UNIVERSITYOF BIRMINGHAM

Mathematical modelling of the sporulation-initiation network in Bacillus subtilis revealing the dual role of the putative quorum-sensing signal molecule PhrA

Jabbari, S.; King, J.R.; Heap, J.T.

Document Version
Peer reviewed version

[Link to publication on Research at Birmingham portal](#)

Mathematical modelling of the sporulation-initiation network in *Bacillus subtilis* revealing the dual role of the putative quorum-sensing signal molecule PhrA

Sara Jabbari[1], John T. Heap[2], John R. King[1]

[1] *Centre for Mathematical Medicine and Biology, School of Mathematical Sciences, University of Nottingham,*

Nottingham, NG7 2RD, UK

[2] *Centre for Biomolecular Sciences, University of Nottingham, Nottingham, NG7 2RD, UK*

January 28, 2010

Abstract

Bacillus subtilis cells may opt to forgo normal cell division and instead form spores if subjected to certain environmental stimuli, for example nutrient deficiency or extreme temperature. The resulting spores are extremely resilient and can survive for extensive periods of time, importantly under particularly harsh conditions such as those mentioned above. The sporulation process is highly time and energy consuming and essentially irreversible. The bacteria must therefore ensure that this route is only undertaken under appropriate circumstances. The gene regulation network governing sporulation initiation accordingly incorporates a variety of signals and is of significant complexity. We present a model of this network that includes four of these signals: nutrient levels, DNA damage, the products of the competence genes and cell population size. Our results can be summarised as follows: (i) the model displays the correct phenotypic behaviour in response to these signals; (ii) a basal level of *sda* expression may prevent sporulation in the presence of nutrients; (iii) sporulation is more likely to occur in a large population of cells than in a small one; (iv) finally, and of most interest, PhrA can act simultaneously as a quorum-sensing signal and as a timing mechanism, delaying sporulation when the cell has damaged DNA, possibly thereby allowing the cell time to repair its DNA before forming a spore.

Keywords: *Bacillus subtilis*, gene regulation networks, mathematical modelling, quorum sensing, sporulation.
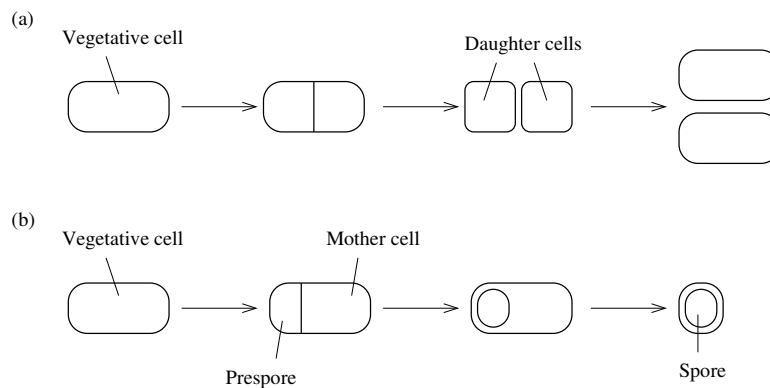
Figure 1: (a) The symmetrical cell division of vegetative growth: the contents of a parent cell are shared roughly equally between the two daughter cells. (b) Sporulation replaces the symmetric cell division of vegetative growth under appropriate conditions. The mother cell envelopes the prespore and manufactures a protective coating around the latter before dying. This coating enables the spore to tolerate harsh conditions for extensive periods of time.

1 Introduction

Bacillus subtilis is the classic model organism for Gram-positive bacteria [1, 2] (in particular those with a sporulation phenotype such as the anaerobic *Clostridium* species) being well-understood and easily genetically manipulated. Sporulation is a specialised type of cell division that enables a bacterium to survive extreme conditions, whereby the symmetrical cell division of vegetative growth is replaced by asymmetric division resulting in a mother cell and a prespore. The latter is engulfed by the mother cell, which manufactures a highly resistant coating around the newly formed spore before dying; see, for example, [3–5] or [6]. Figure 1 depicts this process schematically. The resulting spore can survive for extensive periods of time, which can be for many years.

Spores are resistant to a variety of harsh conditions (for example, those associated with nutrient limitation, pH or temperature) and, accordingly, sporulation is used by a number of bacteria for assorted reasons [1]. Thus the importance of understanding sporulation extends into many different fields. Most obvious, perhaps, is the importance of spores in the diseases caused by *Clostridium difficile*, *Clostridium botulinum* and *Bacillus anthracis* [3]; additionally, sporulation is associated with the production of valuable chemicals, including biofuels, by harmless bacteria like *Clostridium acetobutylicum* [7].

Sporulation requires a great deal of time and energy and is essentially irreversible [4], making it crucial for a cell to monitor its surroundings efficiently and ensure that sporulation is embarked upon at only the most appropriate times. The wrong decision can be catastrophic: a vegetative cell will die if the conditions are too harsh, while bacteria forming spores in an environment which is conducive to vegetative growth will be out-competed. Thus many different signals, each reflecting an aspect of the environment of the cell,

feed into an intricate network of regulatory genes and proteins that monitor the process at a number of stages, see for example [2] or [6]. The number of proteins and pathways involved in such a process often renders the network cumbersome and difficult to understand fully. Mathematical modelling is a useful tool for investigating gene regulation networks in order to analyse the roles of the various components involved in the overall process of what is essentially decision making.

There is a number of previous models concerning sporulation in *B. subtilis*. Jeong *et al.* [8] formulate a model of *B. subtilis* growth and sporulation at the metabolic level. There are accordingly points of contact between their model and ours but there is no real overlap between the two: while we concentrate on the proteins forming the sporulation-initiation network, [8] focusses on metabolic species such as glucose, acetyl-CoA and GTP (the last of which is the only variable to appear in both models). de Jong *et al.* [9] use a qualitative method to represent the sporulation-initiation network of *B. subtilis*, using inequalities, rather than estimated values, to represent parameters. Only one kinase is included and it is assumed that a nutrient-related signal influences this kinase. The *phr/rap* system is omitted, making Spo0E the only phosphatase acting on the phosphorelay, so that certain behavioural aspects may be missing (see Figure 2 for an indication of how the proteins mentioned here relate to the gene network governing sporulation initiation; further details will be provided in §2). Voigt *et al.* [10] concentrate on a smaller section of the network, namely the *sin* operon, and in particular on the effect of parameter variations on the behaviour of this operon. The level of phosphorylated Spo0A (Spo0A∼P) is taken to be a parameter in the model, meaning that the focus is on the effect of Spo0A∼P on the *sin* operon, rather than the inverse, in contrast to what follows. They demonstrate that subtly different parameter sets can produce four distinct behaviours - monostability, bistability (both of which also arise from our model), oscillations and a pulse generator - and focus on the fact that these are alterations that a cell could make without necessitating genetic mutations, giving a simple way to adapt to the varying requirements of different environmental conditions. Similarly Bischofs *et al.* [11] (which appeared in print during the preparation of this paper) model the Spo0 phosphorelay with two competing signals: one un-specified (but assumed to be related to nutrient levels) signal activating the kinases and one signal related to population size. A steady-state analysis of the resulting ordinary differential equations indicates that it is the ratio of these two signals which is ultimately responsible for either activating or suppressing the sporulation response, and this will be similarly addressed in our results.

Our approach is similar to that of [10] and [11] in that we represent the protein-protein interactions by deterministic ordinary differential equations. Our model covers almost the full known network responsible for detecting sporulation-related signals, relaying this information through the cell and initiating the first step of sporulation: it neglects only the less-understood components or those which are not entwined in any feedback sub-loop (we will clarify in §2.1 which aspects are missing). We also include studies of the time-dependent behaviour (in addition to steady-state analyses) in order to examine the dynamics of the system, crucially allowing us to shed light on the timing of the onset of sporulation. We thus present what we believe to be the most comprehensive model of this gene regulation network produced so far. We include

State	Population size	DNA	Competent	Nutrients	Sporulation desirable?
1	Small	Healthy	No	High	No
2	Small	Healthy	No	Low	Yes
3	Small	Healthy	Yes	High	No
4	Small	Healthy	Yes	Low	No
5	Small	Damaged	No	High	No
6	Small	Damaged	No	Low	No
7	Small	Damaged	Yes	High	No
8	Small	Damaged	Yes	Low	No
9	Large	Healthy	No	High	No
10	Large	Healthy	No	Low	Yes
11	Large	Healthy	Yes	High	No
12	Large	Healthy	Yes	Low	Yes
13	Large	Damaged	No	High	No
14	Large	Damaged	No	Low	No
15	Large	Damaged	Yes	High	No
16	Large	Damaged	Yes	Low	No

Table 1: The possible qualitative combinations of the four different signals and whether or not the cell ought to enter sporulation in each of these states, see §2.2-§2.4. In §4 we see that our model reproduces the phenotype for all inputs, except State 14 for which our model does give potentially sporulation-inducing levels of Spo0A∼P but, importantly, only after a significant time delay. An explanation of why we expect sporulation in State 12 but not State 4 is given in §4.3.

four distinct signals: nutrient levels, DNA damage, population size and the products of the competence genes, and are thus able to capture a variety of signal types including environmental and metabolic ones, in addition to one directly related to the cell cycle. Competence has been proposed as a survival mechanism, distinct from (and incompatible with) sporulation, whereby a bacterium can take up DNA from an adjacent cell or its environment, providing a relatively fast means to adapt genetically to the environment (thus by competence we refer to natural competence, as opposed to artificial competence that is induced under laboratory conditions).

Certain combinations of these signals will provide an environment in which sporulation is the best option for survival, see Table 1. Of the four signals, it is a lack of nutrients that would result in vegetative growth being unsustainable. If the cell has damaged DNA, sporulation (and indeed any cell division)

should be prevented since the damage would be inherited by the spore (or daughter cells) where it would be irreparable. Competence could serve to prevent sporulation since the two cannot occur simultaneously, though sporulation would be beneficial in more extreme circumstances (while competence can make a cell fitter, sporulation is a matter of life and death) and may therefore be able to override the competence genes. Regarding population size, we include the PhrA protein which has been postulated as a possible quorum-sensing signal [12] (quorum sensing is a cell-cell communication mechanism allowing a population of bacteria to monitor its size and coordinate its behaviour accordingly). There is some debate, however, as to whether PhrA is in fact a quorum-sensing signal or simply a protein that enforces a time delay in the sporulation-initiation process [13–15] (for simplicity, however, we refer to this signal throughout by 'population size'). We are able to identify a suitable parameter set and investigate the interactions between these four competing signals and the implications for spore formation. The results of our model are consistent with the biological hypothesis that sporulation occurs only under the correct combination of these signals. Importantly, we are able to demonstrate that the PhrA protein can, where appropriate, act simultaneously as a means of communication with other cells and as a timing device. The implications of this will be discussed in §5.

2 Model development

2.1 Preliminaries

We seek here to incorporate the known components of the gene regulation network governing sporulation initiation in a single *B. subtilis* cell; the model can easily be scaled up to the population level if desired. We deduce the network (see Figure 2) largely from [1], where regulatory proteins both upstream and downstream of sporulation initiation are discussed (we include the upstream aspects only), but include additional details from other literature. Certain components, such as those associated with *kipI/kipA*, are neglected from the model as there is no known feedback into their regulation from other components in the network, and any effect they have on the network itself can be absorbed into the parameter choice (KipI inhibits *kinA* transcription, and KipA that of *kipI*, but we do not have information on the effect of other components of the network on KipI or KipA). Other neglected components will be identified in subsequent sections.

The network incorporates a number of different constituents, starting with the signals which feed into the kinases. Although five kinases have been identified which are capable of initiating sporulation in *B. subtilis* [16] (KinA, B, C, D and E), little is known about the signals that may promote or prevent activation of these kinases. For this reason we include only two kinases (KinA and KinB), as we have information on how two specific sporulation-inducing signals affect the activation of these two proteins [17, 18]. In addition, of the five, these are the two proteins which dominate the sporulation phenotype [16, 19]. The kinases initiate a phosphorylation cascade through the phosphorelay proteins (Spo0F, Spo0B and Spo0A), the cascade being monitored by a number of phosphatases and transcription activators and repressors, see

Figure 2, that provide checkpoints at each step. Phosphorylated Spo0A is the ultimate response regulator of the system, which, in addition to downregulating *abrB* (a gene encoding a sporulation repressor), upregulates a number of genes downstream of this system that encode the proteins required for spore formation (those illustrated in Figure 2 principally affect the decision of whether or not to enter the sporulation phase rather than the process of spore formation itself).

In developing the model we adopt the following general assumptions.

- The contents of the cell are well-mixed and of a high copy number so that we can neglect any spatial and stochastic considerations and model the system using ordinary differential equations.

- All mRNA is translated into its protein, so that we do not need to track the concentration of mRNA separately from that of the protein itself, i.e. we take the reactions governing mRNA concentration to be in quasi-steady state. A consequence of this is that within the model any transcriptional regulation impacts directly on the corresponding protein translation, e.g. protein X inhibiting transcription of gene y is reflected in the model by protein X inhibiting translation of protein Y (i.e. the protein product of gene y). This asymptotic reduction is an acceptable approximation given that we assume all mRNA is translated into its corresponding protein.

- Sigma factors are subunits of bacterial polymerase which attach to the core polymerase and are in effect responsible for recognising the promoter sequence on DNA and initiating transcription. Different promoters require different sigma factors and we assume that plentiful supplies of all sigma factors except σ^H are available for transcription and we therefore do not track their concentrations. This is in line with [20], where σ^A (σ^A and σ^H are the principal sigma factors involved in the sporulation initiation network) levels are shown to be constant through the transition from vegetative growth to sporulation, and with [9], where σ^A is taken to be of constant concentration. By contrast, we track the concentration of σ^H as part of our model because its transcription is regulated by AbrB, another component of the model.

