **185**, 4883–4890.
30. Yoshimura, M., Asai, K., Sadaie, Y. and Yoshikawa, H. (2004) Interaction of Bacillus subtilis extracytoplasmic function (ECF) sigma factors with the N-terminal regions of their potential anti-sigma factors. *Microbiology*, **150**, 591–599.

31. Zellmeier, S., Hofmann, C., Thomas, S., Wiegert, T. and Schumann, W. (2005) Identification of sigma(V)-dependent genes of Bacillus subtilis. *FEMS Microbiol. Lett.*, **253**, 221–229.
32. Asai, K., Yamaguchi, H., Kang, C. M., Yoshida, K., Fujita, Y. and Sadaie, Y. (2003) DNA microarray analysis of Bacillus subtilis sigma factors of extracytoplasmic function family. *FEMS Microbiol. Lett.*, **220**, 155–160.
33. Helmann, J. D. and Moran, C. P. Jr. RNA polymerae and sigma factors. In: Sonenshein, A. L. and Losick, R. (eds.), *Bacillus subtilis and Its Relatives: From Genes to Cells*, pp. 289–312. ASM Press, Washington D.C. 2002.