- Protein oligomer formation occurs at a relatively fast rate so that we can take the differential equation representing the corresponding protein in its monomer form to be quasi-steady, i.e. monomer levels are quasi-steady functions of the oligomer ones.

- All phosphorylated proteins can spontaneously dephosphorylate and, unless stated otherwise, all protein complexes can separate.

A detailed analysis of the consequences of these assumptions is beyond the scope of this work; however, we refer the reader to [21], which examines such assumptions in models of general signal transduction mechanisms. We split our discussion of the remaining assumptions into sub-sections covering different aspects of the network, as illustrated by the divisions in Figure 2 (in Figure 3 we also present a simplified version of this network for clarity). These are the sporulation-related signals, the kinases, the phosphorelay (which relays and integrates information from the signals to Spo0A), the transition-state regulators (AbrB
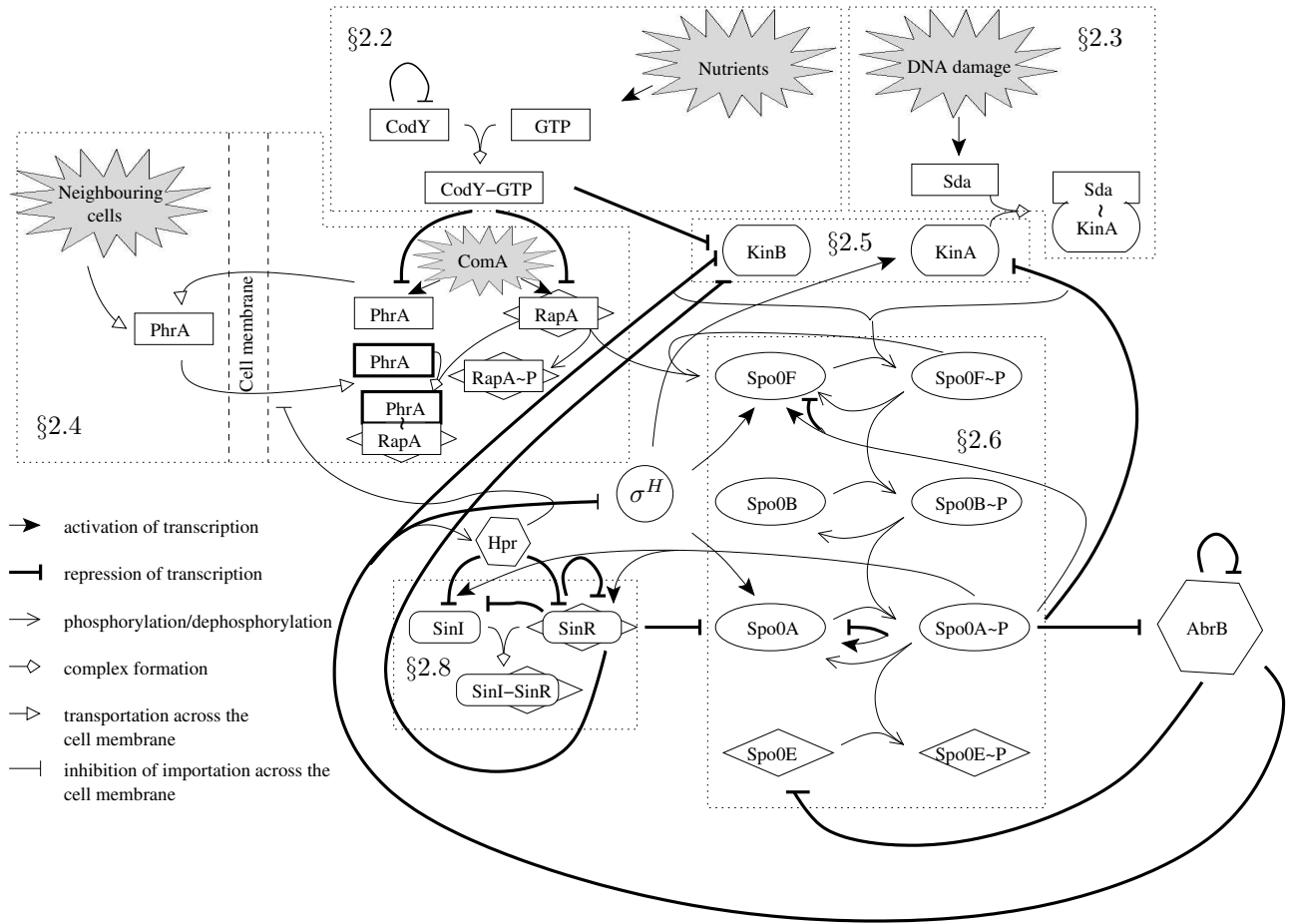
Figure 2: A schematic representation of the network controlling the initiation of sporulation in *B. subtilis*. The dotted lines divide the network into the sub-networks discussed in §2 (each box has been labelled with the appropriate section number) and different shapes have been used to illustrate the different types of components of the network: ellipses represent the phosphorelay, diamonds are used for inhibitors of this phosphorelay, rectangles for the molecules that mediate the signals, which are themselves illustrated by the shaded jagged star shapes (RapA both mediates a signal and inhibits the phosphorelay), rectangles with rounded ends are the Sin proteins (SinR is also an inhibitor of the phosphorelay), hexagons the transition state regulators and the circle represents the only sigma factor explicitly included in the model. All binding reactions except that of SinI/SinR are assumed to be reversible, as are the phosphorylation reactions in the phosphorelay. Although we have not included this explicitly in the diagram, the kinases autophosphorylate and subsequently transfer their phosphate to Spo0F (KinA can also directly phosphorylate Spo0A). All phosphorylated proteins (represented by ∼P) are subject to spontaneous dephosphorylation. For a simplified version of this diagram see Figure 3.

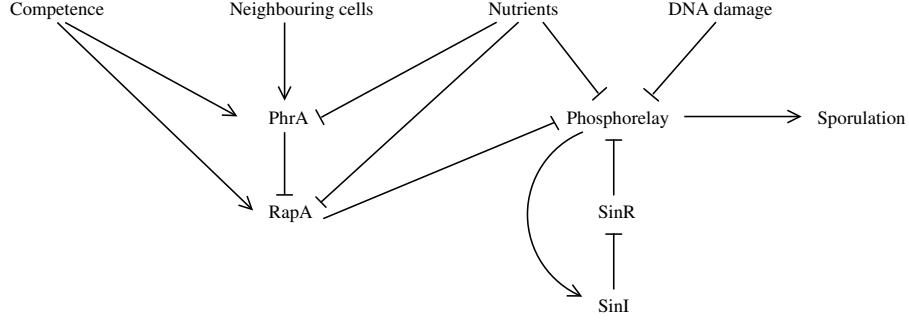Figure 3: A simplified version of Figure 2 in order to give an overview of how the signal pathways interact.

and Hpr), the *sin* operon and, finally, the sigma factor σ^H. Definitions of the parameters and variables used throughout are given in Tables 2 and 3 respectively.

2.2 Nutrient levels

The initiation of sporulation is controlled by *B. subtilis* in order that this process be limited to the correct environmental and cellular conditions, corresponding to a suitable combination of mixed signals [22]. It is favourable for *B. subtilis* to produce spores when nutrient levels are low as vegetative cells would starve when nutrients become exhausted.

GTP is a measure of the nutritional conditions: a plentiful nutrient supply leads to high levels of GTP, which can bind to the regulatory protein CodY [23]; we assume that this complex can spontaneously separate. CodY is a DNA binding protein [18] with a basal transcription level (in particular, σ^H is not required for its transcription) that is self-regulated, and for this we assume GTP is not required so that CodY can self-regulate regardless of nutrient levels (whereas in the regulation of other proteins involved in this network, we assume that CodY must be GTP-bound so that its ability to regulate its target genes correlates with nutrient supply).

We derive the following equations to represent the sporulation-related signal GTP and the GTP-mediating protein CodY in both its un-bound and GTP-bound form:

$$\frac{dG}{dt} = c_G - \lambda_G G - \beta_{C_Y^G} C_Y G + \gamma_{C_Y^G} C_Y^G, \tag{1}$$

$$\frac{dC_Y}{dt} = \frac{c_{C_Y} U_{C_Y}^{C_Y}}{B_{C_Y}^{C_Y} C_Y + U_{C_Y}^{C_Y}} - \beta_{C_Y^G} C_Y G + \gamma_{C_Y^G} C_Y^G - \lambda_{C_Y} C_Y, \tag{2}$$

$$\frac{dC_Y^G}{dt} = \beta_{C_Y^G} C_Y G - \gamma_{C_Y^G} C_Y^G - \lambda_{C_Y^G} C_Y^G, \tag{3}$$

with

$$G(0) = G_0, \qquad\qquad C_Y(0) = C_{Y0}, \qquad\qquad C_Y^G(0) = C_{Y0}^G. \tag{4}$$

8

Parameter	Rate of
c_X	production of variable X (if a protein then where its gene has only one promoter)
c_X^i	production of variable X directed by promoter i (where its gene has multiple promoters)
c_X^{il}	production of variable X directed by promoter i at some specified lower rate
c_X^{ih}	production of variable X directed by promoter i at some specified higher rate
c_Q	production of external PhrA by neighbouring cells in the population
B_X^Y	binding of the regulatory molecule X to a promoter site of the gene encoding for Y
U_X^Y	unbinding of the regulatory molecule X from a promoter site of the gene encoding for Y
β_Z	formation of complex Z
γ_Z	spontaneous separation of complex Z
λ_X	degradation of X
α_X	autophosphorylation of KinX
ϕ_X^Y	phosphotransfer from molecule X to molecule Y
δ_X	dephosphorylation by phosphatase X
ψ_X	spontaneous dephosphorylation of X
μ_e	externalisation (and therefore activation) of PhrA
μ_i	internalisation of PhrA (in its active form)
θ_H	concentration of Hpr at which internalisation of PhrA is half the maximal rate

Table 2: Definitions of the parameters required for the model described in §2. As a result of the equations governing mRNA concentration being taken to be quasi-steady, when we refer in this table to production directed by a promoter, we in fact mean protein translation from the mRNA whose transcription has been directed by that particular promoter.

Variable	Concentration
C_Y	CodY
C_Y^G	CodY∼GTP complex
G	GTP
S_d	Sda
P_i	internal PhrA
P_e	external PhrA
P_A	active and internalised PhrA
R_A	RapA
R_A^P	RapA∼P
P_A^R	PhrA∼RapA complex
K_A	KinA dimer
K_A^P	KinA∼P
K_A^S	Sda-bound KinA dimer
K_B	KinB
K_B^P	KinB∼P

Variable	Concentration
S_F	Spo0F
S_F^P	Spo0F∼P
S_B	Spo0B dimer
S_B^P	Spo0B∼P
S_A	Spo0A
S_A^P	Spo0A∼P dimer
S_E	Spo0E
S_E^P	Spo0E∼P
A	AbrB tetramer
H	Hpr
I	SinI
R	SinR tetramer
I_R	SinI-SinR complex
σ^H	the sigma factor σ^H

Table 3: Definitions of the variables used in the model described in §2. We have split the table into sections corresponding to those of §2.

Our specific choice of initial conditions will be discussed in §4.1. Notice that the Michaelis-Menten expression in the production term in (2) (and all equations henceforth which contain similar functions to represent production) stems from assuming protein binding and unbinding with DNA binding sites to be relatively fast.

2.3 DNA damage

When the DNA of a *B. subtilis* cell is damaged, transcription of *sda* is induced and the Sda protein is produced [24]. Cell division (both vegetative and that required for sporulation) is unfavourable if the cell has damaged DNA because successful replication of the chromosome is required, i.e. in terms of sporulation, both the mother cell and prespore must receive a full copy of the healthy chromosome. Sda binds to the KinA dimer to prevent autophosphorylation of the latter [17]. It may be possible for two molecules of Sda to bind to a KinA dimer [17], but we assume that one Sda protein suffices to block KinA phosphorylation. The equation representing Sda is given by

$$\frac{dS_d}{dt} = c_{S_d} - \lambda_{S_d} S_d - \beta_{K_A^S} S_d K_A + \gamma_{K_A^S} K_A^S, \tag{5}$$

with

$$S_d(0) = S_{d0}. \tag{6}$$

2.4 Competence and population size

Both competence and population size feed into the network through the *phr/rap* operon, and we thus group the discussion of these two signals into one section. The Rap proteins are a family of sporulation antagonists. Some are phosphatases acting on Spo0F∼P and others interfere with transcription via binding to regulatory proteins [25–27] while the Phr proteins modulate this activity [12, 15]. There are several proteins in the Phr/Rap family, but we focus our attention on PhrA and RapA due to a lack of information about how the others may feature in the network illustrated in Figure 2. *rapA* and *phrA* are contained within the same operon [28], and are transcribed at the same rate. Their transcription is σ^A-dependent and is inhibited by GTP-bound CodY [18] (thus giving nutrient levels an extra point of entry into the network) and activated by ComA [28], the response regulator of a two-component system responsible for detecting suitable conditions for the cell to become competent[1]. Thus a competent cell can increase RapA production to interrupt the phosphorylation cascade in an attempt to prevent sporulation (competence and sporulation being incompatible processes).

PhrA (whose production is also activated by ComA) is secreted from the cell and subsequently re-imported. During this transportation PhrA is processed into its active form as a pentapeptide [12,13]. Once

---

[1]The principal protein responsible for transforming the cell into a competent state is ComK, see [29] for a review of this process. ComK does not, to our knowledge, affect any components of the network in Figure 2 and is therefore not included in the model. However, for models of this protein and its role in competence-induction see [30] or [31].

active, PhrA inhibits the phosphatase activity of RapA on Spo0F∼P, most likely by binding to RapA [13,32], thus acting in favour of sporulation in the competition with the competence genes (despite the fact that PhrA is activated by ComA).

In theory, the transportation process should allow communication between cells since external PhrA could be imported into any *B. subtilis* cell, not just the one from which it was produced. Since a larger population would result in more external PhrA, this could be a means by which the cells could monitor the size of their population, making PhrA a quorum-sensing signal, although this has not been unambiguously shown experimentally thus far. As mentioned earlier, it has also been argued [13–15] that the Phr proteins are simply a control on the timing of sporulation initiation, with the export and import processes generating a delay between their transcription and their becoming active. We introduce the parameter c_Q to represent PhrA production from surrounding cells (large values of c_Q imply a large population of *B. subtilis* cells).

The import of PhrA is performed by the Opp group of proteins. Transcription of the *opp* operon is inhibited by Hpr. Rather than include each of the Opp proteins in this model, we treat import of PhrA as being directly affected by Hpr levels.

The following equations emerge from the above:

$$\frac{dR_A}{dt} = \frac{c_{R_A} U_{C_Y^G}^{R_A}}{B_{C_Y^G}^{R_A} C_Y^G + U_{C_Y^G}^{R_A}} - \delta_{R_A} R_A S_F^P - \beta_{P_A^R} R_A P_A + \gamma_{P_A^R} P_A^R + \psi_{R_A^P} R_A^P - \lambda_{R_A} R_A, \tag{7}$$

$$\frac{dR_A^P}{dt} = \delta_{R_A} R_A S_F^P - \psi_{R_A^P} R_A^P - \lambda_{R_A^P} R_A^P, \tag{8}$$

$$\frac{dP_i}{dt} = \frac{c_{R_A} U_{C_Y^G}^{R_A}}{B_{C_Y^G}^{R_A} C_Y^G + U_{C_Y^G}^{R_A}} - \mu_e P_i - \lambda_{P_i} P_i, \tag{9}$$

$$\frac{dP_e}{dt} = c_Q + \mu_e P_i - \frac{\theta_H \mu_i}{\theta_H + H} P_e - \lambda_{P_e} P_e, \tag{10}$$

$$\frac{dP_A}{dt} = \frac{\theta_H \mu_i}{\theta_H + H} P_e - \beta_{P_A^R} R_A P_A + \gamma_{P_A^R} P_A^R - \lambda_{P_A} P_A, \tag{11}$$

$$\frac{dP_A^R}{dt} = \beta_{P_A^R} R_A P_A - \gamma_{P_A^R} P_A^R - \lambda_{P_A^R} P_A^R, \tag{12}$$

with

$$R_A(0) = R_{A0}, \quad R_A^P(0) = R_{A0}^P, \quad P_i(0) = P_{i0}, \quad P_e(0) = P_{e0}, \quad P_A(0) = P_{A0}, \quad P_A^R(0) = P_{A0}^R. \tag{13}$$

2.5   The kinases

We include two kinases, KinA and KinB. Under conditions which are conducive to sporulation, these proteins autophosphorylate and transfer their phosphate to Spo0F, thus initiating the phosphorelay described in the following section.

kinA transcription is directed by a σ^H-dependent promoter [33] that is inhibited by Spo0A∼P [34] at a separate binding site to that of σ^H [35]. The protein KinA is active as a homodimer [36], so we will

assume that it can autophosphorylate only in its dimer form. As described in §2.3, DNA damage results in production of the protein Sda, which binds to the KinA dimer to prevent autophosphorylation of the latter, thus inhibiting the initiation of sporulation. The phosphorylated KinA dimer transfers its phosphate principally to Spo0F [25], but also to Spo0A [19] at a much slower rate (so that $\phi_{K_A^P}^{S_A} << \phi_{K_A^P}^{S_F}$). Kinases often also display phosphatase activity on their target protein but it has been demonstrated that this is not the case for KinA on Spo0F~P [36]. We assume that this is also the case for Spo0A~P and take the above reactions to be irreversible.

kinB transcription is directed by a σ^A-dependent promoter [19] and its expression is inhibited by GTP-bound CodY [18], the SinR tetramer [37] (which will be discussed in §2.8) or the sporulation repressor AbrB [37, 38]. We will assume that all three of these regulatory proteins bind at distinct sites (though it has yet to be shown that the SinR tetramer binds directly to a DNA binding site). KinB is a transmembrane protein [37] but, to keep the model as simple as possible, we will assume that there is no distinction between transmembrane and cytoplasmic KinB, and we therefore do not include distinct variables to represent each form. Once KinB has autophosphorylated, it acts only on Spo0F [25]; given no evidence to the contrary, we will assume this phosphate transfer to be reversible. KinA and KinB have a similar affinity for Spo0F [16] ($\phi_{K_A^P}^{S_F} \approx \phi_{K_B^P}^{S_F}$).

We thus obtain the following equations to represent the various forms of KinA and KinB:

$$\frac{dK_A}{dt} = \frac{1}{2}\left(\frac{c_{K_A} B_{\sigma^H}^{K_A} U_{S_A^P}^{K_A} \sigma^H}{(B_{S_A^P}^{K_A} S_A^P + U_{S_A^P}^{K_A})(B_{\sigma^H}^{K_A} \sigma^H + U_{\sigma^H}^{K_A})}\right) - \alpha_A K_A - \beta_{K_A^S} S_d K_A + \gamma_{K_A^S} K_A^S$$

$$+ \phi_{K_A^P}^{S_F} K_A^P S_F + \phi_{K_A^P}^{S_A} K_A^P S_A + \psi_{K_A^P} K_A^P - \lambda_{K_A} K_A, \tag{14}$$

$$\frac{dK_A^P}{dt} = \alpha_A K_A - \phi_{K_A^P}^{S_F} K_A^P S_F - \phi_{K_A^P}^{S_A} K_A^P S_A - \psi_{K_A^P} K_A^P - \lambda_{K_A^P} K_A^P, \tag{15}$$

$$\frac{dK_A^S}{dt} = \beta_{K_A^S} S_d K_A - \gamma_{K_A^S} K_A^S - \lambda_{K_A^S} K_A^S, \tag{16}$$

$$\frac{dK_B}{dt} = \frac{c_{K_B} U_{C_Y^G}^{K_B} U_A^{K_B} U_R^{K_B}}{(B_{C_Y^G}^{K_B} C_Y^G + U_{C_Y^G}^{K_B})(B_A^{K_B} A + U_A^{K_B})(B_R^{K_B} R + U_R^{K_B})}$$

$$- \alpha_B K_B + \phi_{K_B^P}^{S_F} K_B^P S_F - \phi_{S_F^P}^{K_B} K_B S_F^P + \psi_{K_B^P} K_B^P - \lambda_{K_B} K_B, \tag{17}$$

$$\frac{dK_B^P}{dt} = \alpha_B K_B - \phi_{K_B^P}^{S_F} K_B^P S_F + \phi_{S_F^P}^{K_B} K_B S_F^P - \psi_{K_B^P} K_B^P - \lambda_{K_B^P} K_B^P. \tag{18}$$

We assume that dimer formation by KinA is rapid in comparison with other reactions, meaning that changes in the expression of the monomer can be directly interpreted as changes in dimer concentration. The scaling of the first term in (14) by 1/2 results from two monomers being required to form a dimer. The same principle applies to all other proteins existing in oligomer form (namely Spo0B, AbrB and SinR).

For the kinases we impose the initial conditions

$$K_A(0) = K_{A0}, \qquad K_A^P(0) = K_{A\,0}^P, \qquad K_A^S(0) = K_{A\,0}^S, \qquad K_B(0) = K_{B0}, \qquad K_B^P(0) = K_{B\,0}^P. \tag{19}$$

2.6 The phosphorelay

Signals are transmitted to the appropriate regulatory proteins in the cell via a multicomponent phosphorelay [39]. The kinases phosphorylate Spo0F, which transfers its phosphate to Spo0B which, in turn, donates the phosphate to Spo0A, the ultimate response regulator of the system [39] (it is Spo0A~P, i.e. phosphorylated Spo0A, which downregulates the sporulation repressor AbrB and activates the sporulation genes downstream of this initiation network). Each phosphate transfer between these Spo proteins is reversible [39] and KinA~P, to a much lesser extent ($\phi_{K_A^P}^{S_A} << \phi_{K_A^P}^{S_F}$), can act directly on Spo0A [19].

Transcription of both *spo0F* and *spo0A* is directed by two promoters [40, 41], one promoter being σ^A-dependent and the other σ^H-dependent [33, 40, 42]. In both, the σ^A-dependent promoter is inhibited by Spo0A~P [43, 44], while the σ^H-dependent one is activated by Spo0A~P [42–44]. For *spo0A*, the σ^H-dependent promoter can also be inhibited by the SinR tetramer [45]. For both *spo0F* and *spo0A* we assume that the σ^A and σ^H binding sites are distinct, but for *spo0A* those of σ^H and SinR overlap [45] (the latter assumption results in there being only two factors on the denominator of the second term of (24) instead of three). *spo0B* transcription, on the other hand, is directed by a single promoter and does not require σ^H [46]. The protein, Spo0B, is active only as a dimer [47]; Spo0B~P thus refers to two Spo0B molecules and one phosphate.

Specific phosphatases work on the phosphorelay for additional regulation: namely the Rap proteins on Spo0F~P [32] and Spo0E on Spo0A~P [48]. We will assume that these reactions are irreversible. *spo0E* transcription is directed by a σ^A-dependent promoter [49], which is inhibited by the AbrB tetramer [49].

Spo0A forms a dimer upon phosphorylation and reverts to two monomers [50] upon dephosphorylation. We assume that this dimerisation occurs once one Spo0A protein has been phosphorylated (unphosphorylated Spo0A is much less likely to form dimers [50]), so that a single phosphate is then shared by two Spo0A proteins.

The following equations follow from the above considerations.

$$\frac{dS_F}{dt} = \frac{c_{S_F}^1 U_{S_A^P}^{S_F}}{B_{S_A^P}^{S_F} S_A^P + U_{S_A^P}^{S_F}} + \frac{c_{S_F}^{2l} B_{\sigma^H}^{S_F} U_{S_A^P}^{S_F} \sigma^H + c_{S_F}^{2h} B_{S_A^P}^{S_F} B_{\sigma^H}^{S_F} S_A^P \sigma^H}{(B_{S_A^P}^{S_F} S_A^P + U_{S_A^P}^{S_F})(B_{\sigma^H}^{S_F} \sigma^H + U_{\sigma^H}^{S_F})} - \phi_{K_A^P}^{S_F} K_A^P S_F$$

$$- \phi_{K_B^P}^{S_F} K_B^P S_F + \phi_{S_F^P}^{K_B} K_B S_F^P + \delta_{R_A} R_A S_F^P + \phi_{S_F^P}^{S_B} S_F^P S_B - \phi_{S_B^P}^{S_F} S_F S_B^P + \psi_{S_F^P} S_F^P - \lambda_{S_F} S_F, \qquad (20)$$

$$\frac{dS_F^P}{dt} = \phi_{K_A^P}^{S_F} K_A^P S_F + \phi_{K_B^P}^{S_F} K_B^P S_F - \phi_{S_F^P}^{K_B} K_B S_F^P - \delta_{R_A} R_A S_F^P - \phi_{S_F^P}^{S_B} S_F^P S_B + \phi_{S_B^P}^{S_F} S_F S_B^P - \psi_{S_F^P} S_F^P - \lambda_{S_F^P} S_F^P,$$

$$\qquad (21)$$

$$\frac{dS_B}{dt} = \frac{c_{S_B}}{2} - \phi_{S_F^P}^{S_B} S_F^P S_B + \phi_{S_B^P}^{S_F} S_F S_B^P + \phi_{S_B^P}^{S_A} S_B^P S_A - \phi_{S_A^P}^{S_B} S_B S_A^P + \psi_{S_B^P} S_B^P - \lambda_{S_B} S_B, \qquad (22)$$

$$\frac{dS_B^P}{dt} = \phi_{S_F^P}^{S_B} S_F^P S_B - \phi_{S_B^P}^{S_F} S_F S_B^P - \phi_{S_B^P}^{S_A} S_B^P S_A + \phi_{S_A^P}^{S_B} S_B S_A^P - \psi_{S_B^P} S_B^P - \lambda_{S_B^P} S_B^P, \qquad (23)$$

$$\frac{dS_A}{dt} = \frac{c^1_{S_A} U^{S_A}_{S^P_A}}{B^{S_A}_{S^P_A} S^P_A + U^{S_A}_{S^P_A}} + \frac{c^{2l}_{S_A} B^{S_A}_{\sigma^H} U^{S_A}_{S^P_A} U^{S_A}_R \sigma^H + c^{2h}_{S_A} B^{S_A}_{S^P_A} B^{S_A}_{\sigma^H} U^{S_A}_R S^P_A \sigma^H}{(B^{S_A}_{S^P_A} S^P_A + U^{S_A}_{S^P_A})(B^{S_A}_{\sigma^H} U^{S_A}_R \sigma^H + B^{S_A}_R U^{S_A}_{\sigma^H} R + U^{S_A}_{\sigma^H} U^{S_A}_R)}$$

$$- 2\phi^{S_A}_{S^P_B} S^P_B S_A + 2\phi^{S_B}_{S^P_A} S_B S^P_A - 2\phi^{S_A}_{K^P_A} K^P_A S_A + 2\delta_{S_E} S_E S^P_A + 2\psi_{S^P_A} S^P_A - \lambda_{S_A} S_A, \quad (24)$$

$$\frac{dS^P_A}{dt} = \phi^{S_A}_{S^P_B} S^P_B S_A - \phi^{S_B}_{S^P_A} S_B S^P_A + \phi^{S_A}_{K^P_A} K^P_A S_A - \delta_{S_E} S_E S^P_A - \psi_{S^P_A} S^P_A - \lambda_{S^P_A} S^P_A, \quad (25)$$

$$\frac{dS_E}{dt} = \frac{c_{S_E} U^{S_E}_A}{B^{S_E}_A A + U^{S_E}_A} - \delta_{S_E} S_E S^P_A + \psi_{S^P_E} S^P_E - \lambda_{S_E} S_E, \quad (26)$$

$$\frac{dS^P_E}{dt} = \delta_{S_E} S_E S^P_A - \psi_{S^P_E} S^P_E - \lambda_{S^P_E} S^P_E, \quad (27)$$

with

$$S_F(0) = S_{F0}, \quad S^P_F(0) = S^P_{F0}, \quad S_B(0) = S_{B0}, \quad S^P_B(0) = S^P_{B0},$$
$$S_A(0) = S_{A0}, \quad S^P_A(0) = S^P_{A0}, \quad S_E(0) = S_{E0}, \quad S^P_E(0) = S^P_{E0}. \quad (28)$$

2.7 Transition-state regulators

AbrB and Hpr (also known as ScoC) are both transition-state regulators (i.e. they mediate the production of a number of proteins during the transition between vegetative growth and sporulation) although the former is generally considered to be more influential in determining whether or not a cell enters the sporulation process. AbrB is a repressor of many of the genes required for sporulation [51] (including some that are part of the sporulation-initiation network) and inhibition of AbrB is therefore one of the ultimate aims of the active phosphorelay. Transcription of *abrB* is directed by two σ^A-dependent promoters [52], one being inhibited by AbrB itself [52] and the other by Spo0A~P; these bind to distinct sites and their binding abilities are independent [53].

AbrB is active as a tetramer or a dimer, binding to *spo0E* in either form; however [54] suggests that it is most likely that AbrB is active *in vivo* as a tetramer and accordingly we include only this AbrB oligomer in the model. Transcription of *hpr* is positively regulated by AbrB [55], presumably in its oligometric state, while the Hpr protein inhibits transcription of the *sin* and *opp* operons, along with other genes which do not form part of our sporulation-initiation network.

The above considerations yield the following equations.

$$\frac{dA}{dt} = \frac{1}{4}\left(\frac{c^1_A U^A_A}{B^A_A A + U^A_A} + \frac{c^2_A U^A_{S^P_A}}{B^A_{S^P_A} S^P_A + U^A_{S^P_A}}\right) - \lambda_A A, \quad (29)$$

$$\frac{dH}{dt} = \frac{c_H B^H_A A}{B^H_A A + U^H_A} - \lambda_H H, \quad (30)$$

with

$$A(0) = A_0, \qquad\qquad\qquad H(0) = H_0. \qquad\qquad (31)$$

2.8 The *sin* operon

The *sin* operon provides regulation at two levels: it affects the kinases and the response regulator, Spo0A~P, via the transcription of *kinB* and *spo0A* respectively. It contains two genes, *sinI* and *sinR*, and has three promoters: P1 is σ^A-dependent and functional for both genes, P2 is active only after sporulation so can be neglected from our model, and P3 is σ^A-dependent and produces only SinR. Transcription driven by P1 is activated by phosphorylated Spo0A and inhibited by both the SinR tetramer and Hpr, all at separate binding sites [56, 57]. In addition, while the AbrB tetramer is capable of binding to the *sin* operon, we neglect this due to its binding ability here being relatively weak [57].

As indicated above, SinR is active as a tetramer [58], composed of two SinR dimers [59]. The dimer bond is much stronger than the tetramer bond [59] so we will assume that, while tetramers can spontaneously separate to give two dimers, the dimers cannot split into two monomers. SinI, on the other hand, is active as a monomer [59] and for this reason we consider only the possibility that it exists in a monomer form in the model.

While SinR inhibits the initiation of sporulation by slowing down transcription of *kinB* and *spo0A* (see §2.5 and §2.6), SinI serves to disrupt this inhibition by forming a complex with SinR to render the latter inactive [60]. In [10] it is assumed that SinI can bind only to a SinR monomer; however, there is evidence that it can also displace a SinR monomer from its tetramer and bind to it, leaving a heterodimer and a SinR monomer and homodimer [59, 61]. We therefore include both possibilities in our model. The interaction between SinI and SinR is very tight and hence essentially irreversible [59].

Regulation of the Sin proteins is therefore described by the following equations:

$$\frac{dI}{dt} = \frac{c_I B_{S_A^P}^I U_R^I U_H^I S_A^P}{(B_{S_A^P}^I S_A^P + U_{S_A^P}^I)(B_R^I R + U_R^I)(B_H^I H + U_H^I)} - \beta_{I_R} IR - \lambda_I I, \qquad (32)$$

$$\frac{dR}{dt} = \frac{1}{4}\left(\frac{c_R^1 B_{S_A^P}^R U_R^R U_H^R S_A^P}{(B_{S_A^P}^R S_A^P + U_{S_A^P}^R)(B_R^R R + U_R^R)(B_H^R H + U_H^R)} + c_R^3\right) - \frac{1}{4}\beta_{I_R} IR - \lambda_R R, \qquad (33)$$

$$\frac{dI_R}{dt} = \beta_{I_R} IR - \lambda_{I_R} I_R, \qquad (34)$$

with

$$I(0) = I_0, \qquad\qquad R(0) = R_0, \qquad\qquad I_R(0) = I_{R0}. \qquad (35)$$

2.9 Sigma factors

As discussed previously, we assume all sigma factors are present in abundance except for σ^H, which connects a number of components of the sporulation-initiation network. The transcription of *spo0H* (the gene from which the protein σ^H is produced) is directed by a σ^A-dependent promoter [62] that is inhibited by the transition-state regulator AbrB [38, 62]. We thus take

$$\frac{d\sigma^H}{dt} = \frac{c_{\sigma^H} U_A^{\sigma^H}}{B_A^{\sigma^H} A + U_A^{\sigma^H}} - \lambda_{\sigma^H} \sigma^H, \tag{36}$$

with

$$\sigma_H(0) = \sigma_{H0}. \tag{37}$$

Equations (1)-(13),(15)-(18),(14)-(37) form our model of the sporulation-initiation network in a *B. subtilis* cell.

3 Parameter choice

3.1 Signal levels

In order to investigate the effects of the four signals on the sporulation-initiation network, we focus on the roles of four parameters:

- c_G, GTP production rate (associated with nutrient levels);

- c_{S_d}, Sda production rate (associated with DNA damage);

- c_{R_A}, (ComA-activated) production rate of RapA/PhrA (associated with competence);

- c_Q, production rate of external PhrA by surrounding cells (associated with population size).

We discuss the range over which they will be varied in the subsequent section concerning production rates. The remaining parameters will be kept fixed, unless otherwise stated, and their default values are displayed in Table 4. We next briefly discuss our reasoning behind the choices.

3.2 Production rates

We largely take our transcription and translation rates from Voigt *et al.* [10] where transcription from the *sin* operon is taken to be either 0.15 nM sec^{-1} or 0.28 nM sec^{-1}; for simplicity, we will assume all transcription occurs at 0.2 nM sec^{-1}. In [10] mRNA degradation is taken to occur at 0.005 sec^{-1} (based on the two minute half-life quoted in [63]). We have assumed mRNA dynamics to be quasi-steady in our model, implying that mRNA levels for all variables are equal to the transcription rate divided by the degradation rate, i.e. 40 nM (we recall that we treat any regulation as occurring at the protein level rather than at the mRNA level).

Rate	Parameter	Default Value	Units
Production	$c_{C_Y}, c^1_{S_F}, c^{2l}_{S_F}, c_{S_B}, c^1_{S_A}, c^{2l}_{S_A}, c_{S_E}, c^1_A, c^2_A, c_{\sigma_H}, c_H$	0.4	nM sec^{-1}
	c^1_R	0.42	nM sec^{-1}
	c^3_R	0.783	nM sec^{-1}
	c_{K_B}	10	nM sec^{-1}
	c_I	24	nM sec^{-1}
	$c_{K_A}, c^{2h}_{S_F}, c^{2h}_{S_A}$	40	nM sec^{-1}
DNA binding and unbinding	U^Y_X / B^Y_X for all relevant X, Y	20	nM
Complex separation	$\gamma_{C^G_Y}, \gamma_{K^S_A}$	0.1	sec^{-1}
	$\gamma_{P^R_A}$	1	sec^{-1}
Complex formation	$\beta_{C^G_Y}$	10^{-4}	nM^{-1} sec^{-1}
	$\beta_{K^S_A}, \beta_{I_R}, \beta_{P^R_A}$	0.083	nM^{-1} sec^{-1}
Degradation	$\lambda_{S^P_A}$	10^{-5}	sec^{-1}
	λ_{σ_H}	4×10^{-4}	sec^{-1}
	$\lambda_{C_Y}, \lambda_{C^G_Y}, \lambda_{K_A}, \lambda_{K^P_A}, \lambda_{K^S_A}, \lambda_{K_B}, \lambda_{K^P_B}, \lambda_{S_F}, \lambda_{S^P_F}, \lambda_{S_B}, \lambda_{S^P_B}, \lambda_{S_A}, \lambda_A, \lambda_R, \lambda_{I_R}, \lambda_{R_A}, \lambda_{R^P_A}, \lambda_{P^R_A}, \lambda_H$	0.002	sec^{-1}
	$\lambda_G, \lambda_{S_d}, \lambda_{S_E}, \lambda_{S^P_E}, \lambda_I, \lambda_{P_i}, \lambda_{P_e}, \lambda_{P_A}$	0.02	sec^{-1}
Autophosphorylation	α_A, α_B	0.1	sec^{-1}
Phosphotransfer	$\phi^{S_A}_{K^P_A}, \phi^{K_B}_{S^P_F}$	1.7×10^{-8}	nM^{-1} sec^{-1}
	$\phi^{S_F}_{K^P_A}, \phi^{S_F}_{K^P_B}, \phi^{S_B}_{S^P_F}, \phi^{S_F}_{S^P_B}, \phi^{S_A}_{S^P_B}, \phi^{S_B}_{S^P_A}$	10^{-6}	nM^{-1} sec^{-1}
Dephosphorylation	$\delta_{R_A}, \delta_{S_E}$	4×10^{-4}	nM^{-1} sec^{-1}
	$\psi_{K^P_A}, \psi_{K^P_B}, \psi_{S^P_F}, \psi_{S^P_B}, \psi_{S^P_A}, \psi_{S^P_E}, \psi_{R^P_A}$	4×10^{-4}	sec^{-1}
Import/export	μ_e, μ_i	1	sec^{-1}
	θ_H	1	nM

Table 4: Our default parameter set. We vary c_G, c_{S_d}, c_{R_A} and c_Q to examine the effect of the four different signals; the values employed for these parameters in each simulation are stated in the text. In §4.4 we investigate alternative values of the rates associated with import and export of PhrA.

For protein translation, two rates are adopted in [10], one fast (0.8 sec^{-1}) and one slow (0.014 sec^{-1}). For simplicity we will use 1 sec^{-1} and 0.01 sec^{-1}, respectively (how we choose which rates to be fast and which to be slow is described below). Thus, for mRNA levels at 40 nM, fast production takes place at 40 nM sec^{-1} and slow at 0.4 nM sec^{-1}.

The protein production rates that we will take to be high are $c_{S_F}^{2h}, c_{S_A}^{2h}$ (because *spo0F* and *spo0A* have two promoters each, the second with a higher output) and c_{K_A}. The latter is chosen because the transcription of *kinA* is σ^H-dependent and for *spo0F* and *spo0A* the σ^H-induced transcription is higher than constitutive transcription [35, 41, 64]. The values of c_I, c_{R1} and c_{R3} are taken directly from [10] and we will vary c_{R_A} between 0.4 and 40 nM sec^{-1} (the latter value implies full ComA induction of the *phr/rap* operon), in order to investigate the effect of the competence genes.

The production rates of KinB, c_{K_B}, and Sda, c_{S_d}, are chosen following mathematical investigations under varying nutrient levels and Sda (with $c_Q = 0$ and $c_{R_A} = 0.4$ nM sec^{-1}), see Figure 4. For c_{K_B} sufficiently small (what we believe to be) sporulation-inducing levels of Spo0A∼P are either always reached in the absence of Sda regardless of nutrient levels (Figure 4(a,i)) or never reached with a low level of Sda (Figure 4(a,ii)). On the other hand, increasing both c_{S_d} and c_{K_B} (Figure 4 (b,ii)) does allow a switch from high to low Spo0A∼P levels to occur as we increase c_G. We thus fix c_{K_B} at 10 nM sec^{-1} and take the minimum value of c_{S_d} to be 10 nM sec^{-1} (i.e. a basal level of Sda is present at all times regardless of DNA damage). We then vary c_{S_d} between 10 and 200 nM sec^{-1} (the latter value, as we shall see in §4, allowing for the possibility that damaged DNA can have a negative effect upon sporulation). Similarly c_G will range between 0 and 200 nM sec^{-1}, with c_Q between 0 and 10^4 nM sec^{-1}. While it may be biologically feasible that a cell activates its response to Sda only above a certain threshold level, an alternative and perhaps more likely scenario for the need to take the minimum value of c_{S_d} to be strictly positive is that there is an additional signal involved serving to inhibit KinA phosphorylation (or even transcription) in the absence of suitable sporulation conditions.

3.3 DNA binding and unbinding

Because the binding rates, B_i^j, and the unbinding rates, U_i^j, of proteins to and from their binding sites appear only in the ratio U_i^j/B_i^j (we have assumed protein binding and unbinding is relatively fast, enabling quasi-steady assumptions to be made on the levels of DNA-bound proteins), we need to specify only this combination; the dissociation constant of AbrB and the *spo0E* binding site is approximately 20nM [65, 66], i.e. $U_A^{S_E}/B_A^{S_E} \approx 20$nM, as is that of Hpr and the *sin* operon binding site (U_H^I/B_H^I) [57]. Indeed, in the absence of any other information we will assume that the dissociation constants of all the regulatory proteins and their respective binding sites take this value.

The dissociation constant of a protein from a binding site influences the level of protein required to activate or inhibit the gene in question. Spo0A triggers different responses according to the level at which it is
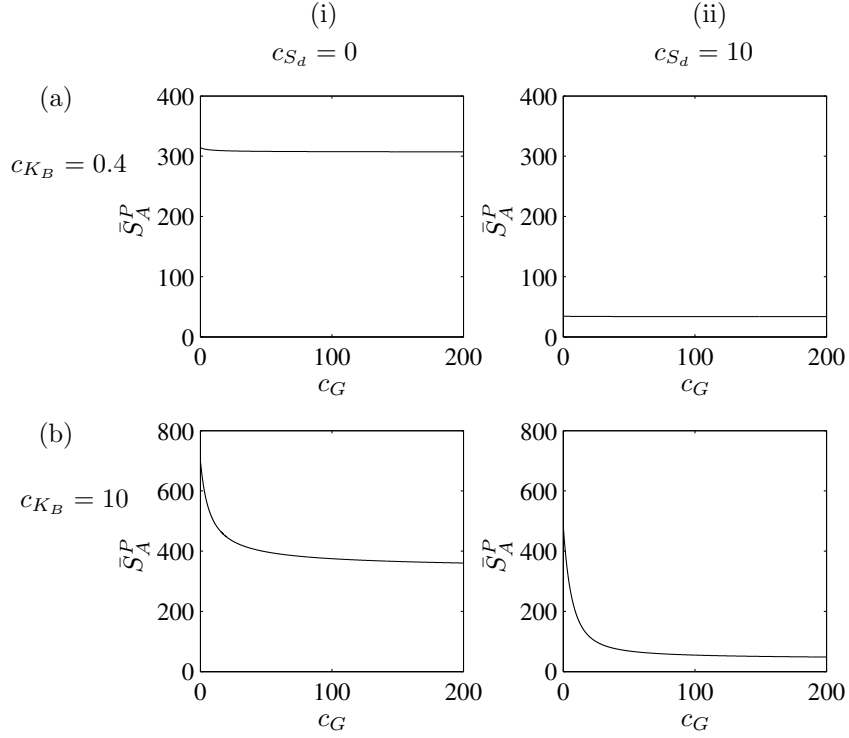
Figure 4: The steady-state curves for S_A^P (units nM), $\bar{S}_A^P$, where in column (i) $c_{S_d} = 0$, column (ii) $c_{S_d} = 10$, row (a) $c_{K_B} = 0.4$ and row (b) $c_{K_B} = 10$ for a solitary ($c_Q = 0$) non-competent ($c_{R_A} = 0.4$) cell, against the nutrient supply, c_G (all rates c_i here and in all subsequent figures have units nM sec^{-1}). We expect high levels of $\bar{S}_A^P$ for low c_G and low levels for high c_G in the absence of any DNA damage (i.e. low c_{S_d}). In row (a) we see that the KinB production rate, c_{K_B}, is too small and a switch in $\bar{S}_A^P$ levels cannot occur, see either (a,i) or (a,ii). In row (b) all conditions can be satisfied as long as a basal level of Sda is permitted (notice that in (b,i) large c_G does not enforce sufficiently low levels of $\bar{S}_A^P$, whereas this does occur in (b,ii)). Given these simulations, we infer that $c_{K_B} = 10$ is a satisfactory choice and that we must have a basal level of Sda production, i.e. $c_{S_d} = 10$ (in reality this could come from an additional signal also interacting with KinA).

present (for example, high levels trigger sporulation while low levels are conducive to biofilm formation [67]), meaning that Spo0A∼P may have different dissociation constants for genes involved in different processes. However, since all the genes considered here form part of the sporulation initiation network, we assume that it is acceptable to use the same dissociation constant for Spo0A and all relevant binding sites; we have undertaken computational investigations that suggest that, while altering the dissociation constants naturally affects the quantitative model behaviour, the same qualitative conclusions result from such more general assumptions.

3.4   Complex separation

There are three rates of disaggregation, i.e. of GTP from CodY, Sda from KinA and RapA from PhrA. Since the first two have been shown to bind we assume that a strong interaction exists. According to [10] the minimum rate of complex fragmentation is 0.1 sec^{-1}, thus we take $\gamma_{K_A^S} = \gamma_{C_Y^G} = 0.1$ sec^{-1}. RapA and PhrA have not yet been shown unequivocally to bind thus we assume their interaction occurs at a more moderate rate and that less energy is required to break the bond between these two proteins, taking $\gamma_{P_A^R} = 1$ sec^{-1} (the maximal rate according to [10] is 1000 sec^{-1}).

3.5   Complex formation

SinI and SinR are taken in [10] to join at a rate $\beta_{I_R} = 0.083$ nM^{-1}sec^{-1} (the desolvation-mediated diffusion-limited rate). We assume this value will serve for all protein-protein complexes. In addition, the dissociation constant of CodY and GTP is in the μM range [68], thus (having fixed $\gamma_{C_Y^G} = 0.1$ sec^{-1} above) we can calculate the estimate $\beta_{C_Y^G} = 10^{-4}$ nM^{-1}sec^{-1}.

3.6   Degradation

SinI (which has a molecular weight of 6.6kDA [59]) is taken in [10] to degrade at rate $\lambda_I = 0.02$ sec^{-1} and SinR (molecular weight 13.0 kDa [59]) at rate $\lambda_R = 0.002$ sec^{-1}. In the absence of other information we assume proteins of comparable sizes will degrade at comparable rates and use the above as a guide for these rates. In Table 5 we list the molecular weights of all the proteins and the degradation rates which we estimate for them. The only exceptions are σ^H and phosphorylated Spo0A, for which information is available on their half-lives [69,70], and are thus able to take $\lambda_{\sigma_H} = 4 \times 10^{-4}$ and $\lambda_{S_A^P} = 10^{-5}$ sec^{-1}. While it is clear that other factors exist, such as the stability of the molecules and the mechanisms of degradation, we believe the estimates in Table 5 to be a satisfactory approximation for this study.

3.7   Dephosphorylation rates

[72] investigates the interaction between Spo0A and its phosphatase Spo0E, these proteins being governed by the reaction Spo0A∼P + Spo0E → Spo0A + Spo0E∼P. Data are presented from which certain

Molecule	Approximate molecular weight, kDa	Estimated degradation rate, sec^{-1}
CodY dimer	60 [71]	0.002
GTP	0.6	0.02
CodY$\sim$GTP	60.6	0.002
Sda	5.6 [17]	0.02
KinA dimer	69 [36]	0.002
KinA$\sim$Sda	75	0.002
KinB	48 [19]	0.002
KinB$\sim$P	48	0.002
Spo0F	14 [46]	0.002
Spo0F$\sim$P	14	0.002
Spo0B dimer	23 [46]	0.002
Spo0B$\sim$P	23	0.002
Spo0A dimer	30 [46]	0.002
Spo0A$\sim$P	30	1×10^{-5} [70]
Spo0E	1 [46, 51]	0.02
Spo0E$\sim$P	1	0.02
σ^H	25 [51]	4×10^{-4} [69]
SinI	6.6 [59]	0.02 [10]
SinR tetramer	52 [59]	0.002 [10]
SinI$\sim$SinR	20	0.002
RapA	45 [28]	0.002
RapA$\sim$P	45	0.002
Internal PhrA	0.5 [28]	0.02
External PhrA	0.5	0.02
Re-internalised PhrA	<0.5	0.02
AbrB tetramer	42 [54]	0.002
Hpr	24 [51, 56]	0.002

Table 5: Molecular weights of the species in the model and the degradation rates that have been chosen for them. These degradation rates, unless we have information about the molecule's half-life, are guided by [10]. When information is available on the degradation rate, the corresponding reference is given in the final column to distinguish rates taken from the literature from estimated ones.

dephosphorylation-reaction rates can be estimated, namely the spontaneous rate $\psi_{S_A^P}$ of Spo0A~P, and $\psi_{S_E^P}$ of Spo0E~P dephosphorylation and the rate δ_{S_E} of Spo0A~P dephosphorylation by Spo0E. By fitting appropriate simple models (i.e. the models only contain the reactions mentioned above) to these data, we have estimated all three to be 4×10^{-4}, with units sec^{-1} for the first two and $\text{nM}^{-1}\text{sec}^{-1}$ for the last (it being a coincidence that the two types of dephosphorylation are assigned the same numerical values). Since the above estimates include both spontaneous and regulatory dephosphorylation, we take all dephosphorylation constants to have this value.

3.8 Phosphorylation rates

In a similar manner to that described in the previous section, $\phi_{K_A^P}^{S_F} = 10^{-6}\ \text{nM}^{-1}\text{sec}^{-1}$ and $\phi_{K_A^P}^{S_A} = 1.7 \times 10^{-8}$ $\text{nM}^{-1}\text{sec}^{-1}$ can be estimated from [39] (for these we have used the dephosphorylation rates calculated in §3.7). We allow for two phosphotransfer rates and from the literature we expect $\phi_{K_A^P}^{S_A}$ and $\phi_{S_F^P}^{K_B}$ to be small, and the others to be relatively large. Thus for $\phi_{K_A^P}^{S_A}$ and $\phi_{S_F^P}^{K_B}$ we use $1.7 \times 10^{-8}\ \text{nM}^{-1}\text{sec}^{-1}$, while we take the remaining phosphorylation rates to be $10^{-6}\ \text{nM}^{-1}\text{sec}^{-1}$.

We anticipate that the autophosphorylation rates be fast as the phosphorylated forms of KinA and KinB should typically be their dominant form. For example, [16] states that the autophosphorylation rate of KinA should be faster than its phosphotransfer rate. Thus we take $\alpha_A = \alpha_B = 0.1\ \text{sec}^{-1}$.

Note that our choices of $\alpha_A, \phi_{K_A^P}^{S_A}, \psi_{S_A^P}$ and $\psi_{K_A^P}$ match the data presented in [16] for KinA phosphorylation of Spo0A.

3.9 PhrA import/export

We have been unable to find any data on the export and import of PhrA that would enable calculation of μ_e, μ_i or θ_H. Thus, for the time being we take the first two to be $1\ \text{sec}^{-1}$ and the third to be $1\ \text{nM}$, but in §4.4 we investigate the effect of varying these parameters and indicate that this is a satisfactory choice.

4 Numerical solutions

4.1 Initial conditions

In Table 1 we listed the possible qualitative combinations of the different signals and how we would expect a cell to respond to each of these. We simulate each of these states in order to see if the model solutions match the corresponding phenotype. Appropriate initial data is to have the cells at their equilibrium state for nutrient rich conditions: we achieve this by first simulating the model with (for simplicity) all initial conditions set to zero and allow the system to evolve to its steady state, with a plentiful nutrient supply, namely

$$c_G = 200 \quad \text{for} \quad 0 \le t < t_S, \tag{38}$$

where $t_S = 4$ hours (which suffices for the solution to approach the required steady state; since the zero initial conditions are being adopted purely as a matter of expedience, the behaviour for $t < t_S$ is not relevant to the investigation). We then remove the nutrient supply, i.e.,

$$c_G = 0 \quad \text{for} \quad t \geq t_S, \tag{39}$$

to see if the cell will stimulate a switch in Spo0A~P levels in response to this starvation.

4.2   No DNA damage, not competent

We begin by examining nutrient levels in an otherwise unstressed cell, i.e. no DNA damage, basal RapA/PhrA production (so the cell is not competent) and either a small or large population. In Figure 5 we illustrate the solution for a selection of variables. The dotted line has $c_Q = 0$ (isolated cell) and the solid line $c_Q = 10^4$ (large population of cells). Remember that c_Q only represents PhrA production from neighbouring cells so that an isolated cell is still able to produce PhrA (at rate c_{R_A}, see (9)). Notice that for most variables the two solutions are indistinguishable, i.e. the presence of other cells makes very little difference. It is evident that, as soon as the nutrient supply is removed, there is a large switch in Spo0A~P levels, implying that sporulation would be activated. Thus the model describes States 1, 2, 9 and 10 appropriately. We will take the Spo0A~P concentration of the order reached in Figure 5 after nutrient supply has been removed to represent a cell in which sporulation will occur.

4.3   Damaged DNA and/or competent

For a single cell, if we introduce ComA (Figure 6(a) where $c_{R_A} = 40$ nM sec^{-1}) or DNA damage (Figure 6(b) where $c_{S_d} = 200$ nM sec^{-1}) or both (Figure 6(c)) the sporulation-associated switch seen in §4.2 and in Figure 5 is prevented (in fact removing nutrients actually lowers the Spo0A~P level), i.e. a solitary cell does not undergo sporulation if it has damaged DNA and/or is competent. Thus the model is displaying the correct phenotype for States 1 through to 10.

For a large population of cells, however, the situation is somewhat different. We begin by considering the competence genes. While quorum sensing in many bacteria, for example *Streptococcus pneumoniae* [73], serves to promote competence (as indeed does the ComQXPA system in *B. subtilis* [29]), our results suggest that for the *phrA* system in *B. subtilis* the opposite occurs. In Figure 7(a) we see how the quorum-sensing system can override the ComA input (ComA incidentally being the response regulator of the ComQXPA quorum-sensing system) to give high levels of Spo0A~P (compare this with Figure 6(a)), thus switching the cells from competence to sporulation (moving from State 11 to 12). Certain components of this sporulation-initiation network also affect genes required for the cell to be competent, so it is not only that high levels of Spo0A~P cause sporulation but, crucially, high levels of Spo0A~P will also inhibit competence [74], for example SinR (which is negatively regulated by Spo0A~P) is required for competence [60], so that sporulation
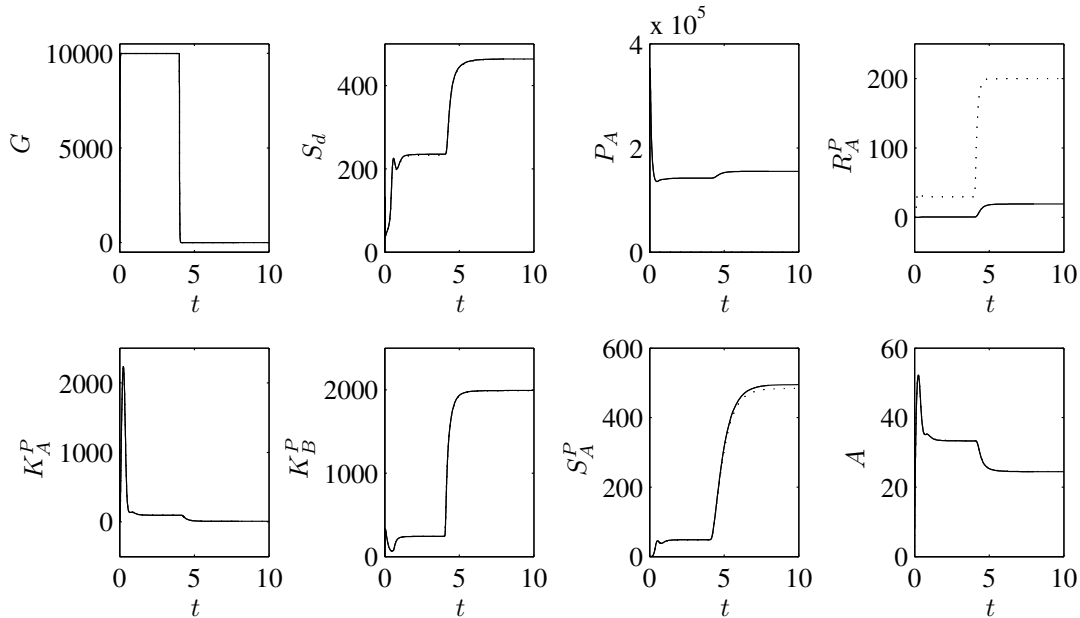
24

Figure 5: Time-dependent solutions (here and henceforth, the units of time are hours) for a selection of concentrations (with units nM) for a solitary cell (dotted line, $c_Q = 0$) and a cell in a large population (solid line, $c_Q = 10^4$). Notice that for G, K_A^P, K_B^P, A and much of the time courses of S_d and S_A^P the two solutions are indistinguishable. The cell is not competent ($c_{R_A} = 0.4$) and has healthy DNA ($c_{S_d} = 10$). The nutrient supply is removed at $t = 4$ hours (i.e. the cell is moving from State 1 to State 2 of Table 1 for the dotted line, and from State 9 to State 10 for the solid line) and this is reflected in the graph of $G(t)$ where GTP (a reflection of nutrient levels) drops to zero. In both the solitary cell and large population cases we see a rapid switch in S_A^P levels, implying that under nutrient limitation the cell would initiate the sporulation process. For most variables the two solutions are indistinguishable, meaning that the size of the population makes little difference in this scenario. In each case, GTP levels drop after $t = 4$, permitting higher levels of KinB (for brevity this is not shown) and thus KinB~P also. This feeds through the phosphorelay to stimulate higher levels of Spo0A~P which represses both AbrB (a negative regulator of sporulation) and KinA (it is for this reason that we have lower levels of KinA~P, and higher Sda, there being less KinA to which Sda can bind). Evidently, there is much more active (re-internalised) PhrA in a cell which is surrounded by other cells secreting PhrA than in an isolated cell and hence there is much less RapA~P in the former scenario.
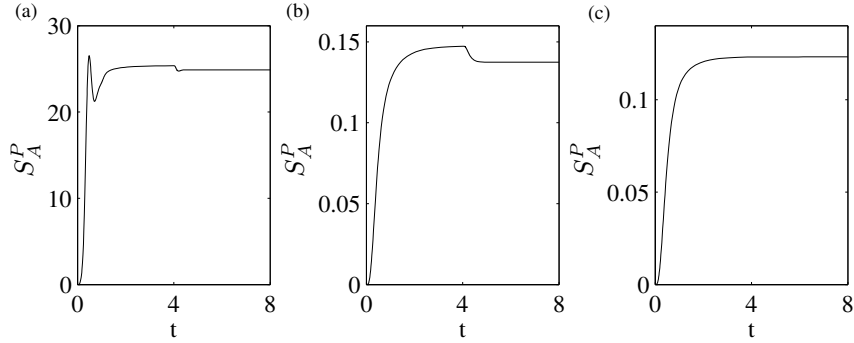
25

Figure 6: Numerical solutions for S_A^P (units nM) when the cell is assumed to be isolated, i.e. $c_Q = 0$, and has nutrients removed at time $t_S = 4$ hours ($c_G = 200$ for $0 \leq t < t_S$, and $c_G = 0$ for $t_S \leq t < 8$) (a) when the cell is competent: $c_{R_A} = 40, c_{S_d} = 10$ (i.e. the cell switches from State 3 to 4); (b) when the cell has damaged DNA: $c_{R_A} = 0.4, c_{S_d} = 200$ (i.e. the cell switches from State 5 to 6); and (c) when the cell is both competent and has damaged DNA: $c_{R_A} = 40$ and $c_{S_d} = 200$ (State 7 to 8). In each case there is no increase in Spo0A∼P after $t = 4$ and hence sporulation is prevented from occurring.

in effect takes over from competence (we recall that the two processes cannot occur simultaneously).

The following biological hypotheses might explain why the cells have evolved this mechanism by which they can vary their behaviour according to population size in this instance. Detecting population size effectively allows a cell to estimate the future nutrient supply and gauge the necessity of forming a spore. The more cells there are the faster nutrients will be depleted and the more urgent it becomes that, if they are to increase their chances of survival, at least a subset of the cells must form spores. Thus competence is replaced by sporulation. This is consistent with [11] whereby it is proposed that the quorum-sensing system is in place to titrate the amount of "food per cell", i.e. a cell must average the nutrient level over population size in order to predict whether sporulation would be advantageous.

An alternative explanation for the difference in behaviour between a solitary cell and that of a large population centres on competition with neighbouring cells: sporulation is high-risk for a single cell because if it turns out that the environment is in fact suitable for vegetative growth, it will be out-competed by any neighbouring cells which have remained vegetative. On the other hand, if the cell is in a large population it is acceptable to risk 'sacrificing' a certain number of cells to the fate of sporulation (in general, sporulation is not undertaken by every cell in a population, which could be attributed to the spatial inhomogeneity of a signal molecule such as PhrA) as at worst the population will have only lost a portion of its cells if a new nutrient supply is discovered and, at best, if the vegetative cells fail to survive, the population has spores which are not only more likely to survive, but after germination are also likely to face much reduced competition for resources. Since sporulation is higher risk for the single cell, under some conditions which would lead to sporulation in a large population, the single cell effectively defers this decision by continuing
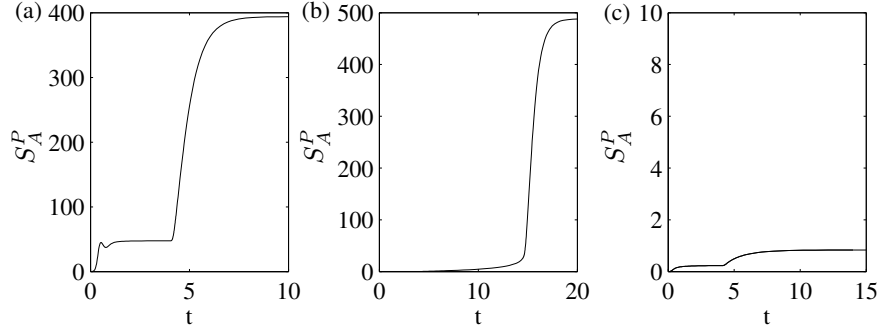
Figure 7: Numerical solutions for S_A^P (units nM) when the cell is in a large population, i.e. $c_Q = 10^4$, and has nutrient supply removed at time $t = 4$ hours. In (a) the cell is assumed to be competent (so $c_{R_A} = 40, c_{S_d} = 10$) and we see that the quorum-sensing system will override that of competence to push the cells into a state capable of entering sporulation. In (b) the cell is not competent but has damaged DNA ($c_{R_A} = 0.4, c_{S_d} = 200$) and sporulation occurs but after a significant time delay which may allow the cell time to repair its DNA before it is replicated in the spore. Finally in (c) the cell is both competent and has damaged DNA ($c_{R_A} = 40, c_{S_d} = 200$) and this combination prevents sporulation from occurring. However, in §4.4 we will see that this can be reversed for an alternative parameter set.

vegetative growth (whereby it can still form a spore at a later timepoint if conditions alter). Thus, here PhrA serves as a communication signal between cells leading to different phenotypes for different population sizes.

In Figure 7(b) we illustrate what can arise if a cell is in a large population, is not competent but has damaged DNA and the nutrients are removed at time $t = 4$ hours: the damaged DNA can no longer prevent the switch in Spo0A~P levels, but the PhrA from neighbouring cells causes a time delay so that high levels of Spo0A~P are reached significantly after the nutrients are removed (initially insufficient internal PhrA will be present to inhibit the sporulation-antagonist RapA, but eventually enough will have entered the cell to counteract RapA; this also indicates that it must be the combination of RapA and Sda antagonising the sporulation pathway, even at low RapA levels), potentially allowing time for the cell to repair its DNA before entering the sporulation process. It would be illuminating to test experimentally whether an overexpressed *sda* gene could enforce delayed sporulation in a population of *B. subtilis* cells. Increasing c_Q decreases the time delay (this is logical since the more cells there are the quicker nutrient supplies will be used up and the more urgent it will be for the cells to enter sporulation), which tends towards the limit displayed in Figure 7(b) (i.e. the switch can occur no earlier than that shown there). Incidentally reducing $\gamma_{P_A^R}$ (which represents the strength, or indeed weakness, of PhrA's ability to act on RapA) will also reduce the time delay, but a significant lag between nutrient removal and the attainment of high Spo0A~P levels remains. Thus we demonstrate that, with identical parameter sets (c_Q aside), the PhrA intercellular-signalling process

27

can be a means both to detect population size and to control the timing of sporulation initiation.

Finally, in Figure 7(c) we see what happens for our default parameter set when nutrients are removed from a competent cell with damaged DNA: the cumulative effect of these two sporulation-negative signals is overwhelming and the cell cannot enforce a suitable switch in Spo0A∼P levels. However, in the following subsection, we will see that altering the transport parameters, μ_i and θ_H, can enable the cell to sporulate (albeit with a time delay) even when the cell begins in a competent state with damaged DNA.

4.4   Transport-parameter sensitivity studies

In §3.2-3.8 we were able to calculate what we believe to be suitable estimates for all of the parameters bar three, namely μ_e, μ_i and θ_H, all of which play a role in the transport process of PhrA, the putative quorum-sensing signal. We remark that the cells are assumed to be well-mixed and thus there is no spatial dependency in the model, meaning that the 'transport' process refers solely to the import and export of this protein from the cell. Solving the system for different μ_e, we find that changing this export parameter has little effect upon the outcome: while minor quantitative changes occur, the qualitative behaviour is insensitive to the value of μ_e (data not shown).

For μ_i (the rate of PhrA import) and θ_H (concentration of Hpr at which internalisation of PhrA is half the maximal rate), however, the situation is somewhat different. Although the phenotype cannot be altered when the cell is assumed to be in isolation, i.e. $c_Q = 0$, and thus import of PhrA is less significant (data not shown), when the cell is in a large population ($c_Q = 10^4$), setting μ_i or θ_H sufficiently small prevents the cell from undergoing sporulation, either when the cell is competent (Figure 8 for example) or has damaged DNA (see Figure 9). On the other hand, sufficiently large μ_i or θ_H give rise to sporulation in a cell which is both competent and has damaged DNA (remember that in §4.3 with $\mu_i = \theta_H = 1$ sporulation was prevented in this scenario), again with the time delay that we saw for the damaged DNA and non-competent case for our default parameter set; see Figure 10.

Biologically it is plausible that these parameters are sufficiently large that the latter behaviour occurs, i.e. a suitably large population could override the competence system and allow a cell to undergo sporulation after a time delay during which the cell could attempt to repair its DNA. Conversely, it is much less likely that either θ_H or μ_i will in practice be small enough that a cell in effect ignores 'communication' from its neighbours. Thus we deem our choice of transportation parameters acceptable, with the possible exception that θ_H or μ_i could be increased to allow a cell in a large population to undergo sporulation despite being both competent and having damaged DNA.

5   Discussion

The model reproduces all the phenotypes displayed in Table 1 bar one, namely a cell in a large population with damaged DNA. We anticipated that in this scenario the cell would not undergo sporulation since the
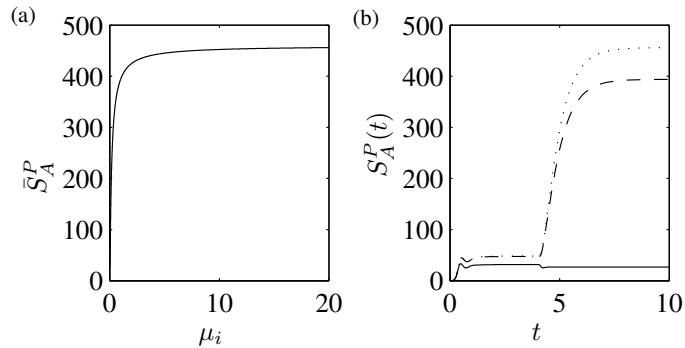
Figure 8: (a) The solution curve for varying μ_i (units sec^{-1}) when $c_G = 0, c_Q = 10^4, c_{R_A} = 40, c_{S_d} = 10$ and all other parameters are taken from the default set in Table 4 (i.e. the cell is competent with healthy DNA and in a large population). Under this parameter set we would expect sporulation to be desirable (this is State 12 from Table 1) and we see that high levels of S_A^P (units nM) will be attained (which is indicative of sporulation) unless μ_i is sufficiently small. In (b) we illustrate some time-dependent solutions (with units of hours) to reflect this behaviour: the solid line has $\mu_i = 0.01$ and sporulation does not arise, whereas for the dashed ($\mu_i = 1$) and dotted ($\mu_i = 20$) lines, a switch in S_A^P occurs (nutrient supply is removed at time $t = 4$ hours). Qualitatively, altering θ_H under these conditions has the same effect as altering μ_i.

damaged DNA would be replicated in the spore, causing permanent harm. However, our model suggests that the sporulation mechanisms would still be activated if nutrient supply were to be removed but, crucially, after a significant time delay which would provide the cell with time to attempt to repair its DNA before the first steps of spore formation. Thus the PhrA protein is able to act not only as a quorum-sensing signal molecule permitting communication between neighbouring cells but also as a timing device, delaying the onset of sporulation. This possibility has been noted in the literature [13,14] and our model provides support for its feasibility and the circumstances under which it could arise.

It is intriguing that production of this quorum-sensing signal molecule (a sporulation activator) is activated by ComA, a protein that should be present when the cell is competent (we recall that sporulation and competence are incompatible). It is perhaps possible that ComA activates transcription of both PhrA and RapA (a negative regulator of sporulation) in order to heighten the sensitivity of the receiver part of the quorum-sensing system and make the best decision regarding whether to be competent or to form a spore. If the cell is in a small population and remains competent instead of forming a spore, excess PhrA will have been produced at a cost to the cell. However, the payoff from choosing sporulation over competence if the cell is part of a large population (with quicker nutrient depletion) is likely to be higher than this cost.

The model also sheds light on the differences between small and large populations of cells, given that making the wrong decision regarding sporulation is catastrophic for an individual cell. One or a few cells
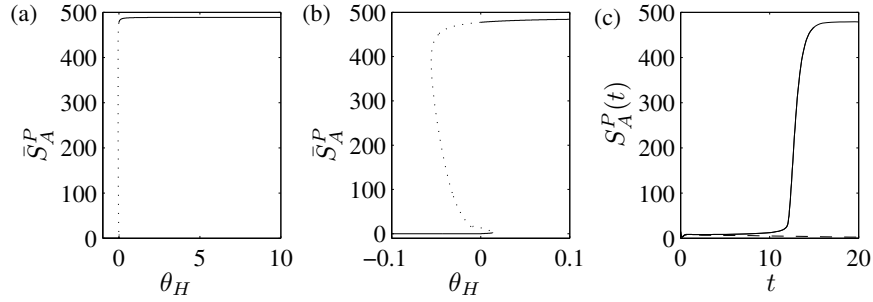
Figure 9: (a) and (b) The solution curve of S_A^P (units nM) for varying θ_H (units nM) when $c_G = 0, c_Q = 10^4, c_{R_A} = 0.4, c_{S_d} = 200$, with all other parameters taken from the default set in Table 4 (i.e. the cell has damaged DNA and is not competent), (b) showing a blow up part of (a). The solid lines are stable steady states and the dotted line an unstable one. The system has a small area of bistability where it can switch between attaining either a state in which sporulation is repressed or one in which it is activated. Note that part of the curve in $\theta_H < 0$ is unphysical. We remark that the change in stability on the upper branch of the hysteresis curve corresponds to a Hopf bifurcation but given that this occurs at $\theta_H < 0$ we disregard this behaviour and do not track it on the diagrams. In (c) we plot some time-dependent solutions with $\theta_H = 0.01$ (and $c_G = 0$ throughout the time course) to illustrate the bistable behaviour which can result from this system. Default initial conditions are maintained for all variables except S_A^P for which we use values close to the unstable equilibrium state, namely eleven (leading to the lower branch) and twelve (leading, after a time lag, to the upper branch). For values of θ_H outside of the bistable region only one steady state can arise: if θ_H is sufficiently large the cell will sporulate and if sufficiently small it remains vegetative. Qualitatively, the dependence on μ_i is similar to that in (a).
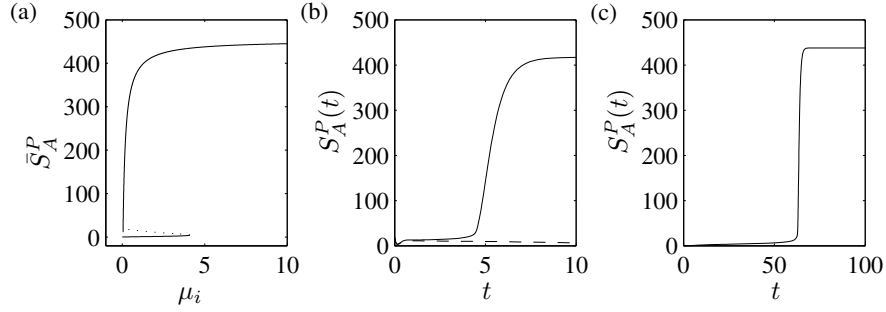
Figure 10: In all three graphs $c_Q = 10^4, c_{R_A} = 40, c_{S_d} = 200$, with all other parameters taken from the default set in Table 4 (i.e. the cell is competent and has damaged DNA). (a) depicts the solution curve of S_A^P (units nM), $\bar{S}_A^P$, for varying μ_i (units sec^{-1}) when there is no nutrient supply ($c_G = 0$). Solid lines illustrate stable steady states and the dotted line an unstable one. The system is bistable for a range of μ_i and for small μ_i may not be able to stimulate sporulation (depending upon the initial conditions of the system). However, increasing μ_i sufficiently will enable the cell to enter the sporulation process in spite of the opposing signals. Qualitatively, the dependence on θ_H is similar. In (b) we illustrate the bistable behaviour with two time-dependent solutions where $\mu_i = 2$ (and $c_G = 0$ throughout the time course). Default initial conditions are maintained for all variables except S_A^P for which we use values close to the unstable equilibrium state, namely eighteen (leading to the lower branch) and nineteen (leading, after a time lag, to the upper branch). Finally in (c) we show the numerical solution when $\mu_i = 5$, the initial conditions for all variables in the model are zero and the nutrient supply is removed at time $t = 4$ hours. The result is that high levels of Spo0A~P are reached but, as we saw in Figure 7(b), only after a substantial delay from when the nutrient supply is removed.

pursuing an inappropriate fate will have a relatively greater impact on a small population than a large one. Logically, sporulation is the higher-risk option. This is reflected by our results, whereby in circumstances in which the benefits that accrue are less clear-cut (for example when the cell is already competent) sporulation is more likely to manifest itself if the cell is part of a large population. Thus the cells plausibly incorporate a quorum-sensing system into their sporulation-initiation network in order to monitor the risks associated with such an irreversible switch.

6 Summary

We have presented what we believe to be the most detailed model to date of the gene regulation network governing sporulation initiation in *B. subtilis*. Having identified a suitable parameter set, we have been able to demonstrate that the model displays the correct phenotypic behaviour under various environmental and cellular conditions, i.e. nutrient supply, population size, whether or not the cell is competent and whether it has healthy or damaged DNA. Of most interest is the quorum-sensing signal molecule, PhrA, which, due to its export and import, can combine two roles, enabling communication between cells whilst also delaying the first steps of spore formation until conditions are appropriate. Additionally, our model requires that a basal level of *sda* expression must exist in order to prevent a cell sporulating in the presence of nutrients (Figure 4). To our knowledge, it is not known whether this occurs *in vivo* and it would be interesting to test whether this threshold level of expression does indeed exist.

Gaining intuition into the sporulation-initiation network in Figure 2 without the aid of a mathematical model would be challenging given the number of ways in which the four signals can interact to determine the fate of a cell. For example, if one were to anticipate the temporal regulation provided by the PhrA protein, it would be reasonable to assume that this might occur in relation to the competence genes given the link between *phr* transcription and ComA, rather than to the *sda* gene representing DNA damage. However, our mathematical model provides a relatively straightforward means by which to examine these interactions and develop hypotheses concerning the roles of various proteins within the network. These hypotheses could be examined by constructing strains in which the expression (and its timing) of one or more relevant genes is experimentally controlled. Other parameters could also be varied *in vitro*, such as nutrient availability in the culture medium, or the PhrA signal by the addition of exogenous PhrA pentapeptide or its precursor.

While many other signals, such as the pH or temperature of the environment, will also have a bearing on a cell's decision to enter sporulation, we believe that our model as it stands (which includes environmental, metabolic and cell cycle signals) is useful in predicting cell behaviour and elucidating some of the reasons behind why such a complicated and intricate network has evolved to govern sporulation initiation in *B. subtilis*. It also provides a base on which to add additional signals (such as those mentioned above) or components of the network, in order to attempt to further our understanding of when and why a *B. subtilis* cell will embark upon the irreversible process of sporulation. For example, there are multiple Phr/Rap proteins

involved in the gene regulation network [15] and further quorum-sensing networks. Indeed, competence induction is itself controlled by the ComX quorum-sensing signal molecule [29] and it would be interesting to model the combined effect of these two quorum-sensing networks. Given the low numbers of molecules which may be present in a single cell, it would also be worthwhile to consider the influence of stochastic effects upon cell behaviour. Additionally, the model can be adapted to fit the sporulation-initiation networks of other bacteria. For instance, there is a great deal of overlap between the network presented here and the corresponding network in the *Clostridium* species [1]. Thus this study may aid in the understanding of when and why either the pathogenic *C. difficile* or the solvent-producing *C. acetobutylicum* form spores. Similarly the model can be easily built upon to represent a population of cells in which a spatial dimension is incorporated; a study of the potentially inhomogeneous distribution of signal molecules such as PhrA may help to explain why only a number of cells within a population sporulate.

Acknowledgements

References

[1] Paredes C.J., Alsaker K.V., Papoutsakis E.T. (2005) A comparative genomic view of clostridial sporulation and physiology. *Nature Rev. Microbiol.,* 3, 969–978.

[2] Sonenshein A.L. (2000) Control of sporulation initiation in *Bacillus subtilis*. *Curr. Opin. Microbiol.,* 3, 561–566.

[3] Wilson M., McNab R., Henderson B. (2002) Bacterial disease mechanisms, an introduction to cellular microbiology, Cambridge University Press, .

[4] Piggot P.J., Coote J.G. (1976) Genetic aspects of bacterial endospore formation. *Bacteriol. Rev.,* 40, 908–962.

[5] Eichenberger P., Fujita M., Jensen S.T., Conlon E.M., Rudner D.Z., Wang S.T., Ferguson C., Haga K., Sato T., Liu J.S., Losick R. (2004) The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol.,* 2, 1664–1683.

[6] Piggot P.J., Hilbert D.W. (2004) Sporulation of *Bacillus subtilis*. *Curr. Opin. Microbiol.,* 7, 579–586.

[7] Jones D.T., Woods D.R. (1986) Acetone-butanol fermentation revisited. *Microbiol. Rev.,* 50, 484–524.

[8] Jeong J.W., Snay J., Ataai M.M. (1990) A mathematical model for examining growth and sporulation processes of *Bacillus subtilis*. *Biotechnol. Bioeng.,* 35, 160–184.

[9] De Jong H., Geiselmann J., Batt G., Hernandez C., Page M. (2004) Qualitative simulation of the initiation of sporulation in *Bacillus subtilis*. *Bull. Math. Biol.*, 66, 261–299.

[10] Voigt C.A., Wolf D.M., Arkin A.P. (2005) The *Bacillus subtilis sin* operon: an evolvable network motif. *Genetics*, 169, 1187–1202.

[11] Bischofs I.B., Hug J.A., Liu A.W., Wolf D.M., Arkin A.P. (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpoplation signaling in the *Bacillus subtilis* phosphorelay. *PNAS*, 106, 6459–6464.

[12] Perego M., Hoch J.A. (1996) Cell-cell communication regulates the effects of protein aspartate phosphatases on the phosphorelay controlling development in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA*, 93, 1549–1553.

[13] Perego M. (1997) A peptide export-import control circuit modulating bacterial development regulates protein phosphatases of the phosphorelay. *Proc. Natl. Acad. Sci. USA*, 94, 8612–8617.

[14] Perego M., Brannigan J.A. (2001) Pentapeptide regulation of aspartyl-phosphate phosphatases. *Peptides*, 22, 1541–1547.

[15] Perego M. (1999) Self-signalling by phr peptides modulates *Bacillus subtilis* development. In Dunny G.M., Winans S.C., (ed.), *Cell-cell signaling in bacteria*, pp. 243–258 American Society for Microbiology.

[16] Jiang M., Shao W., Perego M., Hoch J.A. (2000) Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol. Microbiol.*, 38, 535–542.

[17] Rowland S.L., Burkholder W.F., Cunningham K.A., Maciejewski M.W., Grossman A.D., King G.F. (2004) Structure and mechanism of action of Sda, an inhibitor of the histidine kinases that regulate initiation of sporulation in *Bacillus subtilis*. *Mol. Cell*, 13, 689–701.

[18] Molle V., Nakaura Y., Shivers R.P., Yamaguchi H., Losick R., Fujita Y., Sonenshein A.L. (2003) Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript anlaysis. *J. Bacteriol.*, 185, 1911–1922.

[19] Trach K.A., Hoch J.A. (1993) Multisensory activation of the phosphorelay initiating sporulation in *Bacillus subtilis*: identification and sequence of the protein kinase of the alternate pathway. *Mol. Microbiol.*, 8, 69–79.

[20] Fujita M. (2000) Temporal and selective association of multiple sigma factors with RNA polymerase during sporulation in *Bacillus subtilis*. *Genes Cells*, 5, 79–88.

[21] Millat T., Bullinger E., Rohwer J., Wolkenhauer O. (2007) Approximations and their consequences for dynamic modelling of signal transduction pathways. *Math. Biosci.*, 207, 40–57.

[22] Ireton K., Rudner D.Z., Siranosian K.J., Grossman A.D. (1993) Integration of multiple developmental signals in *Bacillus subtilis* through the Spo0A transcription factor. *Genes and Dev.*, 7, 283–294.

[23] Ratnayake-Lecamwasam M., Serror P., Wong K., Sonenshein A.L. (2001) *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. *Genes and Dev.,* 15, 1093–1103.

[24] Burkholder W.F., Kurtser I., Grossman A.D. (2001) Replication initiation proteins regulate a developmental checkpoint in *Bacillus subtilis. Cell,* 104, 269–279.

[25] Perego M., Hanstein C., Welsh K.M., Djavakhishvili T., Glaser P., Hoch J.A. (1994) Multiple protein-aspartate phosphatases provide a mechanism for the integration of diverse signals in the control of development in *Bacillus subtilis. Cell,* 79, 1047–1055.

[26] Bongiorni C., Ishikawa S., Stephenson S., Ogasawara N., Perego M. (2005) Synergistic regulation of competence development in *Bacillus subtilis* by two Rap-Phr systems. *J. Bacteriol.,* 187, 4353–4361.

[27] Auchtung J.M., Lee C.A., Grossman A.D. (2006) Modulation of the ComA-dependent quorum response in *Bacillus subtilis* by multiple Rap proteins and Phr peptides. *J. Bacteriol.,* 188, 5273–5285.

[28] Mueller J.P., Bukusoglu G., Sonenshein A.L. (1992) Transcriptional regulation of *Bacillus subtilis* glucose starvation-inducible genes: control of *gsiA* by the ComA-ComP signal transduction system. *J. Bacteriol.,* 174, 4361–4373.

[29] Lazazzera B.A., Palmer T., Quisel J., Grossman A.D. (1999) Cell density control of gene expression and development in *Bacillus subtilis.* In Dunny G.M., Winans S.C., (ed.), *Cell-cell signaling in bacteria,* pp. 27–46 American Society for Microbiology.

[30] Süel G.M., Garcia-Ojalvo J., Liberman L.M., Elowitz M.B. (2006) An excitable gene regulatory circuit induces transient cellular differentiation. *Nature,* 440, 545–550.

[31] Süel G.M., Kulkarni R.P., Dworkin J., Garcia-Ojalvo J., Elowitz M.B. (2007) Tunability and noise dependence in differentiation dynamics. *Science,* 315, 1716–1719.

[32] Jiang M., Grau R., Perego M. (2000) Differential processing of propeptide inhibitors of Rap phosphatases in *Bacillus subtilis. J. Bacteriol.,* 182, 303–310.

[33] Predich M., Nair G., Smith I. (1992) *Bacillus subtilis* early sporulation genes *kinA*, *spo0F*, and *spo0A* are transcribed by the RNA polymerase containing σ^H. *J. Bacteriol.,* 174, 2771–2778.

[34] Molle V., Fujita M., Jensen S.T., Eichenberger P., González-Pastor, Liu J.S., Losick R. (2003) The Spo0A regulon of *Bacillus subtilis. Mol. Microbiol.,* 50, 1683–1701.

[35] Asai K., Kawamura F., Yoshikawa H., Takahashi H. (1995) Expression of *kinA* and accumulation of σ^H at the onset of sporulation in *Bacillus subtilis. J. Bacteriol.,* 177, 6679–6683.

[36] Wang L., Fabret C., Kanamaru K., Stephenson K., Dartois V., Perego M., Hoch J.A. (2001) Dissection of the functional and structural domains of phosphorelay histidine kinase A of *Bacillus subtilis. J. Bacteriol.,* 183, 2795–2802.

[37] Dartois V., Djavakhishvili T., Hoch J.A. (1996) Identification of a membrane protein involved in activation of the KinB pathway to sporulation in *Bacillus subtilis. J. Bacteriol.,* 178, 1178–1186.

[38] Strauch M.A. (1995) Delineation of AbrB-binding sites on the *Bacillus subtilis spo0H, kinB, ftsAZ,* and *pbpE* promoters and use of a derived homology to identify a previously unsuspected binding site in the *bsuB1* methylase promoter. *J. Bacteriol.,* 177, 6999–7002.

[39] Burbulys D., Trach K.A., Hoch J.A. (1991) Initiation of sporulation in *Bacillus subtilis* is controlled by a multicomponent phosphorelay. *Cell,* 64, 545–552.

[40] Chibazakura T., Kawamura F., Takahashi H. (1991) Differential regulation of *spo0A* transcription in *Bacillus subtilis*: glucose epresses promoter switching at the initiation of sporulation. *J. Bacteriol.,* 173, 2625–2632.

[41] Lewandoski M., Dubnau E., Smith I. (1986) Transcriptional regulation of the *spo0F* gene of *Bacillus subtilis*. *J. Bacteriol.,* 168, 870–877.

[42] Bai U., Lewandowski M., Dubnau E., Smith I. (1990) Temporal regulation of the *Bacillus subtilis* early sporulation gene *spo0F*. *J. Bacteriol.,* 172, 5432–5439.

[43] Chibazakura T., Kawamura F., Asai K., Takahashi H. (1995) Effects of *spo0* mutations on *spo0A* promoter switching at the initiation of sporulation in *Bacillus subtilis*. *J. Bacteriol.,* 177, 4520–4523.

[44] Strauch M.A., Trach K.A., Day J., Hoch J.A (1992) Spo0A activates and represses its own synthesis by binding at its dual promoters. *Biochimie,* 74, 619–626.

[45] Mandic-Mulec I., Doukhan L., Smith I. (1995) The *Bacillus subtilis* SinR protein is a repressor of the key sporulation gene *spo0A*. *J. Bacteriol.,* 177, 4619–4627.

[46] Perego M., Hoch J.A. (1987) Isolation and sequence of the *spo0E* gene: its role in initiation of sporulation in *Bacillus subtilis*. *Mol. Microbiol.,* 1, 125–132.

[47] Varughese K.I., Madhusudan, Zhou X.Z., Whiteley J.M., Hoch J.A. (1998) Formation of a novel four-helix bundle and molecular recognition sites by dimerization of a response regulator phosphotransferase. *Mol. Cell,* 2, 485–493.

[48] Ohlsen K.L., Grimsley J.K., Hoch J.A. (1994) Deactivation of the sporulation transcription factor Spo0A by the Spo0E protein phosphatase. *Proc. Natl. Acad. Sci. USA,* 91, 1756–1760.

[49] Perego M., Hoch J.A. (1991) Negative regulation of *Bacillus subtilis* sporulation by the *spo0E* gene product. *J. Bacteriol.,* 173, 2514–2520.

[50] Lewis R.J., Scott D.J., Brannigan J.A., Ladds J.C., Cervin M.A., Spiegelman G.B., Hoggett J.G., Barák I., Wilkinson A.J. (2002) Dimer formation and transcription activation in the sporulation response regulator Spo0A. *J. Mol. Biol.,* 316, 235–245.

[51] Hoch J.A. (1993) Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu. Rev. Microbiol.,* 47, 441–465.

[52] Strauch M.A., Perego M. Burbulys D., Hoch J.A. (1989) The transition state transcription regulator AbrB of *Bacillus subtilis* is autoregulated during vegetative growth. *Mol. Microbiol.,* 3, 1203–1209.

[53] Strauch M. Webb V., Spiegelman G., Hoch J.A. (1990) The Spo0A protein of *Bacillus subtilis* is a repressor of the *abrB* gene. *Proc. Natl. Acad. Sci. USA,* 87, 1801–1805.

[54] Benson L.M., Vaughn J.L., Strauch M.A., Bobay B.G., Thompson R., Naylor S., Cavanagh J. (2002) Macromolecular assembly of the transition state regulator AbrB in its unbound and complexed states probed by microelectrospray ionization mass spectrometry. *Anal. Biochem.,* 306, 222–227.

[55] Strauch M.A., Hoch J.A. (1993) Transition-state regulators: sentinels of *Bacillus subtilis* post-exponential gene expression. *Mol. Microbiol.,* 7, 337–342.

[56] Kallio P.T., Fagelson J.E., Hoch J.A., Strauch M.A. (1991) The transition state regulator Hpr of *Bacillus subtilis* is a DNA-binding protein. *J. Biol. Chem.,* 266, 13411–13417.

[57] Shafikhani S.H., Mandi-Mulec I., Strauch M.A., Smith I., Leighton T. (2002) Postexponential regulation of *sin* operon expression in *Bacillus subtilis*. *J. Bacteriol.,* 184, 564–571.

[58] Gaur N. K., Oppenheim J., Smith I. (1991) The *Bacillus subtilis sin* gene, a regulator of alternate developmental processes, codes for a DNA-binding protein. *J. Bacteriol.,* 173, 678–686.

[59] Scott D.J., Leejeerajumnean S., Brannigan J.A., Lewis R.J., Wilkinson A.J., Hoggett J.G. (1999) Quaternary re-arrangement analysed by spectral enhancement: the interaction of a sporulation repressor with its antagonist. *J. Mol. Biol.,* 293, 997–1004.

[60] Bai U., Mandic-Mulec I., Smith I. (1993) SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein-protein interaction. *Genes and Dev.,* 7, 139–148.

[61] Lewis R.J., Brannigan J.A., Smith I., Wilkinson A.J. (1996) Crystallisation of the *Bacillus subtilis* sporulation inhibitor SinR, complexed with its antagonist, SinI. *FEBS Lett.,* 378, 98–100.

[62] Weir J., Predich M., Dubnau E., Mair G., Smith I. (1991) Regulation of *spo0H*, a gene coding for the *Bacillus subtilis* σ^H factor. *J. Bacteriol.,* 173, 521–529.

[63] Vilar J.M.G., Kueh H.Y., Barkai N., Leibler S. (2002) Mechanisms of noise-resistance in genetic oscillators. *PNAS,* 99, 5988–5992.

[64] Strauch M.A., Wu J., Jonas R.H., Hoch J.A. (1993) A positive feedback loop controls transcription of the *spo0F* gene, a component of the sporulation phosphorelay in *Bacillus subtilis*. *Mol. Microbiol.,* 7, 967–974.

[65] Xu K., Strauch M.A. (1996) *In vitro* selection of optimal AbrB-binding sites: comparison to known *in vivo* sites indicates flexibility in AbrB binding and recognition of three-dimensional DNA structures. *Mol. Microbiol.,* 19, 145–158.

[66] Vaughn J.L., Feher V., Naylor S., Strauch M.A., Cavanagh J. (2000) Novel DNA binding domain and genetic regulation model of *Bacillus subtilis* transition state regulator AbrB. *Nat. Struct. Biol.,* 7, 1139–1146.

[67] González-Pastor J.E. (2007) Multicellularity and social behaviour in *Bacillus subtilis*. In Graumann P., (ed.), *Bacillus: cellular and molecular biology*, pp. 419–449 Caister Academic Press.

[68] Blagova E.V., Levdikov V.M., Tachikawa K., Sonenshein A.L., Wilkinson A.J. (2003) Crystallization of the GTP-dependent transcriptional regulator CodY from *Bacillus subtilis*. *Acta Cryst.*, D59, 155–157.

[69] Healy J., Weir J., Smith I., Losick R. (1991) Post-transcriptional control of a sporulation regulatory gene encoding transcription factor σ^H in *Bacillus subtilis*. *Mol. Microbiol.*, 5, 477–487.

[70] Ladds J.C., Muchová K., Blaškovič D., Lewis R.J., Brannigan J.A., Wilkinson A.J., Barák I. (2003) The response regulator Spo0A from *Bacillus subtilis* is efficiently phosphorylated in *Escherichia coli*. *FEMS Microbiol. Lett.*, 223, 153–157.

[71] Levdikov V.M., Blagova E., Joseph P., Sonenshein A.L., Wilkinson A.J. (2006) The structure of CodY, a GTP- and isoleucine-responsive regulator of stationary phase and virulence in Gram-positive bacteria. *J. Biol. Chem.*, 281, 11366–11373.

[72] Stephenson S.J., Perego M. (2002) Interaction surface of the Spo0A response regulator with the Spo0E phosphatase. *Mol. Microbiol.*, 44, 1455–1467.

[73] Dunny G.M., Leonard B.A.B. (1997) Cell-cell communication in Gram-positive bacteria. *Annu. Rev. Microbiol.*, 51, 527–64.

[74] Smits W.K., Bongiorni C., Veening J., Hamoen L.W., Kuipers O.P., Perego M. (2007) Temporal separation of distinct differentiation pathways by a dual specificity Rap-Phr system in *Bacillus subtilis*. *Mol. Microbiol.*, 65, 103–120.

[75] Hicks K.A., Grossman A.D. (1996) Altering the level and regulation of the major sigma subunit of RNA polymerase affects gene expression and development in *Bacillus subtilis*. *Mol. Microbiol.*, 20, 201–212.